# Structural and statistical aspects in joint modelling of artesunate pharmacometrics and malarial parasite lifecycle

by

**Adam James Hall**

Submitted for the degree of

**Doctor of Philosophy**

in

**Mathematics and Statistics**

at the University of Warwick

Mathematics and Statistics Centre for Doctoral Training

September 2015

# Contents

**Contents** i

**List of Tables** v

**List of Figures** vi

**Acknowledgements** xi

**Declaration** xiii

**Abstract** xiv

**1. Introduction** 1

1.1. Background . . . . . 1

1.2. Aims and objectives . . . . . 4

1.3. Structure of this thesis . . . . . 5

**2. Background and literature review** 7

2.1. Malaria . . . . . 7

2.2. Artemisinins . . . . . 10

2.3. Data . . . . . 12

2.3.1. Mahidol_PK dataset . . . . . 13

2.3.2. In vitro uptake data . . . . . 15

2.3.3. MORU_ARC3_PD dataset . . . . . 18

2.4. Modelling techniques . . . . . 19

2.4.1. Compartmental models . . . . . 20

2.4.2. Non-compartmental models . . . . . 23
2.5. Structural identifiability . . . . . 26
2.5.1. Overview and importance . . . . . 26
2.5.2. Definitions . . . . . 28
2.5.3. Example to illustrate definitions . . . . . 34
2.5.4. Abridged summary of available methods . . . . . 37
2.5.5. Taylor series coefficient approach . . . . . 38
2.5.6. Non-structural identifiability . . . . . 39
2.6. Existing models for ARS . . . . . 42
2.6.1. Four compartment pharmacokinetic model . . . . 45
2.6.1.1. Model structure . . . . . 45
2.6.1.2. Model analysis . . . . . 49
2.7. Statistical methods, inference and quantification of results 50
2.7.1. Bayesian vs. non-Bayesian frameworks . . . . . 50
2.7.2. Likelihood functions . . . . . 52
2.7.3. Censoring . . . . . 54
2.7.4. Maximum likelihood estimation and uncertainty quantification . . . . . 55
2.7.5. Goodness of fit statistics . . . . . 56
2.7.6. Monte Carlo methods . . . . . 56
2.7.6.1. Sequential Monte Carlo (SMC) . . . . . 57
2.7.6.2. Choice of SMC proposal . . . . . 65
2.7.6.3. Choice of SMC tempering schedule . . . 68
2.7.6.4. Assessing convergence of SMC runs . . . 68
2.7.6.5. Chi-squared distance between distributions 69
2.7.6.6. Further evaluating SMC output . . . . . 70
2.7.6.7. SMC examples on simple low-dimensional models . . . . . 72

**3. New applications of statistical analysis techniques 77**
3.1. SMC and identifiability . . . . . 77

3.1.1. SMC applied to structurally unidentifiable models 78
3.1.2. SMC applied to structurally locally identifiable models . . . . . 82
3.1.3. Using the posterior distribution to inform on identifiability . . . . . 84
3.2. SMC for computing profile likelihoods/posteriors . . . . 86

**4. Model development and analysis 88**
4.1. Pharmacokinetic model development and analysis . . . . 88
4.1.1. Structural identifiability . . . . . 88
4.1.2. Statistical analysis . . . . . 90
4.1.2.1. Weight cap . . . . . 90
4.1.2.2. Maximum likelihood results . . . . . 92
4.1.2.3. Sequential Monte Carlo results . . . . . 97
4.2. Coupled parasite lifecycle model . . . . . 100
4.2.1. Model development . . . . . 100
4.2.2. Model . . . . . 101
4.2.2.1. Submodel . . . . . 101
4.2.2.2. Submodel structural identifiability . . . 105
4.2.2.3. Submodel parameter estimates . . . . . 106
4.2.2.4. Ageing model . . . . . 111
4.2.2.5. Full model . . . . . 120
4.2.2.6. Model parameters . . . . . 128
4.2.3. Model simulation . . . . . 133

**5. Closing 138**
5.1. Conclusions and future work . . . . . 138

**Bibliography 141**

**Appendices** **157**

**Appendix A. Plots** **158**

A.1. Four compartment model pairwise posterior marginal distributions . . . . . 158

# List of Tables

2.1. Table of artemisinin *in vitro* red blood cell uptake data extracted from Vyas et al. [48] . . . . . 17

4.1. Structural identifiability analysis results for the four compartment model under different combinations of constraints 91
4.2. Table of parameter estimates and their uncertainties for patient A using the 4-compartment pharmacokinetic model. . 93
4.3. Correlation matrix for the parameters in Table 4.2 for the 4-compartment pharmacokinetic model. . . . . 94
4.4. Fitted parameter values, aggregated . . . . . 95
4.5. Table of parameters in the artemisinin red blood cell uptake submodel . . . . . 105
4.6. Table of parameter estimates in the artemisinin red blood cell uptake submodel . . . . . 107
4.7. Table of parameters for coupled model . . . . . 130
4.8. Table of state variables used in the coupled model . . . . . 131
4.9. Table of intermediate functions used in the coupled model . 132

# List of Figures

2.1. Main blood stages of the *Plasmodium falciparum* malaria parasite lifecycle and their durations . . . . . 10
2.2. Diagram representing receptor-ligand binding and associated rate constants . . . . . 25
2.3. Diagram representing receptor-ligand binding and associated non-linear rate flow rates when used as a component in an otherwise compartmental model . . . . . 25
2.4. Two compartment model used to illustrate structural identifiability concepts and for testing SMC sampler . . . . . 34
2.5. System diagram of the four compartment model for ARS (left part) and DHA (right part). Upper compartments represent the absorption compartments (unobserved). Lower compartments are the circulation compartments (observed). . . . . 46
2.6. 3D density plot of known correct posterior distribution from model used for testing . . . . . 73
2.7. Smoothed 3D density plot of output from SMC sampler applied to known-posterior testing model with 1,000 iterations, 10,000 particles, tempering exponent 2 . . . . . 74
2.8. Smoothed 3D density plot of output from SMC sampler applied to known-posterior testing model with 10 iterations (insufficient), 10,000 particles, tempering exponent 10 . . . . . 75

2.9. Smoothed 3D density plot of output from SMC sampler applied to known-posterior testing model with 5,000 iterations, 1,000 particles (insufficient), tempering exponent 10 . . . . 76

3.1. SMC results on a structurally unidentifiable model: marginal histogram density plot of $b\alpha$ and $k_{21}$ . . . . . 79
3.2. SMC results on a structurally unidentifiable model: marginal histogram density plot of $b\alpha$, $k_{e1}$ . . . . . 79
3.3. SMC results on a structurally unidentifiable model: marginal histogram density plot of $b\alpha$, $k_{e2}$ . . . . . 80
3.4. SMC results on a structurally unidentifiable model: marginal histogram density plot of $k_{21}$, $k_{e1}$ . . . . . 80
3.5. SMC results on a structurally unidentifiable model: marginal histogram density plot of $k_{21}$, $k_{e2}$ . . . . . 81
3.6. SMC results on a structurally unidentifiable model: marginal histogram density plot of $k_{e1}$, $k_{e2}$ . . . . . 81
3.7. SMC results on a structurally locally identifiable model: marginal histogram density plot of $b\alpha$, $k_{21}$ . . . . . 83
3.8. SMC results on a structurally locally identifiable model: marginal histogram density plot of $b\alpha$, $k_{e2}$ . . . . . 83
3.9. SMC results on a structurally locally identifiable model: marginal histogram density plot of $k_{21}$, $k_{e2}$ . . . . . 84

4.1. Example of model predicted ARS and DHA quantities/concentrations in each compartment for patient A using the 4-compartment pharmacokinetic model. Error bars are representative of assay error. Confidence bands are also shown but are very narrow. . . . . 93
4.2. Distribution of the coefficient of determination (%) over the Mahidol_PK dataset . . . . . 94
4.3. SMC estimate of posterior marginal distribution for $b\alpha$ for patient A . . . . . 97

4.4. SMC estimate of posterior marginal distribution for $k_{21}$ for patient A . . . . . 98
4.5. SMC estimate of posterior marginal distribution for $k_{42}$ for patient A . . . . . 98
4.6. SMC estimate of posterior marginal distribution for $k_{43}$ for patient A . . . . . 98
4.7. SMC estimate of posterior marginal distribution for $k_{e4}$ for patient A . . . . . 99
4.8. ARS model fit using point with maximal posterior density for patient A . . . . . 99
4.9. DHA model fit using point with maximal posterior density for patient A . . . . . 99
4.10. *In vitro* submodel for drug uptake into red blood cells . . . 103
4.11. Hematocrit against partitioning coefficient for uninfected cells (with $c_0 = 880$ nM, $t = 2$ hours, $p = 0\%$), data and model . . . . . 107
4.12. Incubation time (hours) against partitioning coefficient for uninfected cells (with $c_0 = 1410$ nM, $\zeta = 33\%$, $p = 0\%$), data and model . . . . . 108
4.13. Initial concentration against partitioning coefficient for uninfected cells (with $t = 2$ hours, $\zeta = 33\%$, $p = 0\%$), data and model . . . . . 108
4.14. Hematocrit against partitioning coefficient for infected cells (with $c_0 = 880$ nM, $t = 2$ hours, $p = 6\%$), data and model . 109
4.15. Incubation time (hours) against partitioning coefficient for infected cells (with $c_0 = 1410$ nM, $\zeta = 33\%$, $p = 7\%$), data and model . . . . . 109
4.16. Initial concentration against partitioning coefficient for infected cells (with $t = 2$ hours, $\zeta = 33\%$, $p = 6\%$), data and model . . . . . 110

4.17. Parasitaemia against partitioning coefficient (with $c_0 = 1410$ nM, $t = 2$ hours, $\zeta = 33\%$), data and model . . . . . 110
4.18. Parasite lifecycle model showing connectivity between live and shadow compartments . . . . . 116
4.19. Parasite lifecycle model simulation with all parasites initially in age class 1 . . . . . 118
4.20. Total density of parasites across all classes in simulation . . 119
4.21. Total density of all observable (non-sequestered) parasites in simulation . . . . . 119
4.22. Diagram of drug model component . . . . . 121
4.23. Illustrative full coupled parasite-drug model simulation, for identical doses of ARS administered at $t = 0$ and $t = 24$ hours 134

A.1. SMC estimate of posterior pairwise marginal distribution for $b\alpha$ and $k_{21}$ for patient A . . . . . 158
A.2. SMC estimate of posterior pairwise marginal distribution for $b\alpha$ and $k_{42}$ for patient A . . . . . 159
A.3. SMC estimate of posterior pairwise marginal distribution for $b\alpha$ and $k_{43}$ for patient A . . . . . 159
A.4. SMC estimate of posterior pairwise marginal distribution for $b\alpha$ and $k_{e4}$ for patient A . . . . . 160
A.5. SMC estimate of posterior pairwise marginal distribution for $k_{21}$ and $k_{42}$ for patient A . . . . . 160
A.6. SMC estimate of posterior pairwise marginal distribution for $k_{21}$ and $k_{43}$ for patient A . . . . . 161
A.7. SMC estimate of posterior pairwise marginal distribution for $k_{21}$ and $k_{e4}$ for patient A . . . . . 161
A.8. SMC estimate of posterior pairwise marginal distribution for $k_{42}$ and $k_{43}$ for patient A . . . . . 162
A.9. SMC estimate of posterior pairwise marginal distribution for $k_{42}$ and $k_{e4}$ for patient A . . . . . 162

A.10. SMC estimate of posterior pairwise marginal distribution for $k_{43}$ and $k_{e4}$ for patient A . . . . . 163

# Acknowledgements

I would like to thank my supervisors Dr Michael Chappell and Professor John Aston first and foremost, for their endless and much needed supply of guidance and encouragement. It has been an absolute pleasure. I also express my sincere gratitude to everyone else who has taught, advised, supported, employed, entertained and/or hosted throughout the last few years.

Deserving of special mention are (et alia): Professor Keith Godfrey, not least for introducing me to Dr Chappell; Dr Adam Johansen for his patient counsel on SMC matters; Dr Neil Evans for many fruitful discussions on structural identifiability; Professor Gareth Roberts for chairing my personal advisory committee; Professor Thomas Schön for acting as external examiner. Professor Steven Ward and Professor Joel Tarning provided me with expert advice and the all-essential data used in this work, and I hope they will find something interesting herein to partly repay them for their very much valued efforts.

Many thanks also to everyone involved in the MASDOC DTC for making this work possible. MASDOC, and in turn this work, was supported in part by the Engineering and Physical Sciences Research Council [grant number EP/HO23364/1].

Huge praise is due to Rob Bircumshaw of Coventry Law Centre for his instrumental help in fighting government interference that threatened the completion of this thesis.

My family have of course been incredibly supportive in every sense

imaginable. Naturally, I dedicate this thesis to Hong.

Thanks for reading!

# Declaration

With the exception of referenced material and the input and advice declared in the acknowledgements, this thesis is the work of its author.

The author also confirms that this thesis has not been submitted for a degree at another university. This thesis incorporates and extends work already submitted by the author for the degree of MSc in Mathematics and Statistics at the University of Warwick [1]. Such work is mainly restricted to parts of subsection 2.4.1, subsection 2.3.1, subsection 2.6.1 and subsection 4.1.1.

This thesis also contains work that has been published in Hall et al. [2] (and its reprint [3]) before the thesis was completed or examined, but was submitted for publication after the beginning of the author's period of study for this degree. Much of the work in Hall et al. [2] also overlaps with that described in the MSc [1].

# Abstract

Malaria is a parasite with a complex lifecycle, and commonly used anti-malarial agents from the artemisinin family have varied effectiveness over different stages of this lifecycle. The pharmacokinetic profile of the artemisinins is also strongly influenced by the parasite burden and lifecycle stage. This work introduces a new pharmacokinetic and pharmacodynamic model incorporating these interdependent drug and lifecycle features, for orally administered artesunate and its principal metabolite dihydroartemisinin.

This model, like the underlying system whose features it attempts to capture, is quite complex and cannot be solved analytically like standard linear first-order compartmental models previously used for pharmacokinetic modelling of these drugs. Therefore, understanding, inference and validity are explored through use of the modern statistical technique of a Sequential Monte Carlo sampler. Structural, numerical and practical identifiability are important concepts for all models, the latter two especially so in this case as the model structure does not admit an algebraic structural identifiability analysis. Motivated by this, the above identifiability concepts are also investigated in connection with the Sequential Monte Carlo technique.

Sequential Monte Carlo is demonstrated to be a useful tool for gaining insight into models whose structural identifiability is not known, just as it is also shown to have significant advantages in parameter inference over the classical approach.

The coupled parasite lifecycle and artemisinin-derivative model is built in stages, starting with an *in vitro* submodel capturing the dynamics of uptake of artemisinins into parasitised and non-parasitised red blood cells. Next, the parasite lifecycle, or 'ageing' model, is introduced, which uses a new concept of shadow compartments to achieve its aims of describing ageing in continuous time and to exhibit sufficient control over the parasite population. Finally, these models are integrated together into the full coupled pharmacokinetic and pharmacodynamic model. More work is needed to fully assess the resultant model on clinical datasets, but the building blocks upon which it was constructed appear to fulfil their aims reasonably well.

# Chapter 1

# Introduction

## 1.1. Background

A brief overview of the background of this work is presented here, so that the aims and objectives can be stated and understood. More details on each aspect is given in the dedicated background and literature chapter following this one, chapter 2.

***Plasmodium falciparum* malaria:** *Plasmodium falciparum* is the most deadly of the parasites that cause a severe blood disease known as malaria in humans. Infection starts when an infected Anopheles mosquito carrying sporozoites injects those sporozoites into the subject while taking a blood meal. Initially the sporozoites develop in the liver, and after some time release merozoites to begin the blood stage of the lifecycle. In this blood stage, which is the symptomatic stage, the parasite has a complicated lifecycle of approximately 48 hours, beginning with the invasion of red blood cells (RBCs) and ending when the RBCs burst to release more morozoites into the blood stream ready to invade again. Some merozoites develop into gametocytes, ready to picked up by another mosquito taking a blood meal later, where they further develop and replicate into sporozoites, from whence they can infect the next person. See section 2.1 for more information.

**Artemisinins / artesunate and dihydroartemisinin:** Artesunate (ARS) and dihydroartemisinin (DHA) are from the artemisinin class of anti-malarial compounds. These are important anti-malarials because they work rapidly and effectively, have good safety characteristics, and limited resistance to these compounds has developed so far. ARS is water-soluble and thus easy to administer, and is rapidly metabolised into DHA, which is therefore responsible for most of the anti-malarial action.

The pharmacokinetics and pharmacodynamics (defined next) of these compounds are thought to vary significantly based on the stage of the parasites within their lifecycle, and the resistance that the parasites have developed to the medication is again thought to manifest most prominently in a limited period of the lifecycle. The details of these stage-specific effects are given in section 2.2.

**Pharmacokinetics (PK):** Pharmacokinetics is the study of the time course of the movement and changes to a substance (typically a drug) after it enters the body. This encompasses the absorption, distribution, metabolism and elimination of the substance. In a quantitative sense, pharmacokinetics of a substance are often described through a mathematical model such as a compartmental model. In such a model, the body is represented in a simplified manner as a number of connected functional compartments, where the kinetic behaviour in each individual compartment is considered identical for all molecules of the drug represented as being lumped into that compartment.

**Pharmacodynamics (PD):** Pharmacodynamics, on the other hand, is the study of the effect on the body resulting from a specific substance. Pharmacodynamics is tightly coupled with the concept of pharmacokinetics since the amount of substance in a particular location in the body is the biggest influence on the effect that the substance has in that location.

Details on the modelling techniques and assumptions present therein

are described in section 2.4.

**Pharmacometrics:** The term pharmacometrics refers to pharmacokinetics, pharmacodynamics and other mathematical methods of analysis and interpretation of drug data.

**Structural, numerical and practical identifiability:** Once a mathematical model has been developed, values need to be assigned to its parameters. The goal is to find parameter values such that the model does a good job of describing real data. However, this is not necessarily a well-posed problem: even in a setting of perfect error-free data, there may be multiple (even uncountably many) parameter values that provide equivalent observable model behaviour, usually with quite different implications for model predictions of unobserved behaviour. This situation is known as *structural unidentifiability* if there are uncountably many such values or *structural local identifiability* if there are countably many. The converse is known as *structural global identifiability*. In the structurally non-global situations, the data alone (even if perfect and error-free as is assumed) is incapable of discriminating between the various predictions. This may mean that one or more desired outcomes from the model may instead have to be applied as assumptions in the modelling process, rather than informing on and providing evidence for those aspects. Alternatively, it may be necessary to modify the model and/or the data collected so that parameters of interest are instead structurally identifiable, which may then potentially permit collection of evidence for or against the desired outcome.

However, even if a model is structurally (globally) identifiable — and can be shown to be so (often this is very challenging to demonstrate) — it does not necessarily follow that the parameters can be reliably determined under the non-perfect conditions of discrete observations, especially where those observations are also made with uncertainty/error. The terms *numerical (un)identifiability* have been adopted in this thesis to describe the situation in this context, where it is known that the model

is perfectly capable of describing the underlying data (e.g. because it has been tested against synthetic data generated using the model).

Finally, when dealing with real data where the model is unlikely to be a perfect description of the underlying experimental system, the problem is known herein as the *practical identifiability* problem. More detail on these identifiability aspects is provided in section 2.5.

## 1.2. Aims and objectives

The main aim of this work was to develop, analyse and validate a mechanistic model coupling artesunate and dihydroartemisinin pharmacokinetics with the corresponding pharmacodynamics, that can account for effects that are specific to certain stages of the malaria parasite lifecycle.

Such a model would ideally:

1. be capable of describing *in vivo* and (*mutatis mutandi*) *in vitro* pharmacokinetics and pharmacodynamics on an individual level, with the potential to do the same on a population level;
2. provide reasonable prediction of dynamics that cannot be measured or observed;
3. be able to reproduce or confirm current knowledge or hypotheses regarding stage specific action and the effect of resistance;
4. be suitable for estimating the effect of different dosing regimens;
5. be able to identify or confirm, and simulate the effect of, potential improvements to treatments;
6. be relatively simple and easy to follow/use;
7. be free of jumps in the time course predictions and observations;

and the analysis would ideally include:

1. showing the model to be structurally identifiable (preferably globally so), as well as numerically and practically identifiable;

2. obtaining Bayesian posterior distributions of model parameters, for various subjects/*in vitro* studies and/or synthetic data;

3. determining the sensitivity of the model to the parameters.

This will therefore necessitate an exploration of the current methods available for analysing models for structural identifiability and determination of their suitability for the developed non-linear model with switching behaviour. Such structural identifiability methods may need to be adjusted and implemented.

As part of the statistical analysis and parameter inference, it will be necessary to implement a Monte Carlo method to cope with the high-dimensional situation and the complicated model. In particular, it is intended to implement a Sequential Monte Carlo (SMC) sampler. This class of samplers are still somewhat computationally demanding but have the advantages that they can be tested for convergence in a simple way and that the parameters do not need to be precisely tuned to behave reasonably. Of course, there is an element of tuning but essentially, given enough particles, enough iterations and enough runs, and a sensible proposal distribution, a good approximation of the posterior distribution will be obtained. Such samplers have apparently been used only rarely in these situations to date, so it will also be helpful to use an SMC sampler on simpler models to understand the behaviour and limitations of the method and situations to be aware of.

## 1.3. Structure of this thesis

Chapter 2 introduces the application area of malaria and motivates investigation of the artemisinin class of anti-malarial drugs. Background material is provided on the modelling and analysis techniques used later

in the thesis, including the important concept of structural identifiability, and some methods for conducting such a structural identifiability analysis, and a discussion of the difficulties that may arise in attempting to carry out such an analysis. Related forms of analysis known as numerical and practical identifiability analyses are also introduced, which complement a structural identifiability analysis and can suggest different angles for attacking the structural identifiability analysis if it has not been possible or straightforward to solve directly.

The specific pharmacokinetic, *in vitro*, and pharmacokinetic/pharmacodynamic datasets used in this thesis are introduced, as are existing models with similar aims to this work in summarised format.The pharmacokinetic model from Hall [1] is also presented again as it is revisited here.

A Sequential Monte Carlo method is introduced in subsubsection 2.7.6.1 for the purposes of parameter inference and for determining numerical and practical identifiability. This method is then applied to a low-dimensional model with known posterior distribution, and the output compared with that known posterior, and then also applied to the pharmacokinetic model.

In section 4.2, a new pharmacokinetic/pharmacodynamic model is then introduced, with some discussion on its behaviour and properties. Finally, section 5.1 outlines the further work relevant to the model that would ideally be conducted if time and funding permitted.

# Chapter 2

# Background and literature review

## 2.1. Malaria

Malaria is an extremely burdensome disease caused by micro-parasites with asexual lifecycles, and is transmitted via *Anopheles* mosquitoes as vectors. There are 5 malaria parasite species (*Plasmodium*) that can infect humans, and together they continue to infect hundreds of millions of people every year, with 3.2 billion people being considered "at risk" by the World Health Organisation. It is still a major killer in parts of the world, despite recent efforts having some success in reducing mortality rates [4]. Even where malaria does not cause death it is still has a severe effect on those infected, and on the health care systems and socioeconomics of the countries and communities where the most vulnerable live [5, 6].

Young children are the most severely affected, with malaria having killed nearly half a million African children in 2013 before they turned five (equivalent to approximately one death every minute) [7]. Indeed, malaria is the cause of 15% of deaths of children under five in parts of Africa [7].

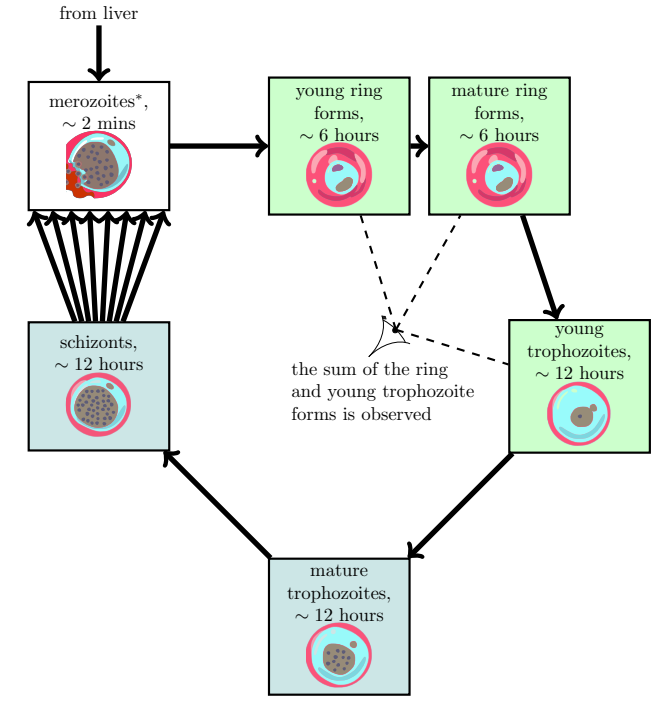
Malaria is (currently) an entirely preventable disease, with a single bed net (if used appropriately) providing an effective barrier from mosquito bites for a few years, spraying indoor insecticide is effective for many months, and prophylactic medicines are available. There are also recent hopes concerning malaria vaccines [8]. Many treatments are also available for malaria, which are often effective if they can be administered to patients in reasonable time. However, resistance to malaria treatments has been detected in many regions and is a major concern [9]. If new treatments cannot be introduced, and current treatments rendered more effective, both as a matter of urgency, then controlling malaria even to the same extent as currently will cease to be possible, and the unpleasant statistics presented above will only get worse.

The most deadly malaria species in humans is *Plasmodium falciparum* [4], and is the species focussed on in this work. After an initial phase in the liver, the parasites emerge into the blood stream and, from there, undergo a 48 hour lifecycle, depicted in Figure 2.1. It is in the blood stage that symptoms begin and this work concentrates on this blood stage, which tends to synchronise *in vivo*, i.e. with the majority of parasites simultaneously at similar stages in their lifecycles [10]. The parasite blood stage lifecycle starts with the invasion of red blood cells, from whence maturation begins. The parasites are seen to have a ring-shape for the first 12 hours or so, and these 12 hours are divided into "early" and "late" ring stages. For the next 12 hours, the parasites are said to be in the young trophozoite stage, and they begin to multiply into around 5–32 (mean 20) new daughter parasites called merozoites [11, 12, 13]. After this 12 hour period, the parasites become mature trophozoites, whereupon the infected red blood cells sequester in the organs to avoid clearance by the spleen [14], meaning that they do not circulate around the body with the remainder of the blood. Therefore, they cannot be detected in blood drawn from patients in the usual way. After the next 12 hours, known as the schizont phase, the cells burst and release the

merozoites into the blood stream to start the cycle afresh, but with potentially fewer red blood cells available for re-invasion. It is this bursting phase of the lifecycle that causes most symptoms [15], though the sequestration in organs can also reduce blood flow to those organs which can cause more severe symptoms [16].

The causes of transition from uncomplicated to severe malaria are not well understood [12]; it is not as simple as developing a high parasitaemia [17]. People in endemic areas are more likely to have gradually developed resistance over repeated exposure [18] and so are not affected by high parasitaemias in the same way as those from non-endemic areas [17]. Non-immune individuals might show symptoms with a parasite load of 10 per microlitre of blood, while partly immune individuals will not show symptoms unless the load reaches 1,000 parasites per microlitre of blood [19].

*Plasmodium falciparum* parasites are most metabolically active during the mature ring stage of their lifecycle, and this explains why many of the drugs for treating malaria are most effective against this stage [20]. The next subsection describes the class of anti-malarials of most interest here, *Artemisinins*.




*Each red blood cell releases between 8 and 20 merozoites when rupturing. These usually reinvade new red blood cells within minutes and start to develop into ring forms, renewing the cycle.



▲ Figure 2.1: Main blood stages of the *Plasmodium falciparum* malaria parasite lifecycle and their durations

## 2.2. Artemisinins

Although also used for other species of malaria and even different purposes altogether, the artemisinin class of compounds is highly effective against at least parts of the blood stage of *Plasmodium falciparum* and have been used with increasing frequency since the 1990s [21]. Artemisinins are generally recommended to be administered together with a concomitant drug with a different mechanism and a longer duration of action in an effort to delay resistance [9]. Such treatments are known as

artemisinin-based combination therapies and are the most effective antimalarials available today [22, 23]. As such, they are the treatment recommended by the World Health Organisation for many classes of patient [9].

Artemisinins alone are the most rapidly acting of all current antimalarial drugs [24] and are able to reduce the parasite biomass ∼10,000-fold per asexual life cycle [25]. Combined with their wide tolerance, it is clear to see that artemisinins are incredibly important to malaria control and treatment efforts. However, this class of anti-malarials have also shown waning efficiency in parts of South-East Asia [4], thought to be due to malarial resistance to the compounds affecting the associated pharmacodynamics in the early stages of the parasite lifecycle [26].

Artemisinin drugs distribute into red blood cells, primarily infected ones, and this is where their anti-malarial action is thought to occur. However, debate remains concerning the exact mechanism of action of these compounds [27, 28, 29]. The general theory is that iron accumulates in red blood cells which have been infected by malaria parasites, which then react with the artemisinin compounds, forming free radicals which in turn damage the parasites [30], although a number of unanswered questions remain [31]. The accumulation of artemisinins in parasitised red blood cells may therefore play an important part in their potency [24]. This suggests that the pharmacokinetics of the drug will differ vastly between healthy individuals and malaria patients, which is indeed found to be the case [32], though the degree of severity of the malaria infection appears to be less important [33]. However, the blood plasma concentration–time profiles and thus the pharmacokinetics of artemisinins have been shown to display high inter-individual variability in many studies, even after accounting for many covariates such as baseline parasitaemia, age, body weight and gender [34], and this variability increases in malaria patients compared to healthy volunteers [35].

Artemisinins are effective throughout a wide period of the parasite

lifecycle [26], though the effectiveness varies throughout that lifecycle [31, 36], especially where resistance has emerged [26, 37]. It is also likely that different derivatives are affected by resistance to differing extents [38].

Further understanding of the pharmacokinetics and pharmacodynamics of artemisinins may assist in informing more effective dosing regimens, discovering more about the effects of resistance, as well as in developing new or complementary antimalarials or artemisinin derivatives with desirable properties.

Artesunate (hereafter ARS) is the most frequently used artemisinin derivative, and is rapidly and almost entirely converted to dihydroartemisinin (hereafter DHA) *in vivo*, mostly by plasma esterases and liver cytochrome P450 CYP2A6 [39, 40, 41]. DHA is the most active of all major artemisinin derivatives, with activity approximately 1.4 times that of ARS [42].

ARS is water soluble, facilitating its absorption [43] (usually assumed to be fast, efficient and first-order [39]). Although it is rapidly hydrolysed into DHA, it may still make a significant contribution to parasite kill [44]. DHA is also rapidly eliminated from plasma, again either due to accumulation in infected red blood cells or through further metabolism (e.g. glucuronidation [45]), but the metabolites of DHA are inactive [46]. Although ARS and DHA are rapidly eliminated from plasma, antimalarial effect is observed to continue for some time after decline of both ARS and DHA plasma concentrations to ineffective levels [47].

## 2.3. Data

It is important when designing a model to consider the data that are available to calibrate or otherwise verify the model. A model that is too ambitious will not be supported by the data, while a model that is too simplistic may not be able to relate to the data in full, or may

not reproduce some of the observed features of the data. Therefore, an overview of the available data is provided now, before any models or modelling techniques are discussed.

Two *in vivo* datasets have been made available for use in this thesis. One consists of pharmacokinetic data only, and will be referred to as the Mahidol_PK dataset. This dataset was also used in Hall [1] and Hall et al. [2]. The other, corresponding to a different study, contains pharmacokinetic and pharmacodynamic data, and will be referred to as the MORU_ARC3_PD dataset.

Additional data were extracted from Vyas et al. [48] and are also presented here.

### 2.3.1. Mahidol_PK dataset

The Mahidol_PK dataset consists of 19 malaria patients from a study carried out at the Department of Clinical Tropical Medicine, Faculty of Tropical Medicine, Mahidol University, Bangkok, 10400, Thailand. Patients were selected based on a diagnosis of adult non-severe *P. falciparum* malaria with a parasite count less than 10,000 parasites per microlitre of blood. The patients were each administered 2mg/kg artesunate in fractions of 50mg oral tablets (body weights not part of the dataset provided to the author) twice daily for three days, in combination with 1800mg fosmidomycin and 750mg azithromycin which are antibiotics and not considered relevant to the modelling (described later). Food was restricted for the first hour after dosing.

The data consist of ARS and DHA concentrations (provided in units of ng/ml but converted to nmol/l prior to analysis) from assayed blood plasma samples over a time course of 12 hours. Blood plasma samples were drawn from the patients immediately after administration of the first dose on the first day, 15 minutes after, 30 minutes after, 1 hour after, 1.5 hours after, 2 hours after, 3 hours after, 4 hours after, 6 hours after,

8 hours after and 12 hours after administration of the first dose on the first day. No samples were taken for subsequent doses or on subsequent days and so cannot be included in the modelling.

Samples were analysed to determine their ARS and DHA concentrations using tandem liquid chromatography-mass spectrometry (on a Thermo Fisher Quantum Access Triple Quad Mass Spectrometer) based on the assay described in Hanpithakpong et al. [49]. (The individual samples were analysed only once but assay robustness was confirmed by a re-analysis of approximately 10% of all samples. Analytical runs included a full calibration curve and three replicate quality control samples.) The assay has an associated *lower limit of quantification (LLOQ)* for each analyte and passed Food and Drug Administration (FDA) validation, for which the requirement is to measure quality control samples and standard curve samples with known concentrations above the respective LLOQ to within ±15% of the nominal value. Specifically, the coefficient of variation for the assay is 15% for both analytes. Values below the respective LLOQ may have significantly greater relative uncertainty or noise. The LLOQ for ARS was $\text{LLOQ}_1 = 3.9$ nmol/l and that of DHA was $\text{LLOQ}_2 = 22.9$ nmol/l. The assumption is that values reported for unknown samples above the respective LLOQ will also be within 15% of the actual value. Observations below the respective LLOQ are felt to be so unreliable that such values are not quantified; they are only reported as being below the limit of quantification (BLQ). In this way, 41% of the ARS data and 8% of the DHA data are censored.

Note that over the 12 hour time span for a single subject, a wide range of drug concentrations was observed, most particularly for DHA. Specifically, for DHA, concentrations smaller than the LLOQ and concentrations above $6,000$ nmol/l were recorded for some patients over the course of the sampling interval. In common with other studies, there was also wide variability between patients in terms of the concentration–time profiles for both ARS and DHA.

The majority of the patients had peak ARS concentrations within 1.5 hours after drug administration (74%), and peak DHA concentrations within 2 hours (63%). However, it was already clear from the data that over half of the patients experienced delayed or possibly double peaks in the concentration–time profiles for both ARS and DHA. These are not thought to be outliers due to the assay validation, and the pattern is quite consistent in some individuals. There are no covariates with these data to allow further analysis and the cause of this phenomenon is not known, nor the frequency of incidence in other artesunate studies as individual patient profiles are often not discernible. This issue does not appear to be widespread or well known in relation to artemisinin drugs, but has been reported with the derivative artemether, which was found by Van Agtmael et al. [50] to have a biphasic absorption profile. As the mechanistic cause of the phenomenon is unknown, the differences in the absorption process have not been accounted for in the present model. This indicates that the model is misspecified and will not be suitable for all the patients, though it is hoped that it will still be applicable for many of the patients. The patients were therefore divided into two groups, one where the concentration–time profiles for both ARS and DHA exhibited only a single peak each within the expected time after drug administration, and the other group for the remaining patients where the absorption profile was unexpected, e.g. being slower to reach the peak concentrations, having multiple peaks and/or having delayed elimination.

### 2.3.2. In vitro uptake data

Vyas et al. [48] have conducted radio-labelled *in vitro* experiments to assess the rate, extent and reversibility of uptake of artemisinin into *Plasmodium falciparum* parasitised and non-parasitised red blood cells. They incubated radio-labelled artemisinin in blood for set times under controlled experimental conditions, varying these conditions to invest-

igate the effect of haematocrit, parasitaemia, temperature, artemisinin concentration and incubation time on the extent of uptake into red blood cells. After incubation, they separated (centrifuged) the red blood cells (precipitate/pellent) from the plasma (supernatant) and counted the disintegrations per minute from each. The partitioning fraction was then calculated as the ratio of the number of disintegrations per minute from the pellent compared to the total number of disintegrations per minute, and thus shows the degree of uptake of the artemisinin into the red blood cells.

They used their results to conclude that artemisinin diffused passively and reversibly across the red blood cell membrane for uninfected cells, but that uptake was higher, saturable and irreversible in infected red blood cells, supporting the belief that the process is carrier mediated.

Their quantitative experimental results were presented in graphical format, showing the uptake into the red blood cells under the various experimental conditions. Some of these results have been extracted from the graphs to facilitate use in models in this work, and are presented in Table 2.1. Results obtained below body temperature (37°C) are not of interest here and have not been extracted.

Although these results concern artemisinin itself, rather than artesunate or dihydroartemisinin which are of interest here, there are no known available data for the latter two derivatives, and because they are from the same family, it will be assumed that they behave similarly.

| Experimental conditions | | | | Observation |
|---|---|---|---|---|
| initial artemisinin concentration (nM) | parasitaemia (fraction) | hematocrit (fraction) | incubation time (hours) | partition in pellet (fraction) |
| 880 | 0 | 0.1 | 2 | 0.20 |
| 880 | 0 | 0.2 | 2 | 0.28 |
| 880 | 0 | 0.33 | 2 | 0.39 |
| 880 | 0 | 0.44 | 2 | 0.48 |
| 880 | 0.06 | 0.1 | 2 | 0.38 |
| 880 | 0.06 | 0.2 | 2 | 0.52 |
| 880 | 0.06 | 0.33 | 2 | 0.62 |
| 880 | 0.06 | 0.44 | 2 | 0.67 |
| 1,410 | 0.07 | 0.33 | 0.083 | 0.33 |
| 1,410 | 0.07 | 0.33 | 0.167 | 0.36 |
| 1,410 | 0.07 | 0.33 | 0.333 | 0.39 |
| 1,410 | 0.07 | 0.33 | 0.5 | 0.46 |
| 1,410 | 0.07 | 0.33 | 0.75 | 0.50 |
| 1,410 | 0.07 | 0.33 | 1 | 0.54 |
| 1,410 | 0.07 | 0.33 | 2 | 0.55 |
| 1,410 | 0.07 | 0.33 | 3 | 0.54 |
| 1,410 | 0 | 0.33 | 0.083 | 0.30 |
| 1,410 | 0 | 0.33 | 0.167 | 0.30 |
| 1,410 | 0 | 0.33 | 0.333 | 0.32 |
| 1,410 | 0 | 0.33 | 0.5 | 0.32 |
| 1,410 | 0 | 0.33 | 0.75 | 0.33 |
| 1,410 | 0 | 0.33 | 1 | 0.31 |
| 1,410 | 0 | 0.33 | 2 | 0.34 |
| 1,410 | 0 | 0.33 | 3 | 0.33 |
| 880 | 0.06 | 0.33 | 2 | 0.61 |
| 1,060 | 0.06 | 0.33 | 2 | 0.63 |
| 1,230 | 0.06 | 0.33 | 2 | 0.52 |
| 1,410 | 0.06 | 0.33 | 2 | 0.49 |
| 1,760 | 0.06 | 0.33 | 2 | 0.49 |
| 2,110 | 0.06 | 0.33 | 2 | 0.44 |
| 2,820 | 0.06 | 0.33 | 2 | 0.40 |
| 3,520 | 0.06 | 0.33 | 2 | 0.37 |
| 880 | 0 | 0.33 | 2 | 0.37 |
| 1,060 | 0 | 0.33 | 2 | 0.36 |
| 1,230 | 0 | 0.33 | 2 | 0.35 |
| 1,760 | 0 | 0.33 | 2 | 0.33 |
| 2,110 | 0 | 0.33 | 2 | 0.35 |
| 2,820 | 0 | 0.33 | 2 | 0.36 |
| 3,520 | 0 | 0.33 | 2 | 0.36 |

Table 2.1: Table of artemisinin *in vitro* red blood cell uptake data extracted from Vyas et al. [48]

### 2.3.3. MORU_ARC3_PD dataset

A pharmacokinetic, pharmacodynamic dataset including ARS and DHA concentration-time courses, and parasitaemia and haematocrit time courses was provided for 80 *Plasmodium falciparum* patients across two study sites, 40 patients from each of Wang Pha (Thailand) and Pailin (Cambodia), which are areas of low transmission intensity. Half of the patients from each site were randomly assigned to receive 2 mg/kg oral artesunate every 24 hours for seven days, and the other half received 4 mg/kg oral artesunate for 3 days together with mefloquine (another anti-malarial agent) on days 3 and 4. The (anonymised) data consist of artesunate and dihydroartemisinin concentrations (ng/ml converted to nmol/l, "<LLOQ" or "No Peak") in blood plasma (not whole blood concentrations) and timings thereof (date, hours, mins) at various intervals over the treatment period, together with parasite counts (per $mm^3$ whole blood, decimal with varying degrees of accuracy) and timings thereof (decimal relative to first dose) and haematocrit levels (%, decimal with varying degrees of accuracy, same timings as parasite counts) during both the treatment (every 4–8 hours) and follow-up periods. The exact time and dates (date, hours, mins) and amounts (mg) of artesunate doses were recorded, as were any adverse events that the patients experienced (e.g. if patient vomited and needed the drug to be re-administered), symptoms, concomitant diseases and medications, and previous relevant medication and comprehensive admission data, including age at time of enrolment, height, weight, biochemistry, haematology and daily clinical observations (heart rate, respiratory rate, temperature, etc).

These data were previously used by Dondorp et al. [51] and Saralamba et al. [37]. Dondorp et al. [51] used these data to conclude that the malaria has reduced *in vivo* susceptibility to artesunate in Pailin compared to Wang Pha, corresponding to slower parasite clearance. Saralamba et al. [37] used these data to support a hypothesis that the resistance mainly

reduces ring-stage susceptibility, and suggests doubling the frequency of artesunate dosing to restore the higher parasite clearance rate.

Mefloquine was not described earlier since it is not a focus of this work, but it is an anti-malarial compound mostly affecting schizonts, and reaches its peak concentration following a single dose in 17.6 hours and has a mean elimination half-life of 18.1 days [52]. This is therefore a much longer-acting compound when compared to the artemisinins and so is a suitable partner drug for artemisinin combination therapy. However, unlike the artemisinins, a single dose of mefloquine does not provide sufficient blood concentrations to achieve a reasonable efficacy, and so it is normally dosed with a high dose daily for 3 days, and lower doses thereafter, to reach such a suitable concentration during the third or forth days [52]. The data for the later days where mefloquine has been used as a concomitant cannot therefore be directly used to determine the efficacy of the the artemisinins to schizonts, but it may be possible to estimate the contribution that the mefloquine made to parasite kill.

## 2.4. Modelling techniques

In order to extract the most information out of a model, sufficient information must be incorporated in the design of the model. Namely, in order to use a model for predictive purposes, knowledge (or predictions) of the mechanisms used by the system being modelled should be incorporated into the model. This is preferable to purely data-driven models, which lack the same kind of predictive power to answer questions about what might happen at times beyond the observed range, or with different inputs or input sites, for example. While a mechanistic model is more likely to suffer from model mis-specification, data can help discriminate between candidate mechanistic models and to provide evidence for or against specific mechanisms. The modelling techniques used in this work are mechanistic and include a combination of compartmental and

receptor-ligand type models. For simplicity, all randomness is assumed to arise from a non-systematic error on the observations; the models themselves are deterministic.

### 2.4.1. Compartmental models

*(Note that much of the material in this section is adjusted from Hall [1].)*

A *compartmental system model* is made up of a finite number of

- *compartments* — kinetically homogeneous amounts of material [53, p. 3] (homogeneous, well-mixed, lumped subsystems [54, p. 1]) (with associated *quantities*);
- *flows* — material transfer between compartments and to and from the environment, including inputs to the system (with associated *flow rates*),
- *observations* — functions of the compartments defining quantities of which measurements are to be made at least once,

such that the quantities of material within each of the compartments may be described by a set of coupled first-order ordinary differential equations [54, p. 1].

By employing compartmental modelling, some key assumptions/simplifications are made:

- mixing of the substance within each compartment is complete (if not, the flow rates will not always be proportional to the present value of the concentration [55, p. 60])
- mixing of the substance within each compartment is instantaneous [55, p. 60].

Note that these assumptions are unlikely to hold in practice, but they are simplifications which allow easy model formulation and often result in models which are successful [54].

Only deterministic compartmental models are considered here. Compartmental models involve mass-balance and can be succinctly described by diagrams, though rate functions and initial conditions usually also need to be specified, and depending on the number and length of such expressions, it can be clearer to do so separately from the diagram.

The compartments are typically represented by circles or rectangles, flows between compartments are represented by simple arrows, and outflows from the system which are a special case of flows (elimination flows) are represented by arrows similarly to flows, but without a receiving compartment at the end of the arrows. Flows which are linear will be represented in this thesis by having a single arrowhead, while non-linear flows will be represented with a double arrowhead. Non-linear flows will generally be given by the flow rate rather than the flow rate constant. Inputs to the system will be represented by double-stroke arrows, and finally, observations will be represented by an eye connected by dashed lines to any compartments whose quantity or concentration of material directly forms part of the observation function, unless it would be too visually distracting to do so.

Throughout, $\mathbb{T}$ will denote the time domain of interest, and without loss of generality will be assumed to start at initial time 0. The quantity of material in compartment $i$ at time $t \in \mathbb{T}$ is often denoted $q_i(t)$ [56, p. 260] (which therefore has the constraint of being non-negative), and the flow rate from compartment $i$ to compartment $j$ is denoted by $k_{j,i}(t)$ (note that conventions for the order of indices of flows rates differ between the pharmacokinetic literature and elsewhere [54, p. 2]), while the environment — representing anything external to the system model — will be denoted by e (some authors use 0 [56, p. 260]).

For linear time-invariant flows, rate *constants* $k_{j,i}$ are generally shown instead of flow rates because they simplify the presentation (and in particular it is easier to fit the information on the model diagram). However, the notation for rate constants is somewhat abusive: the corresponding

flow rate at time $t$ is the rate constant times the quantity of material in the donor compartment at time $t$, i.e. $k_{j,i}(t) = k_{j,i}q_i(t)$, where $k_{j,i}(t)$ represents an (instantaneous) flow rate, and the constant $k_{j,i}$ in the right hand side of the equation is a flow rate constant.

This work will use flow rates which are donor-controlled, as well as flow rates which are not. Donor controlled means that the flow rate is a function only of the quantity of material in the donor compartment, not of any other compartments, i.e. the flow rate $k_{j,i}(t)$ from compartment $i$ to compartment $j$ is a function of $q_i(t)$ only: $k_{j,i}(t) = f(q_i(t))$ for some function $f$. In the linear flow case, this $f$ is linear, and so admits a description by a flow rate constant as above. For flow rates that are not purely donor-controlled, it can be helpful to draw some connection (e.g. a dashed line) between the flow rate and the controlling quantities, but in practice this can overwhelm the diagrams.

The most general form of equation for the quantity of material in compartment $i$ in a non-linear compartmental system, which is derived on the basis of mass-balance principles, is:

$$q_i'(t) = k_{i,\mathrm{e}}(t) + \sum_j k_{i,j}(t) - \left( k_{\mathrm{e},i}(t) + \sum_j k_{j,i}(t) \right),$$

where the sums are taken over all compartments $j \neq i$, and any flows which do not occur have rate function identically 0.

These equations, together with equations for the observation function $\boldsymbol{y}\colon \mathbb{T} \to \mathbb{R}^m$ and initial conditions $\boldsymbol{q}(0) \in \mathbb{R}^n$, make up the general form of a compartmental model:

$$\begin{cases} \boldsymbol{q}'(t) &= \boldsymbol{f}(\boldsymbol{q}, \boldsymbol{p}, \boldsymbol{u}, t) \\ \boldsymbol{q}(0) &= \boldsymbol{q}_0(\boldsymbol{p}) \\ \boldsymbol{y}(t) &= \boldsymbol{h}(\boldsymbol{q}, \boldsymbol{p}, t), \end{cases} \tag{2.1}$$

where $t \in \mathbb{T}$. $\boldsymbol{f}$ has special properties due to the model's compartmental

structure. Here, $\boldsymbol{p} \in \mathbb{R}^d$ denotes the collection of system parameters and $\boldsymbol{u} \in \mathbb{R}^s$ denotes the input. $\boldsymbol{h}$ is called the *observation function*. From a classical systems perspective, $\boldsymbol{h}$ might be referred to as the *output* of the system, especially when considering input-output structure, but during the modelling stage, it is helpful to be lucid about the distinction between material outflow and observations.

### 2.4.2. Non-compartmental models

A model similar to a receptor-ligand binding model will be assumed for the active uptake of artemisinins by red blood cells. Red blood cells are treated as having a varying number of receptors based on whether the cell is infected by malarial parasites, and the stage within the lifecycle of those parasites. Artemisinins are treated as ligands, which form receptor-ligand *complexes* when bound to receptors in the red blood cells. Compartmental models cannot conveniently be applied to model receptor-ligand binding, because unless each complex compartment is duplicated (for each species involved in the complex), the principle of mass-balance is violated. Certain assumptions could generally be made which can enable use of approximation by Michaelis-Menten kinetics, but these assumptions are not valid in the particular receptor-ligand context used here, and so this concept is not discussed.

The approach taken here is to relax the assumptions of "compartmental mass-balance", which also means that simple model diagrams are no longer sufficient to fully describe a model, because the concept of "flow" is altered: a receptor and a ligand need to bind (i.e. a donation needs to be made from 2 source compartments) in order for there to a be a flow into the receptor-ligand complex compartment. Of course, mass-balance equations still apply to such reactions but strictly compartmental models require no merging or splitting of molecules. The other assumptions for compartmental models are carried forward to receptor-ligand

models: flows are first-order and receptors and ligands are well-mixed (equally accessible). It is also assumed that partial binding is not possible, and if the dissociation rate constant is non-zero, that binding does not alter the ligand or receptor so each can again participate in a future binding reaction.

Denoting the molar concentration of receptors by $[R]$, the molar concentration of ligands by $[L]$, the molar concentration of receptor-ligand complex by $[RL]$, the association rate constant by $k_{\text{on}}$ and the dissociation rate constant by $k_{\text{off}}$, the equations describing the system are given by the law of mass action:

$$\frac{\mathrm{d}}{\mathrm{d}t}[R] = -k_{\text{on}}[R][L] + k_{\text{off}}[RL],$$

$$\frac{\mathrm{d}}{\mathrm{d}t}[L] = -k_{\text{on}}[R][L] + k_{\text{off}}[RL],$$

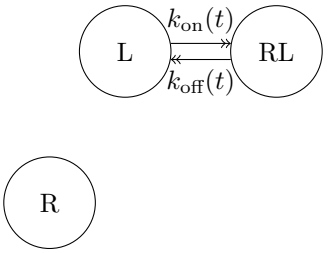
$$\frac{\mathrm{d}}{\mathrm{d}t}[RL] = k_{\text{on}}[R][L] - k_{\text{off}}[RL].$$

It can therefore be seen that $[R] + [RL]$ and $[L] + [RL]$ are conserved (constant) quantities.

Generally if such models are discussed on their own, they are depicted similar to Figure 2.2. However, in this work, they will be incorporated into a more complicated model that would otherwise be compartmental, so the receptors will be shown as disconnected from the ligand and complex compartments, and molecular quantities will be shown instead of molar concentrations, as in Figure 2.3. It is noted that this does not show all required details, but is felt to be clearer than diagrams where such details are attempted to be shown, especially where the same receptor pool can bind to ligands from multiple different compartments.

$$[\mathrm{R}] \quad + \quad [\mathrm{L}] \quad \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} \quad [\mathrm{RL}]$$

▲ Figure 2.2: Diagram representing receptor-ligand binding and associated rate constants




▲ Figure 2.3: Diagram representing receptor-ligand binding and associated non-linear rate flow rates when used as a component in an otherwise compartmental model

The general form of system model equations that will be considered here is:

$$\begin{cases} \boldsymbol{q}'(t) &= \boldsymbol{f}(\boldsymbol{q}, \boldsymbol{p}, \boldsymbol{u}, t) \\ \boldsymbol{q}(0) &= \boldsymbol{q}_0(\boldsymbol{p}) \\ \boldsymbol{y}(t) &= \boldsymbol{h}(\boldsymbol{q}, \boldsymbol{p}, t), \end{cases} \tag{2.2}$$

($t \in \mathbb{T}$) which is very similar to (2.1) except that the compartmental structure of $\boldsymbol{f}$ is no longer assumed (though the positivity constraint still applies).

## 2.5. Structural identifiability

### 2.5.1. Overview and importance

Once a structural model has been designed, it will very often have a number of unknown parameters that might be particular to a certain patient or population, rather than having a value derived from a universal physical law. In order for a model to have maximal utility, it is necessary to estimate those unknown parameters and to specify the degree of certainty to which those estimates are valid. It may be the case that multiple parameter sets have indistinguishable input/output (stimulus/response) behaviour. In such cases, model utility may be limited since any information extrapolated from the model may have a wide range of equally plausible values, without there necessarily being any way to filter those values into a meaningful or manageable range.

Once a candidate model and its input-output behaviour have been selected, it becomes possible to ask the theoretical question of whether that observable input/output behaviour (together with any relevant prior knowledge) is sufficient to determine the necessary information for a particular application of the model, even before considering the issue of data error or sampling frequency. Indeed (except for in certain machine learning situations where the parameters themselves are not of interest), the question should be asked before attempting to collect or analyse data, and where possible, attempts should be made to answer it. This question is known as the *structural identifiability* question.

It is possible that the input/output behaviour is not influenced by some of the system parameters, or that the influence is only in combinations that do not allow to distinguish the parameters' effects separately, and these are the cases when a model is not structurally identifiable.

Structural identifiability is considered under the ideal assumptions of having both a "perfect model and perfect data", and (unless the situation is altered by use of a prior distribution) is a prerequisite for the related

questions of *numerical* and *practical identifiability*, where these assumptions are relaxed (though only partially so in the case of numerical identifiability). A "perfect model" means that the model and the system being modelled are identical [57, p. 20], while "perfect data" involves having data available over a continuum of times (or less paradoxically, arbitrarily dense data), for any time period of interest, and with no observation error [58]. Structural identifiability is therefore a property of a structural model, where that structure is taken to include the controls and observables. It is independent of any specific datasets, data sampling schemes or data quantification methods. It is generally of interest to know not only whether a model is structurally identifiable or not, but which parameters are structurally identifiable or structurally unidentifiable, and to know the complete functional form of other parameter sets that are indistinguishable from the input/output behaviour of the model. Depending on the structure of the model, it can sometimes be easier to solve the structural identifiability problem, involving an algebraic analysis, while for other models, the algebra may be too difficult even for computational algebra systems to solve, and so answering the numerical and/or practical identifiability questions may be easier in those situations, though these cannot directly give the same level of information. However, to determine the structural unidentifiability of the system, one only needs to find one generically indistinguishable alternative parameter set, rather than finding all such indistinguishable parameter sets which most structural identifiability methods unavoidably attempt to do. Note that the related problem of structural distinguishability between two or more models is not considered in this work.

If the result of the structural identifiability analysis is that important parameters are not structurally identifiable, then it may be necessary to modify the model structure (including the observable input-output behaviour) and restart the modelling procedure with a modified model. Such an analysis is therefore ideally conducted before doing any experi-

ments to determine whether those experiments are capable of producing meaningful results, and/or to gain insights on how to improve such experiments, for example by suggesting to make observations of another state variable or variables (or functions thereof). However, it may also be possible to re-parameterise the model in such a way that it becomes at least structurally locally identifiable. It is for this reason that structural unidentifiability is also termed *parameter redundancy* by some authors [59].

Once it has been established that a model is structurally identifiable, numerical and practical identifiability analyses can suggest specific times at which it might be optimal to collect data samples from the system being modelled, and can also be used to give an indication of how many samples it might be appropriate to collect, or might provide the information that, while the model might be structurally identifiable, estimation of the parameters from the data it is planned to collect will still not be practically possible.

Only after conducting these analyses should one proceed to actually attempt to collect real data of interest, or to use already collected data for parameter inference. Attempting to estimate parameters from a model whose structural, numerical and practical identifiability has not been considered is to invite significant and severe problems, to which a researcher may sometimes remain unaware; at best, they would be unable to obtain any parameter estimates at all. At worst, incorrect conclusions could be reached, and with false confidence.

### 2.5.2. Definitions

There are a number of subtly incompatible definitions in the literature for structural identifiability, and even more so when expanding to differently named concepts of identifiability. These differences and incompatibilities are often perhaps (though certainly not always) unintentional.

Fix a parameterised system model $\mathcal{M}(\boldsymbol{p})$ of the form (2.2), and denote by $\boldsymbol{p} = (p_1, \ldots, p_d)$ a parameter vector parameterising the model. Let $\Omega \subseteq \mathbb{R}^d$ denote the space of feasible parameter values. Inequality constraints should be incorporated into the feasible parameter space. Only those parameters which are to be considered unknown should be part of the model parameterisation $\boldsymbol{p}$ in this section. Any parameters which are to be considered as known functions of other parameters should also be eliminated from the parameterisation. If a particular value for a parameter is known, it can sometimes be helpful to use that particular value during the structural identifiability analysis, though equally the converse can also be true: that using a generic parameter value makes the analysis easier, and it is therefore wise to try both generic and specific situations. Any other prior beliefs about the parameters can be considered in a practical identifiability analysis through a prior distribution, rather than being considered at this stage.

Note that a model in the above form includes specified initial conditions, possibly in terms of unknown parameters. Though this setup is intended here, it is worth noting that other authors advocate different approaches and use different terms to describe the setup here [60].

- Let $\mathcal{U} \subseteq \mathbb{R}^s$ be the space of admissible inputs to the system.

  If the system is uncontrolled then take $\mathcal{U} = \{\varnothing\}$ (i.e. the only admissible input is the empty function). It is common to simply refer to an output equation instead of an input/output equation in such circumstances, without needing to involve the empty function.

- Let $\hat{\Sigma}(\boldsymbol{p}){:}\mathcal{U} \to \mathcal{F}(\mathbb{T}; \mathbb{R}^m)$ denote the input/output map for the model $\mathcal{M}(\boldsymbol{p})$. This is the map such that $\hat{\Sigma}(\boldsymbol{p})(\boldsymbol{u}) = \boldsymbol{y}$ where $\boldsymbol{y}{:}\,\mathbb{T} \to \mathbb{R}^m$ are the observations from the model $\mathcal{M}(\boldsymbol{p})$ when its input is $\boldsymbol{u}$. This map plays a key role in structural identifiability, but it is not always straightforward to obtain an algebraic form of the input/output map for models of interest. A "structural identi-

fiability method" generally refers to a method for either generating such (generic) algebraic input/output maps, or generating exhaustive summaries, which generically capture all the same information as the input/output map.

Here, $\mathcal{F}(\mathbb{T};\mathbb{R}^m)$ denotes the space of functions defined on the time domain $\mathbb{T}$ with co-domain $\mathbb{R}^m$.

**Definition 1** (Parameter indistinguishability)**.** Parameters $\boldsymbol{p}$ and $\overline{\boldsymbol{p}}$ are said to be *indistinguishable* under $\mathcal{M}$ iff $\hat{\Sigma}(\boldsymbol{p}) = \hat{\Sigma}(\overline{\boldsymbol{p}})$.

**Definition 2** (Exhaustive (parameter) summary)**.** A function $\Sigma\colon \Omega \to \mathbb{R}^{d_\Sigma}$ is called an *exhaustive summary* (or a *structural invariant vector* [61]) for $\mathcal{M}$ iff for almost all values of $\boldsymbol{p} \in \Omega$,

$$\Sigma(\boldsymbol{p}) = \Sigma(\overline{\boldsymbol{p}}) \iff \hat{\Sigma}(\boldsymbol{p}) = \hat{\Sigma}(\overline{\boldsymbol{p}}).$$

$d_\Sigma$ is allowed to be infinite in this definition, so an exhaustive summary may consist of a countably infinite sequence (e.g. as is the case for Taylor series coefficients — see subsection 2.5.5).

Note that $\hat{\Sigma}$ is trivially an exhaustive summary. Exhaustive summaries generically capture all the information that is available from the observable input/output behaviour of the system model. Identifiability and structural identifiability are concerned with whether or not the information from the observable input/output behaviour of the model is sufficient to distinguish one parameter set from another, i.e. whether there exist distinct (and possibly arbitrarily close) parameters $\boldsymbol{p}, \overline{\boldsymbol{p}}$ with the same observable behaviour $\hat{\Sigma}(\boldsymbol{p}) = \hat{\Sigma}(\overline{\boldsymbol{p}})$.

Formally:

**Definition 3** (Identifiability of parameter combinations at fixed $\boldsymbol{p}$)**.** Fix $\boldsymbol{p} \in \Omega$. With respect to $\mathcal{M}$, a parameter combination $\boldsymbol{c}(\boldsymbol{p})$ is said to be:

- *globally identifiable* at $\boldsymbol{p}$ iff

$$\text{if } \overline{\boldsymbol{p}} \in \Omega \text{ and } \hat{\Sigma}(\boldsymbol{p}) = \hat{\Sigma}(\overline{\boldsymbol{p}}) \text{ then } \boldsymbol{c}(\overline{\boldsymbol{p}}) = \boldsymbol{c}(\boldsymbol{p}); \tag{2.3}$$

- (at least) *locally identifiable* at $\boldsymbol{p}$ iff there exists a neighbourhood $N(\boldsymbol{p}) \subseteq \Omega$ of vectors around $\boldsymbol{p}$ such that

$$\text{if } \overline{\boldsymbol{p}} \in N(\boldsymbol{p}) \text{ and } \hat{\Sigma}(\boldsymbol{p}) = \hat{\Sigma}(\overline{\boldsymbol{p}}) \text{ then } \boldsymbol{c}(\overline{\boldsymbol{p}}) = \boldsymbol{c}(\boldsymbol{p}); \tag{2.4}$$

- *unidentifiable* at $\boldsymbol{p}$ otherwise.

Note that the qualifier "at least" is replaced with the qualifier "non-uniquely" when it is known that the parameter is not also globally identifiable. These qualifiers remove the ambiguity that would otherwise arise when referring to an (unqualified) "locally identifiable" parameter combination.

In practice, one might assess the number of solutions to $\hat{\Sigma}(\boldsymbol{p}) = \hat{\Sigma}(\overline{\boldsymbol{p}})$ as follows. Write $\mathcal{S}_{\boldsymbol{c}}(\Sigma, \boldsymbol{c}(\boldsymbol{p}))$ to denote the set of all $\boldsymbol{c}(\overline{\boldsymbol{p}})$ values within such solutions, i.e.

$$\mathcal{S}_{\boldsymbol{c}}(\Sigma, \boldsymbol{p}) := \left\{ \boldsymbol{c}(\overline{\boldsymbol{p}}) \colon \hat{\Sigma}(\boldsymbol{p}) = \hat{\Sigma}(\overline{\boldsymbol{p}}) \text{ for some } \overline{\boldsymbol{p}} \in \Omega \right\}.$$

The condition of the parameter combination $\boldsymbol{c}$ being globally identifiable (2.3) is equivalent to $\mathcal{S}_{\boldsymbol{c}}(\Sigma, \boldsymbol{p})$ being a singleton set. If the cardinality of the set $\mathcal{S}_{\boldsymbol{c}}(\Sigma, \boldsymbol{p})$ is finite but distinct from (i.e. greater than) one, then clearly $\boldsymbol{c}(\boldsymbol{p})$ is non-uniquely locally identifiable. It is common to take $\boldsymbol{c}$ to be $\boldsymbol{c}(\boldsymbol{p}) = p_i$ (a scalar), to determine the identifiability of the $i$-th component parameter of $\boldsymbol{p}$, where the abbreviated notation $\mathcal{S}_i$ will be used here for the corresponding $\mathcal{S}_{\boldsymbol{c}}$. If $\mathcal{S}_i(\Sigma, \boldsymbol{p})$ contains a closed interval containing $p_i$ then clearly the model is unidentifiable at $p_i$. With more assumptions on the exhaustive summary and/or parameter space, stronger statements can be made.

If it is intended to only use a model with a single fixed input, then it is better to move that input into the model equations directly and to consider the model as an uncontrolled system. This is because, as can be seen from the definition, structural identifiability uses information from all permissible inputs, not a single fixed input. Some authors refer to the concept of a "persistently exciting input" in an attempt to simplify matters as they then need to consider only one input.

Now suppose that $\boldsymbol{p}$ is not fixed. The above definitions become their *structural* counterparts when the conditions are required to hold for almost every $\boldsymbol{p} \in \Omega$:

**Definition 4** (Structural identifiability of parameter combinations)**.** A parameter combination $\boldsymbol{c}(\boldsymbol{p})$ of $\boldsymbol{p}$ is said to be

- *structurally globally identifiable (SGI)* iff for almost every $\boldsymbol{p} \in \Omega$,

$$\text{if } \overline{\boldsymbol{p}} \in \Omega \text{ and } \hat{\Sigma}(\boldsymbol{p}) = \hat{\Sigma}(\overline{\boldsymbol{p}}) \text{ then } \boldsymbol{c}(\overline{\boldsymbol{p}}) = \boldsymbol{c}(\boldsymbol{p}); \tag{2.5}$$

- (at least) *structurally locally identifiable (SLI)* iff for almost every $\boldsymbol{p} \in \Omega$, there exists a neighbourhood $N(\boldsymbol{p}) \subseteq \Omega$ of vectors around $\boldsymbol{p}$ such that

$$\text{if } \overline{\boldsymbol{p}} \in N(\boldsymbol{p}) \text{ and } \hat{\Sigma}(\boldsymbol{p}) = \hat{\Sigma}(\overline{\boldsymbol{p}}) \text{ then } \boldsymbol{c}(\overline{\boldsymbol{p}}) = \boldsymbol{c}(\boldsymbol{p}); \tag{2.6}$$

- *structurally unidentifiable (SUI)* otherwise.

**Definition 5** (Identifiability and structural identifiability of models)**.** The identifiability status of the whole model at $\boldsymbol{p}$ is defined to be the identifiability status of the identity parameter combination ($\boldsymbol{c} =$ identity in Definition 3) at $\boldsymbol{p}$.

Similarly, the structural identifiability status of the whole model is defined to be the structural identifiability status of the identity parameter combination in Definition 4.

Equivalently, this can be expressed in terms of the identifiability of the component parameters as follows (phrased only for structural identifiability for brevity):

- If **all** the component parameters in the model are structurally globally identifiable then the model itself is said to be structurally globally identifiable;
- The model is (at least) *structurally locally identifiable* iff **all** component parameters of $\boldsymbol{p}$ are (at least) structurally locally identifiable;
- The model is *structurally unidentifiable* iff **any** of the component parameters of $\boldsymbol{p}$ are structurally unidentifiable.

  (Some authors prefer the term *partially structurally identifiable* iff **some** but not all of the component parameters of $\boldsymbol{p}$ are structurally unidentifiable.)

A model is therefore structurally globally (locally) identifiable iff there exists a null set $E$ with $\hat{\Sigma}$ (locally) injective when restricted to $\Omega \setminus E$. By the definition of exhaustive summaries, the input/output map in the definitions of identifiability and structural identifiability can be replaced with any exhaustive summary for the model. Most structural identifiability methods work by producing a simpler exhaustive summary than the original input/output map, upon which simpler methods (such as the direct test method) can often be used to determine the generic injectivity and thus structural identifiability.

Instead of solving $\Sigma(\boldsymbol{p}) = \Sigma(\overline{\boldsymbol{p}})$ for $\overline{\boldsymbol{p}}$, it may sometimes be algebraically simpler to solve

$$\boldsymbol{\phi} = \Sigma(\boldsymbol{p}) \text{ for } \boldsymbol{p} \text{ when } \boldsymbol{\phi} \text{ is in the range of } \Sigma. \tag{2.7}$$

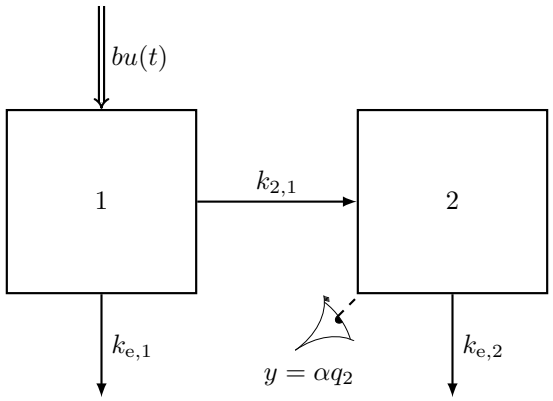
This alternative formulation is just another way of checking the injectivity of the exhaustive summary, and so the number of generic solutions

for $\boldsymbol{c}(\boldsymbol{p})$ still determines the structural identifiability of $\boldsymbol{c}(\boldsymbol{p})$.

Note that structural identifiability depends on each of the feasible parameter space (and the a priori information incorporated into it), the system model structure, the observations, the initial conditions, and the admissible inputs.

### 2.5.3. Example to illustrate definitions

It will of course aid the reader to see examples of how the above definitions work in practice. This subsection will illustrate those definitions in the context of the well-known [54] two-compartment model depicted in Figure 2.4.




▲ Figure 2.4: Two compartment model used to illustrate structural identifiability concepts and for testing SMC sampler

This model is described by the following equations:

$$\begin{cases} q_1'(t) &= bu(t) - (k_{e,1} + k_{2,1})q_1(t) \\ q_2'(t) &= k_{2,1}q_1(t) - k_{e,2}q_2(t) \\ q_1(0) &= 0 \\ q_2(0) &= 0 \\ y(t) &= \alpha q_2(t) \end{cases} \tag{2.8}$$

for $t \in \mathbb{T} := \mathbb{R}^+$.

Write $\boldsymbol{p} = (\alpha, b, k_{2,1}, k_{\mathrm{e},1}, k_{\mathrm{e},2})$ so the input/output equation is given by :

$$\hat{\Sigma}(\boldsymbol{p}) = \left( t \mapsto \alpha b k_{2,1} \frac{e^{-k_{\mathrm{e},2}t} - e^{-(k_{2,1}+k_{\mathrm{e},1})t}}{k_{2,1} + k_{\mathrm{e},1} - k_{\mathrm{e},2}} \right) \otimes u, \tag{2.9}$$

where $\otimes$ is the convolution operator.

The Laplace transform of the above expression is

$$s \mapsto \alpha b k_{2,1} \frac{1}{(s + k_{2,1} + k_{\mathrm{e},1})(s + k_{\mathrm{e},2})} u(s), \tag{2.10}$$

and so the Laplace transfer function is

$$s \mapsto \alpha b k_{2,1} \frac{1}{(s + k_{2,1} + k_{\mathrm{e},1})(s + k_{\mathrm{e},2})}. \tag{2.11}$$

Noting that there is never any pole-zero cancellation, the coefficients of the Laplace transfer function (i.e. with the denominator expanded) are

$$\alpha b k_{2,1}, \quad 1, \quad k_{2,1} + k_{\mathrm{e},1} + k_{\mathrm{e},2}, \quad (k_{2,1} + k_{\mathrm{e},1}) k_{\mathrm{e},2}. \tag{2.12}$$

Indistinguishability of the parameters under the input/output equation is equivalent to indistinguishability of the coefficients of the Laplace transfer function, within the meaning of Definition 1. Hence, the coefficients form an exhaustive parameter summary for the model according to Definition 2. Call the coefficients $\Sigma(\boldsymbol{p})$ and write $\overline{\boldsymbol{p}} = (\overline{\alpha}, \overline{b}, \overline{k}_{2,1}, \overline{k}_{\mathrm{e},1}, \overline{k}_{\mathrm{e},2})$.

From the exhaustive parameter summary, note that if $r > 0$ and $\overline{\boldsymbol{p}} = (\alpha r, \frac{b}{r}, k_{2,1}, k_{\mathrm{e},1}, k_{\mathrm{e},2})$ then $\Sigma(\boldsymbol{p}) = \Sigma(\overline{\boldsymbol{p}})$. As $\forall \quad N(\boldsymbol{p}), \ \exists \quad r > 0$ such that $(\alpha r, \frac{b}{r}, k_{2,1}, k_{\mathrm{e},1}, k_{\mathrm{e},2}) \in N(\boldsymbol{p})$, this means (through Definition 4) that component parameters $\alpha$ and $b$ are unidentifiable at any $\boldsymbol{p}$, and so in the more generic sense of structural identifiability, are also structurally unidentifiable.

Now note that

$$\Sigma(\boldsymbol{p}) = \Sigma(\overline{\boldsymbol{p}}) \implies \overline{k}_{\mathrm{e},1} = r \text{ and either}$$
$$\begin{cases} \overline{\alpha}\overline{b} = \alpha b \frac{k_{2,1}}{k_{2,1}+k_{\mathrm{e},1}-r}, \quad \overline{k}_{2,1} = k_{2,1} + k_{\mathrm{e},1} - r, \quad \overline{k}_{\mathrm{e},2} = k_{\mathrm{e},2} & \text{or} \\ \overline{\alpha}\overline{b} = \alpha b \frac{k_{2,1}}{k_{\mathrm{e},2}-r}, \qquad \overline{k}_{2,1} = k_{\mathrm{e},2} - r, \qquad \overline{k}_{\mathrm{e},2} = k_{2,1} + k_{\mathrm{e},1} \end{cases}$$

for some $r > 0$. In particular, if $\Omega = (0, \infty)^5$,

$$\begin{aligned} \mathcal{S}_\alpha = \mathcal{S}_b = \mathcal{S}_{k_{2,1}} = \mathcal{S}_{k_{\mathrm{e},1}} &= (0, \infty) \\ \mathcal{S}_{k_{\mathrm{e},2}} &= \left\{k_{\mathrm{e},2},\ k_{2,1} + k_{\mathrm{e},1}\right\} \end{aligned}$$

so the only component parameter of the model that is identifiable is $k_{\mathrm{e},2}$, which is locally identifiable at $\boldsymbol{p}$ (non-uniquely unless $k_{\mathrm{e},2} = k_{2,1} + k_{\mathrm{e},1}$) and is non-uniquely structurally locally identifiable. The remaining parameters are (structurally) unidentifiable.

If instead $k_{\mathrm{e},1} = 0$ is known, then the indistinguishability conditions reduce to

$$\begin{cases} \overline{\alpha}\overline{b} = \alpha b, \qquad \overline{k}_{2,1} = k_{2,1}, \quad \overline{k}_{\mathrm{e},2} = k_{\mathrm{e},2} & \text{or} \\ \overline{\alpha}\overline{b} = \alpha b \frac{k_{2,1}}{k_{\mathrm{e},2}}, \qquad \overline{k}_{2,1} = k_{\mathrm{e},2}, \qquad \overline{k}_{\mathrm{e},2} = k_{2,1}, \end{cases} \tag{2.13}$$

and so

$$\begin{aligned} \mathcal{S}_\alpha = \mathcal{S}_b &= (0, \infty) \\ \mathcal{S}_{\alpha b} &= \left\{\alpha b,\ \alpha b \frac{k_{2,1}}{k_{\mathrm{e},2}}\right\} \\ \mathcal{S}_{k_{2,1}} &= \left\{k_{2,1},\ k_{\mathrm{e},2}\right\} \\ \mathcal{S}_{k_{\mathrm{e},2}} &= \left\{k_{\mathrm{e},2},\ k_{2,1}\right\}. \end{aligned}$$

So, consider in turn the parameter combinations $\boldsymbol{c}(\boldsymbol{p}) = \alpha b$, $\boldsymbol{c}(\boldsymbol{p}) = k_{2,1}$,

and $\boldsymbol{c}(\boldsymbol{p}) = k_{e,2}$. If $k_{2,1} = k_{e,2}$ then each of these parameter combinations are globally identifiable at $\boldsymbol{p}$, else they are only non-uniquely locally identifiable at $\boldsymbol{p}$. Considering now the structural sense, where almost every parameter has to be considered rather than a specific $\boldsymbol{p}$, all the above parameter combinations are non-uniquely structurally locally identifiable.

The model (with or without the assumption that $k_{e,1} = 0$) is structurally unidentifiable (Definition 5) because the component parameters $\alpha$ and $b$ are structurally unidentifiable. Properly, in order to consider the model (with the assumption that $k_{e,1} = 0$) to be identifiable, the model ought first to be re-parameterised so that its state equations involve $\alpha$ and $b$ only through the product $\alpha b$ rather than through $\alpha$ and $b$ separately, but this re-parameterisation step is not of interest in this thesis: here it is enough to say that some parameters are considered to be combined, and under this assumption, to treat the model as non-uniquely structurally locally identifiable.

### 2.5.4. Abridged summary of available methods

There are a number of techniques available for determining the structural identifiability of linear systems. The Laplace transform approach was already introduced by way of the example in the previous subsection. This, and other approaches, were discussed in more detail in Hall [1], and are not repeated here. Techniques that are applicable to the more general case of non-linear systems (and thus also to linear systems) are of interest in this work. Such techniques include differential algebra approaches [62, 63, 64, 65], the similarity transformation approach (for compartmental models) [66], the Taylor series approach [67] (discussed in more detail in the next section), and local methods such as the Pohjanpalo rank test [68], the Exact Arithmetic Rank (EAR) approach [69]. Many of these approaches have been tried by the author but only the Taylor

series approach is discussed in this work as it is the simplest of the few approaches that were tractable for the models presented herein.

### 2.5.5. Taylor series coefficient approach

The Taylor series coefficients technique for structural identifiability can be used for both linear and non-linear systems. It is relatively easy to calculate the Taylor series coefficients for most models, though using these coefficients to solve the structural identifiability problem can sometimes prove to be computationally intractable even when other methods are tractable.

Clearly, for given $t_*, \boldsymbol{u}, \boldsymbol{p}, \overline{\boldsymbol{p}}$, the following implication holds:

$$\begin{aligned}&\Big(\forall\ \ t, \boldsymbol{y}(t, \boldsymbol{u}, \boldsymbol{p}) = \boldsymbol{y}(t, \boldsymbol{u}, \overline{\boldsymbol{p}})\Big) \implies \\ &\quad \text{for all } k \in \mathbb{N}_0 \text{ s.t. } \boldsymbol{y}^{(k)} \text{ is defined}, \boldsymbol{y}^{(k)}(t_*, \boldsymbol{u}, \boldsymbol{p}) = \boldsymbol{y}^{(k)}(t_*, \boldsymbol{u}, \overline{\boldsymbol{p}}).\end{aligned} \tag{2.14}$$

If $\boldsymbol{y}$ is analytic at $t_*$ then the reverse implication is also true. Therefore, if $\boldsymbol{y}$ is analytic at $t_*$, the collection of Taylor series coefficients $(\boldsymbol{y}^{(k)}(t_*, \boldsymbol{u}, \boldsymbol{p}))_k$ form an exhaustive summary.

If $t_* = 0$ then the corresponding Taylor series coefficients of a model in the form (2.2) can (in principle) be determined algebraically in terms of the (symbolic) parameters and the initial value of the input functions.

If $\boldsymbol{y}$ is not analytic at 0 then the Taylor series coefficients can be used to give a sufficient but not necessary method for determining structural identifiability [67, p. 23]: generic injectivity of the Taylor series coefficients are sufficient but not necessary for structural identifiability.

Of course the Taylor series coefficients are in general an infinite series which would not be practical to work with, but injectivity when restricted to any $n$ components implies injectivity for the whole series. Therefore, truncating the Taylor series coefficients again gives a sufficient but not generally necessary method for determining structural identifiability.

In certain settings, it can be shown that truncation to a certain num-

ber of coefficients does again give a necessary and sufficient criterion.

### 2.5.6. Non-structural identifiability

Although structural identifiability is a necessary prerequisite to conduct well-posed parameter inference on a model, it is not sufficient. It is also often the case that the structural model is too complicated or cannot be expressed in one of the required algebraic forms in order to conduct a complete structural identifiability analysis. Clearly in these situations, some evaluation of the identifiability of the model would be useful.

There are methods known as numerical identifiability and practical (aka data-based [70]) identifiability methods that can be applied to a wider class of models, though they are only able to produce local outcomes. Instead of an algebraic or structural analysis of the model equations, a numerical or practical identifiability analysis uses numerical methods to analyse the model, at specific numerical instances or collections of parameter values. There is no known agreed distinction between the terms "numerical identifiability" and "practical identifiability", but the following convention will be used here: numerical identifiability will be used to refer to identifiability using arbitrarily dense and noise-free data generated from simulating the model under study, while practical identifiability will refer to the situation where the data are not arbitrarily dense, and/or not noise-free, and/or not simulated data. When simulated data are used, the recovered parameters can be compared against the original parameter values that were used to generate the simulated data. It is generally a good idea to conduct both numerical and practical identifiability analyses: the numerical analysis with dense noise-free data and the practical analysis with data sampling times and signal to noise ratio corresponding to what one expects to see from a real data set, as well as later with an actual real data set. Of course, if practical identifiability is confirmed for a number of distinct parameter sets and model inputs,

then it may not be necessary to consider numerical identifiability as well, but a failure of practical identifiability may be caused by a failure of numerical identifiability rather than the data simply being too sparse or too noisy.

These types of analysis can provide insight to the kind of identifiability problems in the sense of Definition 3, but not in the *structural* sense of Definition 4, because the "almost everywhere" condition within the definition of structural identifiability cannot be carried over to the numeric case (specific instances of parameters always lie within a set of zero measure). However, numerical and practical identifiability have to be considered with respect to fixed inputs, and cannot be applied to all inputs as is required by Definition 3, unless the set of inputs is a singletonr finite set (though a persistently exciting input may be of use here if available). Numerical and practical identifiability will together be referred to as "non-structural identifiability" here. Usefully, numerical and practical identifiability can often allow recovery of functional equations along which the likelihood stays constant [71], and it may then be possible to use those functional equations to test the identifiability of the model in a structural sense. As mentioned previously, it is much easier to see if two given generic parameter sets are indistinguishable than to find all indistinguishable parameter sets, and these identifiability methods can help construct candidate parameter sets to test for indistinguishability.

For systems that are structurally globally identifiable, a failure of numerical or practical identifiability will be due to limitations on the data and/or the relationship between the data and the model, not the structure of the model itself. For systems that are not known to be structurally globally identifiable, a failure of numerical or practical identifiability may be due to lack of structural global identifiability in addition to or instead of the above reasons.

Practical and numerical identifiability generally involve studying the shape of the likelihood surface. Raue et al. [71] defines these kinds of iden-

tifiability in relation to a parameter in terms of the size of a likelihood-based confidence interval for that parameter: the confidence interval for a numerically unidentifiable parameter (in log space) is the whole of $\mathbb{R}$, and extends to infinity in one direction for a practically unidentifiable parameter.

One widely used approach for considering numerical or practical identifiability is the profile likelihood approach [71]. In case a Bayesian analysis is conducted, this can be replaced with a "profile posterior". This approach can be computationally costly, since it involves solving high-dimensional optimisation problems on a fine mesh of points, but can be parallelised and the resultant information is easy to interpret. Unfortunately, it is difficult to reuse information from solving one optimisation problem when solving an adjacent problem (where one of the parameters has changed by a small value). Of course, in this case, another mode of the likelihood may become dominant. A potentially more efficient method for calculating profile likelihoods is proposed later in Section 3.2.

Practical identifiability problems, in the context of solving an optimisation problem, include the case that no solutions are found by the optimiser, or where other solutions are returned by the optimiser in addition to or instead of any global optima. In the context of posterior distributions, practical identifiability problems may occur where the posterior has two or more modes that are not sufficiently separated to be distinguished. In a practical identifiability analysis, the data sampling frequency can be studied too.

Clearly, it may be useful to determine the practical identifiability of a model in addition to other forms of identifiability. Therefore, it is suggested to determine the structural identifiability of a model, or if not possible then the numerical identifiability, followed by the practical identifiability using synthetic data of a similar nature to that seen in real data.

## 2.6. Existing models for ARS

Many existing pharmacokinetic studies for artemisinins have been conducted over the last couple of decades, and have successfully provided some insights into the absorption, elimination and/or multiple dosing behaviour of these drugs, and the covariates that influence these, including the author's own attempts in Hall [1] and Hall et al. [2]. Some studies have restricted their interest to either healthy subjects, uncomplicated malaria, or severe malaria, and either children, adults or pregnant women, while others have been designed specifically to consider the differences between some of these groups. Each study focusses on a specific artemisinin derivative or derivatives, and a specific route or routes of administration, either alone or in combination with other antimalarial agents.

As mentioned in section 2.4, mechanistic models have the most utility. Of the existing models that used mechanistic approaches, some have been used to analyse the effects of differing dosing regimens in different contexts, including cases where the malaria has developed resistance to this class of drugs. They range from being very simple, e.g. with linear absorption and exponential elimination as in Saralamba et al. [37], to being quite complicated, e.g. involving 9 compartments as in Gordi et al. [72], and of various complexities in between, e.g. 4 compartments as in Tan et al. [73].

A goal of this work was to develop a model incorporating the parasite lifecycle and stage-specific effects of the artemisinins. There are a few existing models of parasite lifecycle, either alone or as part of larger pharmacokinetic and pharmacodynamic models.

One such model is from Gravenor et al. [74], which splits the lifecycle into a variable number of compartments of equal duration, each potentially with their own death rates, but these death rates do not relate to any pharmacokinetic data. The model allows for changes in the degree

of synchronicity as the length of the parasite lifecycle is not fixed, but is described by a gamma distribution with a mean of 48 hours and a standard deviation inversely proportional to the square root of the number of compartments used.

A slightly similar model is Svensson et al. [47] which fixes the number of age classes at 4, though this does not include stage-specific effects (but could be adapted to do so, and indeed the authors note the same). In order to account for parasite kill continuing beyond the decline of artemisinin plasma concentrations, the model uses a intermediary pool of damaged parasites still in general circulation, which are later removed by the spleen in a first-order process. Rate of parasite damage was described by first-order but time-varying dynamics, with the rate depending on the time above a minimum inhibitory concentration.

The only known model incorporating stage-specific effects is from Saralamba et al. [37]. The pharmacokinetic aspect of the model is non-mechanistic with linear absorption and exponential elimination with a single exponential term, which was well suited for their data-fitting application but cannot be used to experiment with different dosing strategies. The model makes many useful simplifying assumptions to aid in their data-fitting application, but which are less suitable for a predictive model. The model advances in discrete-time, assumes a normal age distribution of parasites and also assumes that all doses have identical pharmacokinetic profiles. While this model showed excellent performance despite its relative simplicity, for some patients there were clearly systematic deviations from the model, which might be capable of being explained by a more advanced model.

Most pharmacometric studies of artemisinin have shown that there are vastly different volumes of distribution of artemisinin drugs between healthy and infected persons, and the same is true even considering just infected persons. The author speculates that this could possibly be due to the take up in infected cells, and no model is currently known to exist

with the capability to explore this possibility.

It is submitted there is a need for a new model where the pharmacokinetics and pharmacodynamics are both captured in detail, including the interaction between the two, and one aim of this work is to develop such a continuous-time model of the stage-specific effects of ARS and DHA. The model previously published by the author [1, 2] was always intended as a building block towards such a more complete model, and will be revisited in subsection 2.6.1.

As has also been noted by Simpson et al. [75], few existing studies formally estimated pharmacokinetic or pharmacodynamic parameters within a rigorous (preferably Bayesian) statistical framework using individual patient data. The author agrees that a model's behaviour with respect to its parameters should be fully explored, and the sensitivity of the model to its parameters be determined, including whether the parameters are estimable with sufficient precision. Indeed, this latter point is related to the concepts of structural and practical identifiability, which were important issues highlighted by the author in Hall [1] and Hall et al. [2], yet those (still) appear to be the only models applied to artemisinins that have been assessed for structural identifiability to date. One potential problem with developing a new complete model as desired above is that it is likely to present unique mathematical and statistical challenges due to its complexity, and existing structural identifiability techniques may struggle in analysing such a model.

Finally, as noted above, Simpson et al. [75] expressed a preference for conducting model fitting via a Bayesian analysis, another preference shared by the author, despite conducting a non-Bayesian analysis in Hall [1] and Hall et al. [2].

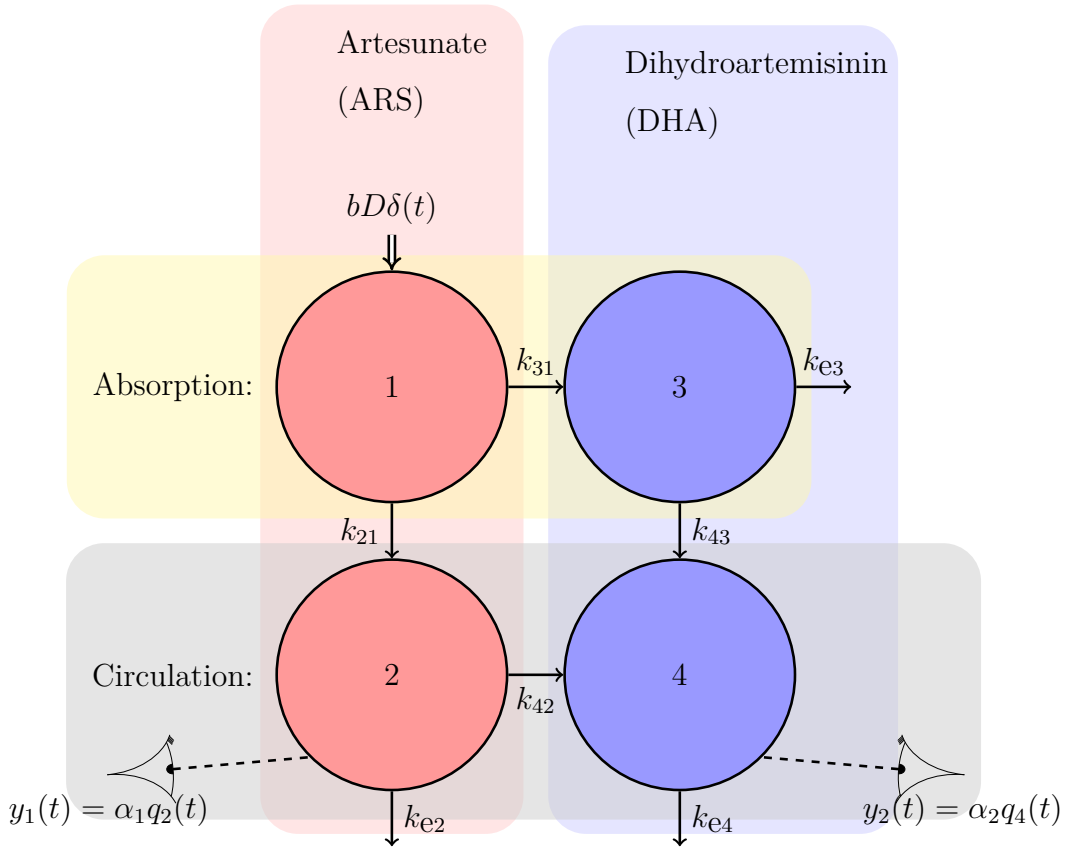
Therefore, the issues of structural and practical identifiability and Bayesian analyses are explored in this work.

### 2.6.1. Four compartment pharmacokinetic model

This section presents the model used by author previously in Hall et al. [2], which was in turn based on one of the models presented in Hall [1]. Further analysis of this model is presented in section 4.1. This model is also used as a base model for the novel model introduced in section 4.2.

#### 2.6.1.1. Model structure

A relatively simple coupled mechanistic model was developed for the pharmacokinetics of orally-administered ARS and its principle metabolite DHA, for situations where blood plasma concentrations of both are observed, and is depicted in Figure 2.5. It consists of four linked compartments, with the parent drug and its metabolite each represented by two compartments: an absorption (gut) compartment and a circulation/plasma compartment. The absorption compartments account for the delay in the drug and metabolite reaching the circulation (and site of measurement) due to the oral route of administration.




▲ Figure 2.5: System diagram of the four compartment model for ARS (left part) and DHA (right part). Upper compartments represent the absorption compartments (unobserved). Lower compartments are the circulation compartments (observed).

The administered oral dose of ARS is considered as a bolus (impulsive) input into its absorption compartment (1 in the diagram). To account for bioavailability, a fraction $b$ of the administered dose $D$ is assumed to reach the systemic circulation. The dose $D$ is prescribed in proportion to the body weight of the patient, and so taken in units of nmol per kg.

Once in the system, ARS is either irreversibly metabolised into DHA (compartment 3) prior to reaching the circulation (compartment 4), or is absorbed into the circulation (compartment 2) and subsequently metabolised (compartment 4 again). Elimination can occur from any com-

partment except the input compartment (compartment 1), and can be caused by either excretion from the body or further metabolism into inactive metabolites which are not of interest.

Observations are made of the drug concentrations in the circulation compartments, with observation gains $\alpha_1$ for ARS ($y_1$) and $\alpha_2$ for DHA ($y_2$). These parameters incorporate the volumes of distribution of the respective drugs. As is standard for the purposes of assessing the identifiability of the structural model, "ideal observation" assumptions are made, namely that observations are available continually over the entire infinite time horizon, and further are available without error. These two assumptions are relaxed later when dealing with the problem of parameter estimation from data.

Note that because metabolism of ARS into DHA takes place in the liver as well as in esterases, metabolism can occur before presentation in the observed circulation compartments. Indeed, in concentration–time profiles of malaria patients (e.g. those analysed in this work), large quantities of DHA are observed in the blood plasma prior to those of ARS, which cannot be attributed solely to being artefacts of differing observation gains (or otherwise to quantification limits). Hence, the presence of compartment 3 is crucial to capture the metabolism-before-absorption route that ARS can take.

The differential equation characterisation of the model is given, for $t \in [0, \infty)$ describing the time in hours since drug administration, by

$$\begin{cases} \boldsymbol{q}'(t) &= \mathbf{A}\boldsymbol{q}(t) + \mathbf{B}\boldsymbol{u}(t) \\ \boldsymbol{q}(0^+) &= \boldsymbol{q}_0 \\ \boldsymbol{y}(t) &= \mathbf{C}\boldsymbol{q}(t). \end{cases} \tag{2.15}$$

Here, $\boldsymbol{q} = \begin{pmatrix} q_1 & q_2 & q_3 & q_4 \end{pmatrix}^{\mathrm{T}}$ represents the state vector of the system model, where each $q_i$ denotes the quantity of the respective drug in compartment $i$, $\boldsymbol{u}(t) = \begin{pmatrix} D\delta(t) & 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}$ denotes the model input and $\boldsymbol{q}_0 = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}$ the initial condition, $\boldsymbol{y}$ denotes the vector-valued

observation function, and the model matrices are

$$\mathbf{A} = \begin{pmatrix} -(k_{21}+k_{31}) & 0 & 0 & 0 \\ k_{21} & -(k_{42}+k_{e2}) & 0 & 0 \\ k_{31} & 0 & -(k_{43}+k_{e3}) & 0 \\ 0 & k_{42} & k_{43} & -k_{e4} \end{pmatrix}, \tag{2.16a}$$

$$\mathbf{B} = \begin{pmatrix} b \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 0 & \alpha_1 & 0 & 0 \\ 0 & 0 & 0 & \alpha_2 \end{pmatrix}. \tag{2.16b}$$

Note that there are different ways to parameterise $\boldsymbol{u}$, $\boldsymbol{q}_0$ and $\mathbf{B}$. The parameterisation used here has been chosen as it more clearly corresponds to the mechanistic concepts.

As mentioned previously, due to the difference in the molecular weights of the parent drug and the metabolite, the $q_i$ are considered in units of molar mass, per kilogram of patient body weight (nmol/kg). Observations, which are concentrations, are assumed to be in units of nmol/l. The observation gains $\alpha_1$ and $\alpha_2$ therefore have units of kg/l, but the volumes of distribution are generally assumed to scale approximately linearly with patient body weight, hence the reason that the dosing is calculated in those terms. (However, there is some debate about whether or not allometric scaling might be more appropriate.)

All flows (absorption, metabolism and elimination) are assumed to be first-order and linear, with rate constants $k_{ij}$ (denoting the flow rate constant to compartment $i$ from compartment $j$, or to the environment when $i = \mathrm{e}$) time-invariant and specified in units of per hour (which are standard units for artemisinin drugs). Note that conversion into inactive unmeasured metabolites and excretion from the body are considered as flows to the environment with respect to the system model.

The system of equations (2.15), with $\boldsymbol{u}(t)$ and $\boldsymbol{q}_0$ as described above,

can easily be solved analytically to yield:

$$\boldsymbol{q}(t) = bDe^{\mathbf{A}t}\,, \quad \boldsymbol{y}(t) = \mathbf{C}\boldsymbol{q}(t)\,. \tag{2.17}$$

The solution for the state variables is thus

$$\begin{aligned}
q_1(t) &= bDe^{-(k_{21}+k_{31})t} \\
q_2(t) &= \frac{bDk_{21}\left(e^{-(k_{42}+k_{e2})t} - e^{-(k_{21}+k_{31})t}\right)}{k_{21}+k_{31}-k_{42}-k_{e2}} \\
q_3(t) &= \frac{bDk_{31}\left(e^{-(k_{43}+k_{e3})t} - e^{-(k_{21}+k_{31})t}\right)}{k_{21}+k_{31}-k_{43}-k_{e3}} \\
q_4(t) &= bD\left(e^{-(k_{21}+k_{31})t}\right. \\
&\quad \frac{k_{21}^2k_{42}+k_{31}k_{43}(k_{31}-k_{42}-k_{e2})+k_{21}(k_{31}(k_{42}+k_{43})-k_{42}(k_{43}+k_{e3}))}{(k_{21}+k_{31}-k_{42}-k_{e2})(k_{21}+k_{31}-k_{43}-k_{e3})(k_{21}+k_{31}-k_{e4})} \\
&\quad - \frac{e^{-(k_{42}+k_{e2})t}k_{21}k_{42}}{(k_{21}+k_{31}-k_{42}-k_{e2})(k_{42}+k_{e2}-k_{e4})} \\
&\quad - \frac{e^{-(k_{43}+k_{e3})t}k_{31}k_{43}}{(k_{21}+k_{31}-k_{43}-k_{e3})(k_{43}+k_{e3}-k_{e4})} \\
&\quad \left. + \frac{e^{-k_{e4}t}(k_{31}k_{43}(k_{42}+k_{e2}-k_{e4})+k_{21}k_{42}(k_{43}+k_{e3}-k_{e4}))}{(k_{21}+k_{31}-k_{e4})(k_{42}+k_{e2}-k_{e4})(k_{43}+k_{e3}-k_{e4})}\right).
\end{aligned} \tag{2.18}$$

These equations are already seen to be somewhat unwieldy despite the relative simplicity of the system model.

#### 2.6.1.2. Model analysis

This model was analysed in Hall [1]. This analysis included a structural identifiability analysis using the Laplace transform method [54] and highlighted problems with the Taylor series method [67] and similarity transformation method [53]. The model was found to be structurally unidentifiable, and so three constraints were identified that together were sufficient to make the model structurally globally identifiable.

This model was also analysed further in Hall et al. [2]. These further results include use of a slightly different form of those constraints, and use of a more systematic approach of testing combinations of the constraints. Some of this as well as a further improved analysis, including statistical analysis, of this model is presented in this thesis in section 4.1.

## 2.7. Statistical methods, inference and quantification of results

### 2.7.1. Bayesian vs. non-Bayesian frameworks

In both Bayesian and classical statistical frameworks, it is normal to have a likelihood function quantifying the support that the given data have for arbitrary model parameters. In the classical framework, a point estimate of the parameters (the maximum likelihood estimate) is generally sought. This can be problematic due to practical difficulties in confirming whether an estimate is genuinely the maximum likelihood estimate (and indeed whether such an estimate is even unique) since global optimisation algorithms cannot identify the global maximum with perfect certainty, and disallows identification or use of a secondary mode of the likelihood function, although that secondary mode may sometimes be more plausible or more compatible with other estimates (e.g. from different studies or from different subjects), and may only have a slightly worse likelihood than the global maximum. In a Bayesian framework, all the information known about the parameters *a priori* is encoded in the form of a prior distribution, and this is combined with the likelihood, which contains all the information the data tell us about the parameters, to give a posterior distribution, which assigns a probability to each potential parameter vector. Such probabilities quantify the degree of belief that the parameter vector matches the data.

Many authors advocate use of Bayesian posterior distributions in scientific applications (e.g. [75]). These contain much more information, especially regarding uncertainty, than simply determining point estimates or confidence bands for the parameters and model predictions. Correspondingly, Bayesian information is much more useful to a practitioner, and allows a more thorough exploration of the relevant scenarios. Bayesian methods also allow borrowing of information across subjects

in mixed effects situations, which would not be so readily possible in a classical framework. (Mixed effects models are a topic for future work and are not discussed in this thesis.)

However, calculating a probability distribution instead of just a point estimate involves much more computation and different assumptions. One immediate difference is that instead of simply specifying a feasible parameter region, it is necessary to quantify a priori how feasible each particular parameter combination is, and this is not always straightforward. This is known as specifying a prior distribution on the parameters. Priors can be informative or uninformative, depending on what information is known a priori and how many assumptions one wants to make at this stage. If information is known from other studies, then including that information as a prior assumption essentially prevents confirmation of that same information. Potential problems can arise when little is known (or incorporated) in the prior distribution, for example if no parameter may be *a priori* unbounded, as a proper prior is required to integrate to a total probability of one. Additionally, priors are not invariant under transformations and so, for example, a prior that may appear to be uninformative may in fact be significantly discriminatory towards a different parameterisation.

Priors also have an effect on identifiability: if a parameter is structurally or practically unidentifiable but the corresponding prior has (for example) a single point with maximal density, then that single point will generally be identified, though this depends on the exact nature of the practical unidentifiability. Indeed, many people impose certain priors deliberately so that they get an identifiable model, but naturally this should only be used to resolve practical identifiability concerns rather than structural ones. It is also possible for a prior to render an otherwise structurally identifiable model into an unidentifiable model, but this is pathological and unlikely to occur in practice without a specific aim to set up such a situation.

Throughout this work, proper uniform priors with finite support are used.

In quantifying the results from a Bayesian analysis, it is largely a case of expressing or otherwise summarising the posterior distribution and its marginals, but in a non-Bayesian analysis, it is common to have to conduct further calculations under further (and possibly exaggerated) assumptions, for example an asymptotic variance-covariance matrix (see subsection 2.7.4).

### 2.7.2. Likelihood functions

Unlike for the structural identifiability analysis, the observations are now assumed to be finite in number and collected at discrete times. Selection of an error model and a predictive model determines the likelihood function. Only observation error is considered, as error resulting from model misspecification is assumed to be dominated by observation error.

So, let $y_i$ denote the $i$th observation, and $h_i(\boldsymbol{\theta})$ denote the corresponding model prediction under parameters $\boldsymbol{\theta}$. Then,

$$\epsilon_i := \boldsymbol{r}(y_i, h_i(\boldsymbol{\theta})),$$

denotes the (modified) model residual under parameters $\boldsymbol{\theta}$ for observation $i$.

Due to the properties of the drug assays used (discussed in section 2.3), it is assumed that the model residuals for drug concentrations are normally distributed, with mean zero, and variance proportional to the model predicted value. For simplicity, the same is assumed for parasite observations. Note that this error model does not account for the fact that the observed concentrations will always be positive, but is nevertheless convenient to work with. For both the drug assays and the model residuals, other distributional assumptions could be made, such as the Laplace distribution (which instead of penalising a large squared

difference from the mean, penalises a large absolute difference from the mean, and so has a sharper peak, fatter tails and is more tolerant to outliers) or Student's $t$-distribution (which maintains a similar shape to the normal distribution but again has fatter tails). These other distributions are relatively easy to incorporate into a Bayesian analysis of the kind conducted here, though no comparison of the effect of these different distributions were made here: the normal distribution was used throughout. The results section later discusses whether a different choice of distribution might ultimately have been beneficial.

In symbols,

$$\epsilon_i \sim N(0, (\delta_i h_i(\boldsymbol{\theta}))^2). \tag{2.19}$$

The drug assays have a coefficient of variation reported at 15%, so $\delta_i = 0.15$.

It is further assumed that the observation errors for observations at different times are independent. (This assumption may not be realistic but is felt to be a good starting point in the absence of any prior information to the contrary.) It is also assumed that there is no correlation between observation errors of drug concentrations and parasitaemia. Observation errors for ARS and DHA observations obtained at the same time are assumed to be correlated with correlation parameter $\rho$ unknown.

It is convenient to view the $y_i$ as forming a one-dimensional vector. Write $\boldsymbol{y}$ for the data and $\boldsymbol{h}(\boldsymbol{\theta})$ for their model predictions. The above specification gives rise to the following log-likelihood function, defined up to an additive arbitrary constant:

$$\ell\Big(\boldsymbol{\theta}|\boldsymbol{y}\Big) = -\frac{1}{2}\left(\log\det \mathbf{V}(\boldsymbol{\theta}) + \underbrace{\Big(\boldsymbol{r}(\boldsymbol{y}, \boldsymbol{h}(\boldsymbol{\theta}))\Big)^{\mathrm{T}} \mathbf{V}(\boldsymbol{\theta})^{-1} \Big(\boldsymbol{r}(\boldsymbol{y}, \boldsymbol{h}(\boldsymbol{\theta}))\Big)}_{\text{weighted residual sum of squares (WRSS)}}\right), \tag{2.20}$$

where $\mathbf{V}(\boldsymbol{\theta})$ is the weighting matrix with $(i,i)$-th element $\sigma_i^2$ and $(i,j)$-th element $\rho\sigma_i\sigma_j$ when $t_i = t_j, i \neq j$, and $i, j$ are observations of drug

concentrations, and 0 otherwise. Here, $\boldsymbol{r}(\boldsymbol{y}, \boldsymbol{h}(\boldsymbol{\theta}))$ is the modified residual according to the rules in the following section, which coincides with the standard residual $\boldsymbol{r}(\boldsymbol{y}, \boldsymbol{x}) = \boldsymbol{y} - \boldsymbol{x}$ when neither argument is below the applicable censoring thresholds.

### 2.7.3. Censoring

In both datasets used in this work, data of drug concentrations below the limit of quantification are not reported with a numerical value but are simply reported as being below the limit of quantification. These data points are therefore said to be "left censored"; the values are below a known threshold but it is not known how far below the threshold. There are also some data points in the MORU_ARC3_PD dataset corresponding to drug concentrations so low that no distinct peak could be identified in the HPLC traces; these are reported as "no peak". This is essentially the same as detecting no drug in the sample, but the HPLC method does not have perfect sensitivity so it would not be quite accurate to report the drug concentration as 0. These data points are therefore also best considered as left censored, but with a lower threshold than the limit of quantification. Indeed, some assays report a limit of detection which is the appropriate threshold in this situation.

Clearly, data points that are below the limit of quantification or below the limit of detection still carry a significant amount of information, so this information needs to be incorporated into the statistical analysis rather than discarding such points. Nevertheless, these points must be treated distinctly from other data points where values are reported.

In this work, left censored data points will be handled by treating the model residual as zero where both the data point and the model prediction lie below the unobserved threshold. Any model predictions or data points below the threshold are treated as though they were at the threshold. If a model prediction and corresponding data point are

both left censored but with different thresholds then the point within the larger threshold is treated as though it was half way between the two thresholds.

### 2.7.4. Maximum likelihood estimation and uncertainty quantification

In a non-Bayesian framework, standard numerical optimisation methods are generally used to find a minimiser $\hat{\boldsymbol{\theta}}$ of the negative of the log-likelihood expression (hereafter referred to as the objective function), and the minimiser is used as "the best" estimate of the parameters, though the likelihood function is often multi-modal and it can be hard to ensure that a global minimum is found. Indeed, in non-practically identifiable situations, such a global minimum may not be unique and/or there may be multiple local minima that exacerbate the problem of finding global minima.

To attempt to quantify the uncertainty in the parameter estimates in a non-Bayesian analysis, the asymptotic (for a large number of observations) distribution of the parameter estimates can be estimated [76]. The asymptotic distribution of the parameter estimate $\hat{\boldsymbol{\theta}}$ is approximately $\text{MVN}(\boldsymbol{\theta}^*, \mathbf{C})$ where MVN denotes the multivariate normal family of distributions, $\boldsymbol{\theta}^*$ is the "true" value of $\boldsymbol{\theta}$ and the variance-covariance matrix $\mathbf{C}$ is described next. Consider the linear approximation to the dependence of the unweighted residuals on the parameters about the estimate $\hat{\boldsymbol{\theta}}$:

$$\mathbf{R}(\hat{\boldsymbol{\theta}}) = \frac{\partial}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \left( \boldsymbol{y} - \boldsymbol{h}(\boldsymbol{\theta}) \right) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \tag{2.21}$$

The inverse of the Fisher information matrix at $\hat{\boldsymbol{\theta}}$ provides an estimate $\mathbf{C}$ of the asymptotic variance-covariance matrix for $\hat{\boldsymbol{\theta}}$,

$$\mathbf{C} = \left( \mathbf{R}(\hat{\boldsymbol{\theta}})^{\mathrm{T}} \mathbf{V}(\hat{\boldsymbol{\theta}})^{-1} \mathbf{R}(\hat{\boldsymbol{\theta}}) \right)^{-1}. \tag{2.22}$$

The variance-covariance matrix $\mathbf{C}$ is easier to interpret by reporting the diagonal elements of $\mathbf{C}$ together with the correlation matrix formed by dividing the respective rows and columns by the square roots of these diagonal elements. This information fully specifies $\mathbf{C}$ but is easier to compare and contrast than $\mathbf{C}$ itself.

### 2.7.5. Goodness of fit statistics

Likelihood function values and WRSS values are not directly comparable between subjects, due to each data set having a different variation to begin with. Instead, the *(weighted) coefficient of determination* can be used. Loosely speaking, this expresses the variation in the data explained by the model as a ratio of the total variation present in the data, and is defined as

$$R^2 := 100\left(1 - \frac{\left(\boldsymbol{r}(\boldsymbol{y}, \boldsymbol{h}(\boldsymbol{\theta}))\right)^{\mathrm{T}} \mathbf{V}(\boldsymbol{p})^{-1} \left(\boldsymbol{r}(\boldsymbol{y}, \boldsymbol{h}(\boldsymbol{\theta}))\right)}{(\boldsymbol{y} - \overline{\boldsymbol{y}})^{\mathrm{T}} \mathbf{V}(\boldsymbol{p})^{-1} (\boldsymbol{y} - \overline{\boldsymbol{y}})}\right) \%, \qquad (2.23)$$

where the elements of $\overline{\boldsymbol{y}}$ are the average of the observed values for the corresponding curve.

The idea is that a larger coefficient of determination should indicate a better fit. However, a large value of this statistic does *not* necessarily correspond to a high likelihood, which is in some ways problematic as the objective is to maximise the likelihood not the coefficient of determination, but it does accord at least qualitatively with a visual analysis of fits. (Note that the baseline model is simply a mean model, which is not contained in the fitted model class, so the ANOVA interpretation of this statistic does not apply.)

### 2.7.6. Monte Carlo methods

It is well-known that many common problems that need to be solved in order to conduct a Bayesian analysis quickly become intractable to

do directly as the dimension increases. This is known as the curse of dimensionality. These common problems include sampling from a distribution, determining the normalising constant for a density, and/or determining the maximal density regions of a distribution. Therefore, high-dimensional problems require specially tailored analysis methods. Most of these methods work by approximate simulation, which can in principle operate to any desired degree of accuracy, and they are known as *Monte Carlo* methods.

A simple but efficient class of such methods are Sequential Monte Carlo (SMC) methods, used extensively here.

#### 2.7.6.1. Sequential Monte Carlo (SMC)

Suppose $\pi$ is a (possibly high-dimensional) distribution of interest (such as a posterior distribution), referred to hereafter as the target distribution. Instead of trying to work directly with this target distribution (due to the reasons identified above), it is often helpful to sequentially estimate some intermediaries $\pi_0, \ldots, \pi_{T-1}$, before finally arriving at the target distribution $\pi_T = \pi$ (where $T$ is not necessarily fixed and may instead be determined adaptively). In an SMC sampler, each distribution is approximated by a collection of weighted particles (random samples). In the application here, each particle is actually a parameter set from the parameter space of the model, and the weight represents (up to a constant of proportionality) the plausibility of that parameter set in terms of describing the observed data under the relevant assumptions, which can be compared relative to the weights of the other particles in the collection. Particles that have a relatively low weight and so contribute little to the approximation of the distribution may ultimately be discarded/replaced.

The idea is that successive distributions are 'close', so generating samples from $\pi_{n+1}$ is easier when samples from $\pi_n$ are already available. An SMC sampler exploits this when propagating the weighted particles from each distribution in the sequence to the next. Hence, a sample from

$\pi$ is constructed sequentially. SMC samplers are applied in situations where sampling from $\pi_0$ and then proceeding sequentially is much easier than sampling directly from $\pi_T$ (despite there being potentially many such intermediate distributions, i.e. with a large value of $T$). The starting distribution $\pi_0$ is chosen to be a distribution that is easy to sample from. Although there are other occasions where SMC samplers are useful, the above situation is the motivation behind using SMC in this work.

As SMC is a simulation method, it requires some tuning, diagnostics and validation, has a high computational cost and only produces approximate output. However, it is felt that each of these aspects are somewhat simpler or more satisfactory with SMC than other methods, and it is a proven and widely adopted and trusted method that can make use of modern computing facilities (which tend to involve parallel computation).

The parameter space $\Theta$ for each distribution in the sequence will be taken to be the same. This is not restrictive for the purposes here. It will be assumed that $\pi$ admits a density with respect to Lebesgue measure, as will always be the case for distributions used in this work, and making this assumption now provides for notational simplicity. For complete details on SMC in its original generality, see Del Moral et al. [77]. Only sufficient details are given in this thesis for the purposes required herein.

When describing distributions used in Monte Carlo methods, it is common to need to distinguish between the distributions themselves, their density functions, and because the latter are normally known only up to a constant of proportionality, kernels or intensities (unnormalised densities) for the distribution with respect to a particular constant of proportionality. Unnormalised densities may also be used where the normalising constant is not relevant (e.g. due to being multiplied by terms that are also unnormalised). In this work, for a distribution denoted by $\pi$, its density function will be denoted $[\pi]$ and an unnormalised density for it will be denoted by $[\![\pi]\!]$. Hopefully this convention will aid the reader

more than using different symbols for each of these functions would. The constant of proportionality is not directly of interest here and no special notation will be introduced for that.

In some applications, there will be a natural sequence of distributions to consider, for example observations could be incorporated sequentially, leading to a sequence of distributions referred to as a data tempered sequence. Such a sequence has the advantage that the intermediate distributions are meaningful; the posterior likelihood can be updated "on-line" and possibly in real-time as new observations arrive. However, in such a sequence, the number of intermediate distributions is constrained by the number of observations available and this may not be an ideal way to reach the target distribution, especially where the analysis is conducted "off-line", i.e. where all the observations are already available.

Here, all observations are already available for incorporation simultaneously rather than sequentially, and so distribution tempering is used to produce appropriate sequences of distributions. Use of such distribution tempering schemes within SMC has not been widely documented but is becoming more popular [78, 79, 80]. The sequence of distributions is taken as $(\pi_0, \ldots, \pi_t, \ldots, \pi_T)$ where each $\pi_t$ is defined by

$$[\![\pi_t]\!] = \boldsymbol{x} \mapsto \left([\![\pi]\!](\boldsymbol{x})\right)^{\phi_t}, \tag{2.24}$$

with $0 = \phi_0 < \cdots < \phi_T = 1$ and $\pi$ is the target distribution of interest (i.e. the posterior).

Naturally, for observations assumed to come from a single fixed model, $\pi$ is taken to be the posterior distribution of the parameters given the data, derived from Bayes' theorem as the prior distribution $\nu$ times the likelihood:

$$[\![\pi]\!] = \boldsymbol{\theta} \in \Theta \mapsto [\![\nu]\!](\boldsymbol{\theta}) L(y_1, \ldots, y_c; \boldsymbol{\theta}), \tag{2.25}$$

where $y_1, \ldots, y_c$ denote the observations (and without loss of generality, can be considered to be scalar-valued), and $L$ denotes the likelihood

function (which may also have irrelevant constant factors removed), so the sequence of distributions used in the distribution tempering scheme is $\{\pi_t\}_t$ where

$$[\![\pi_t]\!] = \boldsymbol{\theta} \in \Theta \mapsto [\![\nu]\!](\boldsymbol{\theta}) L(y_1, \ldots, y_c; \boldsymbol{\theta})^{\phi_t}, \tag{2.26}$$

$\nu$ is the prior distribution. Note that at the final iteration $T$, $\phi_T = 1$ and so $[\pi_T] = \pi$, the target distribution as in (2.25).

With this method, it is necessary to be able to evaluate the likelihoods $L(y_1, \ldots, y_c; \boldsymbol{\theta})$ pointwise, but this is certainly not a problem in the applications here. However, it is worth noting that it will sometimes be necessary to use an ODE solver to compute these likelihoods, which will naturally slow down the computations to a significant, but non-crippling extent. Crucially, it is also the case that the constants of proportionality (i.e. those not depending on $\boldsymbol{\theta}$) of each intermediate distribution need not be known *a priori*, but can (if of interest) be estimated by the SMC sampler.

It is convenient to work with logarithms of the (unnormalised) density functions of the distributions:

$$\log [\![\pi_t]\!] = \boldsymbol{\theta} \in \Theta \mapsto \text{constant} + \log [\![\nu]\!](\boldsymbol{\theta}) + \phi_t \underbrace{\log L(y_1, \ldots, y_c; \boldsymbol{\theta})}_{\ell(\boldsymbol{\theta}|\boldsymbol{y})},$$

and in this situation, the constant additive term, not depending on $\boldsymbol{\theta}$, incorporates the normalising constant and need not be known.

With the same assumptions as presented in (2.20),

$$\ell\Big(\boldsymbol{\theta}|\boldsymbol{y}\Big) = -\frac{1}{2}\left(\log\det \mathbf{V}(\boldsymbol{\theta}) + \underbrace{\boldsymbol{r}(\boldsymbol{y}, \boldsymbol{\theta})^{\mathrm{T}} \mathbf{V}(\boldsymbol{\theta})^{-1} \boldsymbol{r}(\boldsymbol{y}, \boldsymbol{\theta})}_{\text{weighted residual sum of squares (WRSS)}}\right), \tag{2.27}$$

where $\boldsymbol{r}(\boldsymbol{y}, \boldsymbol{\theta}) = \Big(\boldsymbol{y} - \boldsymbol{h}(\boldsymbol{\theta})\Big)$ is the modified residual (modified since, if censoring occurs, it may be cut-off at the limit of quantification). See

again subsection 2.7.2 where other error distributions were discussed.

In practice, the speed of the SMC sampler is important, so where simplifications can be made to the above equation, e.g. where $\mathbf{V}$ is known to be a diagonal or block diagonal matrix, these optimisations are built into the code and provide significant speed ups in those situations. This is worth the extra maintenance effort of having to maintain multiple different versions of the likelihood calculation code (i.e. in the diagonal, block diagonal and dense matrix cases).

During the SMC runs, it often happens that all but a few particles gain negligible weights; in other words, the particle system is degenerate. A convenient indicator of the degree of degeneracy is the "survival diagnostic", or the number of effective particles $N_{\text{eff}}$ (often referred to as the effective sample size, but this terminology is avoided here as it conflicts with the separate notion of effective sample size used in Equation 2.30 later, and is also somewhat misleading; again discussed later) [81]. The survival diagnostic is calculated as [82]:

$$N_{\text{eff}} = \frac{1}{\sum_{i=1}^{N} (w^{(i)})^2},$$

where $w^{(i)}$ is the normalised weight of the $i^{\text{th}}$ particle. After every iteration of the SMC sampler, the survival diagnostic should be tested to see if it is below a given threshold, and if so, to resample the particles: the new particle system (usually with the same number of particles) contains replicates of the previous particles in proportion to their weights, with the weights all set to be equal in the new system. This effectively means that particles with high weights will have lots of replicates in the new system (with the new replicates able to be moved independently in the next iteration), whereas particles with low weights will be omitted completely. There are various resampling schemes available that satisfy the above descriptive constraints, including stratified, multinomial and residual resampling. These schemes are not discussed here, except to

say that residual resampling will be used throughout, as it exhibits lower conditional variance for all configurations of the weights [83]. It is possible for the SMC algorithm to recover from situations where the survival diagnostic falls to a low level, although this may take several iterations. Therefore, it is helpful to warn if the survival diagnostic becomes quite low before resampling, especially if this occurs with $t$ close to $T$. Note that it is somewhat misleading to refer to the survival diagnostic as an effective sample size, because immediately upon resampling, the weights are reset to be equal and so the survival diagnostic is restored to the number of particles $N$, even though many of the particles will not be distinct.

As resampling introduces extra Monte Carlo error, it is not done unless necessary (i.e. the survival diagnostic is below a given threshold), and all calculations should be done before resampling where possible, e.g. adapting the proposal variance, or analysing the final particle distribution from the last iteration of the sampler. However, this makes those calculations slightly harder as they must work with a set of weighted particles rather than particles with uniform weights, as they would be following a resampling step.

The matters of selecting the SMC parameters, and evaluating convergence etc, will be discussed in the next subsection, subsubsection 2.7.6.2.

There are many different (but similar) versions of SMC algorithms. Some differ in the order of the component steps per iteration (e.g. reweight, select/resample, mutate, advance iteration number). The following algorithm describes the one used in this work and is very much like the standard versions [77], but with the simplifying assumptions that $\pi_t \approx \pi_{t-1}$, and with the otherwise arbitrary backwards Markov kernel chosen to be the time-reversed proposal kernel. This has allowed separation of the reweighting and mutation steps, whereas without the above choices, the incremental weights will depend on both the previous and new particle positions.

Under the assumptions that the distributions $\pi_t$ are invariant under their respective mutation kernels $K_t$ — written in symbols as follows:

$$\pi_t = \pi_t K_t \text{ and } \pi_{t-1} \approx \pi_t, \tag{2.28}$$

— the listing in Algorithm 1 summarises the version of the SMC algorithm used in this work.

A framework in `C++` called SMCTC [84] allows implementing such algorithms relatively straightforwardly, and was used here. (Although other frameworks were available, none were found to be as powerful or easy to use, and a comparison of such frameworks is not a focus of this work. Other frameworks became available during this work [85, 86, 87] but arrived after the implementation in SMCTC was already underway and so were not evaluated.)

As alluded to earlier, it is also worth mentioning that, in contrast to standard MCMC methods, SMC samplers are relatively easy to parallelise. The particles can be easily divided into chunks and different processors/machines can operate on different chunks simultaneously. Most operations are on individual particles and are independent of other particles, therefore these steps can readily be done in parallel. Of course, information from different instances needs to be merged when analysing variance, determining total weight, etc, but these steps happen a lot less frequently than the particle mutation steps, and do not require transmission of the entire particle collection, just aggregate information computed in parallel. Work on parallel SMC samplers has begun (e.g. [85]) but have considerably higher complexity than non-parallel variants. As the SMC runs needed in this work were able to be completed in practically acceptable time periods without requiring to take advantage of parallel environments, parallelisation was not employed. However, it may be necessary to employ parallel methods to complete the analysis of the model presented in section 4.2.

**Algorithm 1** SMC algorithm for approximating $\pi_T$ with the assumptions (2.28)

*Initialization:*

Choose an appropriate number of particles $N$ and either the sequence of intermediate distributions $\pi_0, \ldots, \pi_T$, or a procedure for determining the same.

Step $t = 0$: Particles are initialised by sampling from the initial distribution $\pi_0$. $\pi_0$ should be chosen so that this can either be done directly or with standard importance sampling methods. Write $\eta_0$ for the corresponding importance distribution (often $\eta_0 = \pi_0$).

For each particle $i$, sample $\boldsymbol{\theta}_t^{(i)} \sim \pi_0$ and set $w_0^{(i)} = \frac{[\![\pi_0]\!](\boldsymbol{\theta}_t^{(i)})}{[\![\eta_0]\!](\boldsymbol{\theta}_t^{(i)})}$ to be the importance weight.

Normalise the particle weights.

Choose an appropriate initial proposal scale $\Sigma_1$.

*Iteration:*

**for** $t = 1, \ldots, T$ **do**

*Reweighting:*

The weight of each particle $i$ is incremented from its weight in the previous iteration by multiplying by its unnormalised incremental weight, i.e:

$$w_t^{(i)} \leftarrow w_{t-1}^{(i)} \omega_t^{(i)}$$

where the unnormalised incremental weight is

$$\omega_t^{(i)} = [\![\pi_t]\!](\boldsymbol{\theta}_{t-1}^{(i)}) / [\![\pi_{t-1}]\!](\boldsymbol{\theta}_{t-1}^{(i)}).$$

Renormalise the weights if necessary.

Detect if there were any issues in the distance between $\pi_{t-1}$ and $\pi_t$ (e.g. the ESS has fallen too far or the $d_{\chi^2}$ is too large), and warn appropriately.

Optionally, adapt proposal scales $\tilde{\boldsymbol{\Sigma}}_t$ and/or distribution sequence.

*Resampling:*

**if** the survival diagnostic has fallen below the resampling threshold, **then** resample.

**if** the survival diagnostic has fallen below the warning threshold, **then** warn appropriately.

*Mutation:*

For each particle $i$, sample $\boldsymbol{\theta}_t^{(i)} \sim K_{t,\tilde{\boldsymbol{\Sigma}}_t}(\boldsymbol{\theta}_{t-1}^{(i)}, \cdot)$.

Optionally, save the state of the particle system (and/or other information obtained during the iteration) for later inspection.

**end for**

*Output:*

When the above has finished, the collection of weighted particles $(\boldsymbol{\theta}_T^{(i)}, w_T^{(i)})$, form an approximation of $\pi_T$.

#### 2.7.6.2. Choice of SMC proposal

It should be clear that with SMC samplers, there is always a trade off of number of particles, number of intermediate distributions and speed (and memory requirements) of computation. More intermediate distributions means closer successive distributions, while more particles means a potentially finer approximation of each distribution. The larger the number of parameters to be sampled, the larger the number of particles and number of iterations that should be selected, but there is no general rule of thumb for selecting these "meta" parameters.

It is convenient to select $K$ by means of a proposal kernel together with an accept-reject step. Writing $Q$ for the proposal and $\alpha$ for the acceptance probability:

$$K_{t,\tilde{\boldsymbol{\Sigma}}_t}(\boldsymbol{x},\cdot) = A \mapsto \int_A Q_{t,\tilde{\boldsymbol{\Sigma}}_t}(\boldsymbol{x},\boldsymbol{y})\alpha_t(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} + \left(1 - c_{\boldsymbol{x},t,\tilde{\boldsymbol{\Sigma}}_t}\right)\delta_x(A)$$

where

$$c_{\boldsymbol{x},t,\tilde{\boldsymbol{\Sigma}}_t} = \int_{\mathbb{R}^d} Q_{t,\tilde{\boldsymbol{\Sigma}}_t}(\boldsymbol{x},\cdot)\alpha_t(\boldsymbol{x},\cdot),$$

and

$$\alpha_t(\boldsymbol{x},\boldsymbol{y}) = \min\left\{\frac{[\![\pi_t]\!](\boldsymbol{y})Q_{t,\tilde{\boldsymbol{\Sigma}}_t}(\boldsymbol{y},\boldsymbol{x})}{[\![\pi_t]\!](\boldsymbol{x})Q_{t,\tilde{\boldsymbol{\Sigma}}_t}(\boldsymbol{x},\boldsymbol{y})}, 1\right\}$$

This is equivalent to drawing a proposal from $Q_{t,\tilde{\boldsymbol{\Sigma}}_t}(\boldsymbol{x},\cdot)$ and accepting this proposed state (i.e. storing $\boldsymbol{\theta}_t^{(i)} \leftarrow \boldsymbol{x}$) with probability $\alpha_t(\boldsymbol{x},\boldsymbol{y})$, otherwise leaving the state unchanged. Note that this satisfies the assumption that $K_t$ is $\pi_t$-invariant.

For the purposes here, $Q_{t,\tilde{\boldsymbol{\Sigma}}_t}(\boldsymbol{x},\cdot)$ is chosen to be a Metropolis random walk starting at $\boldsymbol{x}$, with covariance matrix $\tilde{\boldsymbol{\Sigma}}_t$. This is a special case of a Metropolis-Hastings random walk with a symmetrical proposal. The symmetry allows for a simpler description of the mutation step as follows: For each particle $i$, draw a sample from $Q_{t,\tilde{\boldsymbol{\Sigma}}_t}(0,\cdot)$ and add it to the current state $\boldsymbol{\theta}_t^{(i)}$ to give a proposed state $\boldsymbol{x}$. The acceptance probability then simplifies to $[\![\pi_t]\!](\boldsymbol{x})/[\![\pi_t]\!](\boldsymbol{\theta}_t^{(i)})$.

If the proposal variance is too small then, depending on the distribution in question and how different it was to the previous distribution, the acceptance rate will be high but the particles will move around the space slowly, which might correspond to the particle cloud from the next iteration not adequately representing that iteration's target distribution. Conversely, if the proposal variance is too large then, again depending on the distributions in question, it is likely that the proposals will be in regions of lower density and so the acceptance rate will be lower. Again this might mean that the particle cloud is not a good representation of the intermediate distribution. This therefore suggests that it is important to select the proposal variance and monitor the acceptance rate carefully, but fortunately SMC samplers are more robust than similar MCMC algorithms, in that the whole probability distribution is estimated at each iteration, and the next distribution in the sequence is not too dissimilar from the previous. Indeed, some thought should be given to the selection of the proposal variance, but generally increasing the number of intermediate distributions will allay such issues, provided the acceptance rate isn't either extremely high or extremely low. When using distribution tempering, and without adapting the proposals, it is usually observed that the accept rate will decrease with every iteration, so adaptation of the proposal variance is important in SMC settings.

Roberts and Rosenthal [88] found an optimal proposal in a certain high-dimensional MCMC setting, which was:

$$K_t(\boldsymbol{x}, \boldsymbol{x} + \cdot) = (1 - \beta)N(0, (2.38)^2 \boldsymbol{\Sigma}_t / d) + \beta N(0, (0.1)^2 I_d / d),$$

where $\boldsymbol{\Sigma}_t$ denotes the empirical covariance matrix of the $t^{\text{th}}$ iteration, and they found the optimal acceptance rate in such settings to be 0.234. As this was an MCMC context rather than an SMC context, it required $t > d$ for it even to make sense, but the condition imposed was $t > 2d$ presumably for stability. $\beta$ was taken to be 0.05 and this component (that

does not depend on the covariance matrix) ensured that the algorithm did not get stuck if the covariance matrix $\mathbf{\Sigma}$ became singular at any point.

Although the context here is not the same, it still seems reasonable to select a similar proposal and aim for a similar acceptance rate. Used here is the following adaptation, which does not necessarily satisfy the diminishing adaptation condition or other assumptions that were necessary in the MCMC context:

$$(1-\beta)N(0, f_t(2.38)^2\tilde{\mathbf{\Sigma}}_t/d) + \beta N(0, (0.1)^2 I_d/d),$$

although if the covariance matrix was found to be singular, $\beta$ was temporarily increased to 1. The factor $f_n$ is another type of adaptation introduced, with $f_1 = 1$ and $f_{t+1} = \max(\min(1 + (a_t - 0.234), 1.1), 0.9)f_t$ where $a_t$ denotes the accept rate of iteration $t$. This is not dissimilar to the scheme proposed by Roberts and Rosenthal [88], but here the diminishing adaptation condition is not required so the adjustment factor is not itself scaled according to the iteration. Other approaches could of course be adopted that use more information about the history of the acceptance rate over previous iterations, which would better facilitate reaching or remaining at the target acceptance rate, but is felt to be unnecessary here, especially as the target acceptance rate is not known to be optimal in this context. In any event, only a reasonable performance is needed, not an "optimal" performance.

The proposal scale $\tilde{\mathbf{\Sigma}}_t$ can be updated (based on the covariance $\mathbf{\Sigma}_t$ of whole particle system) every $K$ SMC iterations (say) for some $K$. As the covariance of the particle system is relatively inexpensive to compute, it is feasible to take $K = 1$, so that is done here.

#### 2.7.6.3. Choice of SMC tempering schedule

The tempering schedule is not adapted here, but a sensible scheme has been proposed by Zhou et al. [78], who suggests aiming for a fixed reduction in the conditional ESS, a quantity also defined by him. Instead, a fixed geometric-based tempering schedule is applied instead: $\phi_k = \left(\frac{k}{T}\right)^\lambda$ with $\lambda \geqslant 1$ chosen so that $\pi_1$ is reasonably close to $\pi_0$, as the jump between these two distributions in particular is substantially different to that of any other adjacent distributions. This could be described as a deterministic adaptation, but the term "adaptive" here is reserved for dynamic adaptation cases.

While tuning the parameters, it may be helpful to run the SMC solver only up to some iteration $t < T$ (e.g. just for the first one or two iterations or so), to check that things are working correctly and the distance between the first two distributions is not too large.

#### 2.7.6.4. Assessing convergence of SMC runs

Diagnostics for sample impoverishment are borrowed from Carpenter et al. [89] (which was a particle filtering context). Somewhat confusingly, these are also referred to as effective sample sizes. Although this is not a particle filtering context, the same concept applies without modification. If one is interested in the property

$$\int g(\boldsymbol{\theta}_t)\,\mathrm{d}\mathcal{L}(x_k|D_k), \tag{2.29}$$

(for instance with $g$ taken to give the marginal means of a parameter) then, with $M$ replicates of weighted particles $(x_k^{(m,j)}, w_k^{(m,j)})$ (for $m = 1, \ldots, M$), the effective sample size of the property is

$$M\overline{v}_k / \sum_{m=1}^{M} \left(z_k^{(m)} - \overline{z}_k\right)^2 \tag{2.30}$$

where

$$\begin{aligned} z_k^{(m)} &= \sum_{j=1}^{N} w_k^{(m,j)} g(x_k^{(m,j)}), \\ v_k^{(m)} &= \sum_{j=1}^{N} w_k^{(m,j)} g^2(x_k^{(m,j)}) - z_k^2 \end{aligned} \tag{2.31}$$

and $\overline{z}_k, \overline{v}_k$ are the average values of the $M$ replicates of $z_k, v_k$ respectively.

#### 2.7.6.5. Chi-squared distance between distributions

Knowledge of the "distance" between distributions is useful to check that distributions of the SMC sampler corresponding to successive iterations are sufficiently close for the SMC to perform effectively. It can also be used to give some insight into the identifiability of the problem, and this aspect is discussed in the next subsection.

The chi-squared distance between two distributions $\phi$ and $\mu$ is defined as [90]:

$$\begin{aligned} d_{\chi^2}(\phi, \mu) &= \mathrm{Var}_\mu\left(\left[\phi(X)\right] / \left[\mu(X)\right]\right) \\ &= \mathbb{E}_\mu(([\phi](X)/[\mu](X))^2) - (\mathbb{E}_\mu([\phi](X)/[\mu](X)))^2 \\ &= \int([\phi]/[\mu])^2 [\mu] - (\int [\phi]/[\mu] [\mu])^2 = (\int [\phi]^2 / [\mu]) - 1 \end{aligned} \tag{2.32}$$

Of course, sufficient knowledge of either or both of $\phi$ and $\mu$ is required to calculate this quantity directly, and is often not available or not feasible to obtain since the curse of dimensionality arises again. Therefore, Monte Carlo approximations of $\chi^2$ distances are calculated in this work, which involve little extra effort when a Monte Carlo approximation of the distribution $\phi$ is already available. The integral is replaced with an importance sampling integral, and the normalising constants of $\phi$ and $\mu$ are also estimated with importance sampling integrals. Obviously multiple replicates should be obtained and averaged but a single replicate can still be informative. However, in this setup, if $\phi$ is not a good importance base distribution for $\mu$, then the corresponding $d_{\chi^2}(\phi, \mu)$ estimate will be unreliable (and possibly poorly conditioned).

The chi-squared distances provide quantities that can be bench-

marked but not directly interpreted.

#### 2.7.6.6. Further evaluating SMC output

During the SMC runs, although the resource requirements of outputting the entire particle distributions for each iteration would be excessive, it is possible to output the particles periodically, not just at the final iteration. It is helpful to output the full particle system at iteration 0, to check that the prior distribution was correct and to help identify whether any oddities seen in later iterations may have been caused by the specific sample used in this initial iteration. Iteration one is also a very important iteration as it is the first one that introduces the likelihood function. If the tempering exponent (or number of iterations) is not well tuned then the distribution after this iteration may either look too similar to the initial distribution, suggesting that there may be many wasted computations in this iteration and following iterations, or it may be drastically different from the initial distribution suggesting that the assumption of being close has been violated and so the distribution from the first iteration may not be accurate, not least because only a few particles from the initial iteration will have survived, resulting in particle degeneracy. Obviously the final particle distribution is the target distribution in this setting, and so needs to be output in full so it can be thoroughly analysed. It may also be helpful to output one or more iterations mid-way through the SMC run, e.g. $\lfloor T/2 \rfloor$.

Further, it is possible to output details about every iteration, such as: the average acceptance rate, the survival diagnostic, the proposal variance, summary statistics for each of the posterior parameters, the chi-squared distance between the current and next iteration, whether resampling was performed, and details of a particle with maximal density. These details will make it much easier to notice if anything looks unusual, and make it easier to identify where and what may have gone wrong.

If the SMC sampler is executed with a large number of particles,

then care needs to be taken when analysing the resulting particles: the memory requirements of loading all the particles into memory on a standard PC can exceed the capability of the machine, and Windows is especially bad at remaining responsive during (or recovering from) low available memory situations.

Obviously, SMC is used mostly when the posterior cannot be evaluated analytically, and so the SMC estimate cannot be compared with the actual posterior. However, it is helpful to run the SMC sampler for simpler situations where the posterior can in fact be evaluated analytically, to check that the results are correct in those cases, and to get a feel for the behaviour and variance of the sampler and things to look out for. The same can be done in cases where the posterior cannot be evaluated analytically but where features of the posterior are known, and it can then be checked whether the SMC sampler has correctly captured those features. See the next subsection for details where this was done.

Finally, if the output from the SMC sampler looks like it might potentially correspond to the correct target distribution, it is wise to increase the number of SMC iterations and/or the number of particles by an order of magnitude, and to check that the resulting estimates are similar. This need only be done once or twice for every model, and not necessarily for every set of input data.

There are of course ways to test whether two independent samples may have come from the same probability distribution, and to test whether a sample may have come from a particular known probability distribution. In one-dimension, the Kolmogorov-Smirnov test for two independent samples and the Kolmogorov-Smirnov goodness-of-fit test [91] are such possibilities, and do not assume that the error follows a normal distribution. In higher dimensions, similar tests are somewhat harder [92] and there has not been time to apply these in this work.

Visualisation of multi-dimensional distributions can be difficult, but marginal distributions are trivial to extract from the weighted particle

collections output by the SMC sampler, so it is generally worth plotting marginal and pairwise marginal distributions for all the parameters. These marginal distributions can often indicate features and/or likely problems with the calculated posterior distribution. Scatter plots can also be used if desired, provided that either the particles are resampled first to have uniform weights or particles are shown with markers whose sizes are relative to their weights.

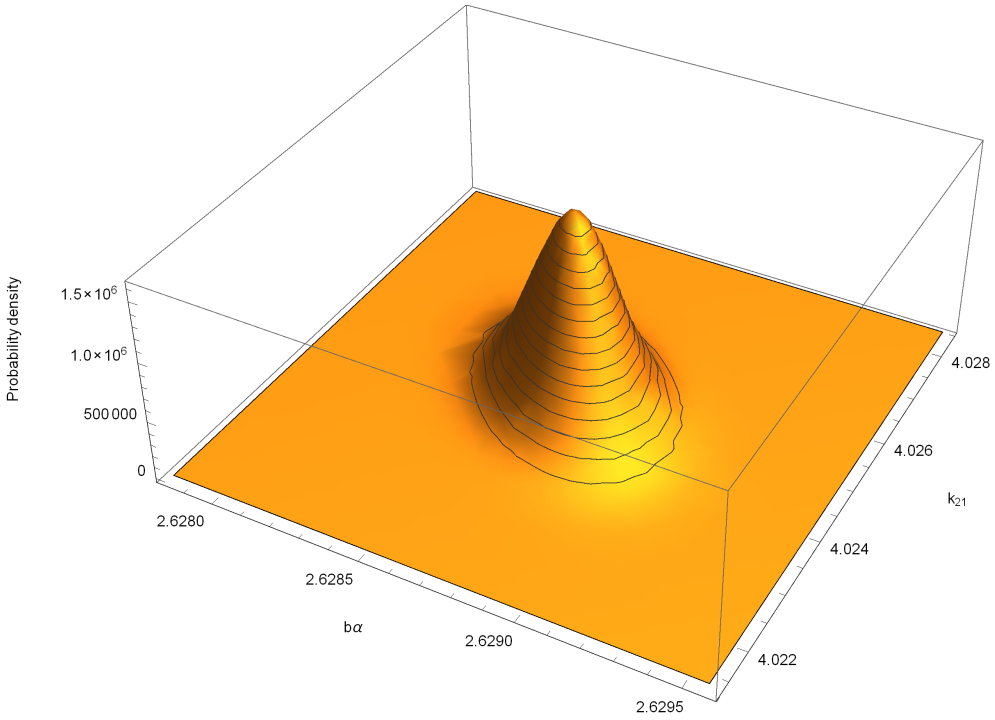
#### 2.7.6.7. SMC examples on simple low-dimensional models

As mentioned in the previous subsection, the SMC sampler should be tested in situations where the true target distribution is known exactly, so the sampler can be evaluated for performance and accuracy, before it is applied to distributions that are less straightforward to verify.

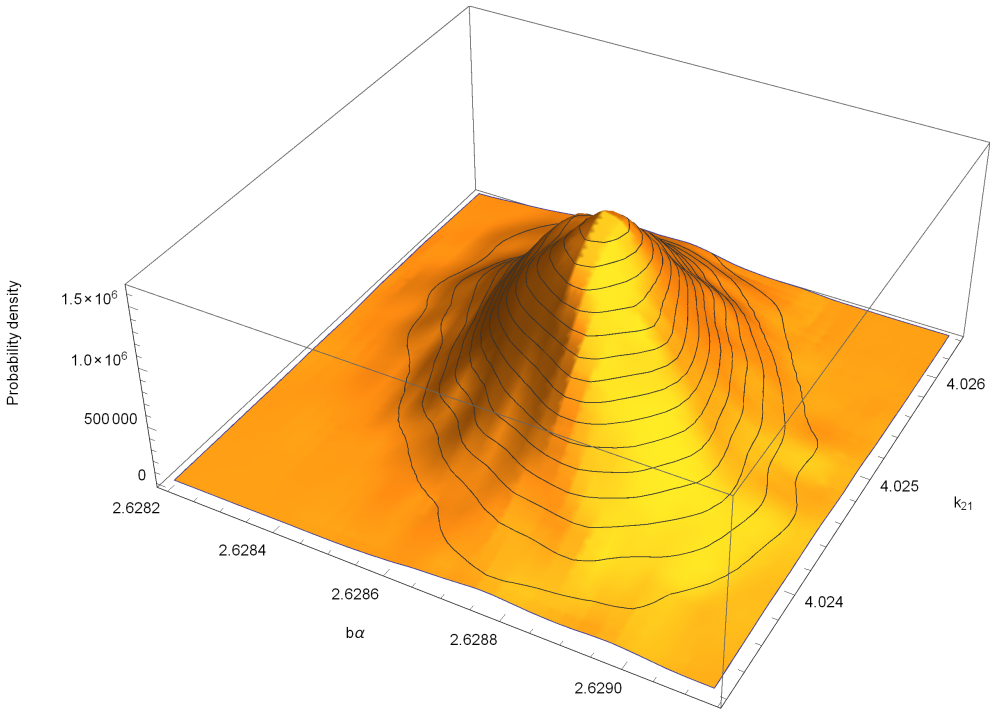
With this in mind, the SMC sampler described was applied to a structurally globally identifiable 2-compartment model with 2 parameters considered unknown. As this is a 2-dimensional problem, the results are easy to see visually in 3-dimensions. This was the model depicted in Figure 2.4 with $\alpha b = 2.6287$, $k_{2,1} = 4.02488$, $k_{\mathrm{e},1} = 0$ and both $u(t) = D\delta(t)$ (an impulsive input) with $D = 4362.76$ and $k_{\mathrm{e},2} = 1.53968$ were considered known. Initially, the sampler was applied to noise-free observations from this model, and because the observations were noise free, the above (normally excessive) accuracy stated in the parameters could be recovered by the sampler.

The correct posterior distribution, computed directly, is shown in Figure 2.6, and the smoothed output from the SMC sampler with 1,000 iterations, 10,000 particles and tempering exponent 2 is shown in Figure 2.7. The two plots have slightly different scales but appear to be in agreement. As mentioned in the previous subsection, a statistical test would ideally be performed here to check quantitatively whether the SMC output appears to be a sample from the reference probability distribution, the true posterior. The author is aware that such a test is not

straightforward in more than one dimension and so presently relies only a visual calculation.

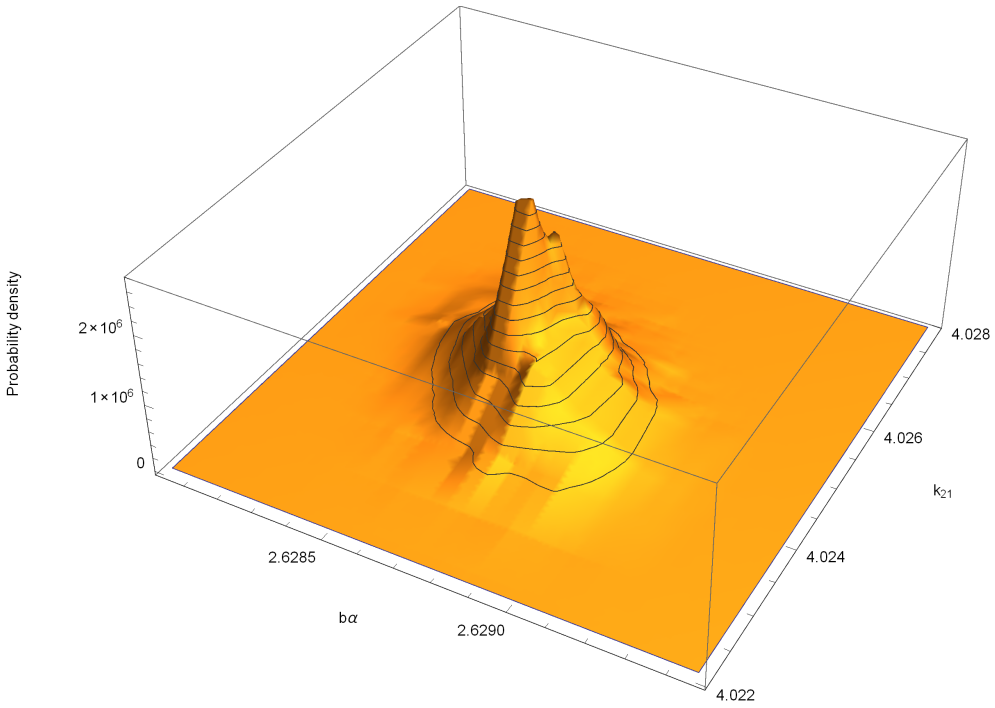



▲ Figure 2.6: 3D density plot of known correct posterior distribution from model used for testing



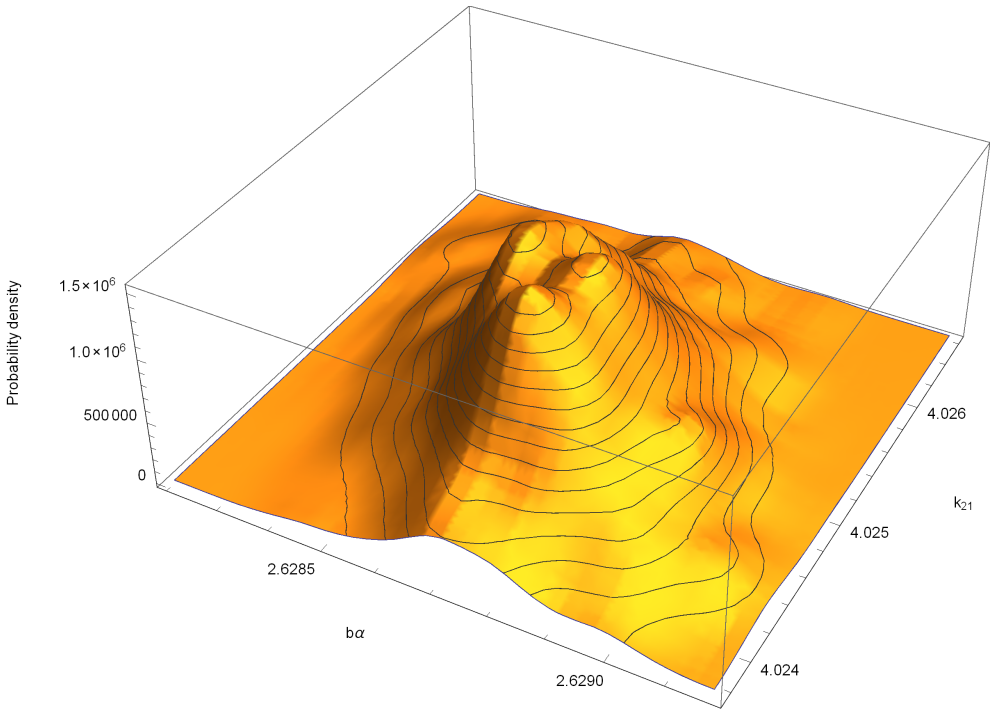

▲ Figure 2.7: Smoothed 3D density plot of output from SMC sampler applied to known-posterior testing model with 1,000 iterations, 10,000 particles, tempering exponent 2

With insufficient iterations with respect to the 2 compartment model with 2 unknown parameters, the particle system output by the SMC sampler misrepresented the true posterior distribution by showing the distribution to have a lower variance than it actually does. Of course, this may not be the only way that insufficient iterations generally affect the output. The histogram in Figure 2.8 shows the output of the SMC sampler when run with only 10 iterations, but with 10,000 particles. This can be compared and contrasted with the true distribution in Figure 2.6.




▲ Figure 2.8: Smoothed 3D density plot of output from SMC sampler applied to known-posterior testing model with 10 iterations (insufficient), 10,000 particles, tempering exponent 10

With insufficient particles with respect to the same situation, the particle system output by the SMC sampler again misrepresents the true posterior distribution by showing the distribution to have a number of nearby modes instead of one distinct mode. The histogram in Figure 2.9 shows the output of the SMC sampler when run with only 1,000 particles (but for 5,000 iterations).




▲ Figure 2.9: Smoothed 3D density plot of output from SMC sampler applied to known-posterior testing model with 5,000 iterations, 1,000 particles (insufficient), tempering exponent 10

Clearly, it is necessary to choose carefully both the number of iterations and the number of particles. A deficiency in the number of particles cannot always be easily corrected by a modest (even geometric) increase in the number of iterations, and vice versa. While a significant increase in the number of iterations could potentially compensate for a low number of particles, this would be inefficient and difficult to justify compared to increasing both the number of particles and the number of iterations to an appropriate (but more modest) level.

# Chapter 3

# New applications of statistical analysis techniques

## 3.1. SMC and identifiability

As mentioned in subsection 2.5.6, it is not always possible to determine the identifiability of a model algebraically. This may be because the algebra becomes too complicated, or because the model does not admit a rational polynomial expression. In such cases, although the identifiability question cannot be definitively resolved structurally, it can be insightful to consider the numerical identifiability of the model. Ideally, this numerical identifiability information can be extracted using the same tool that is used to perform parameter inference on a model, as the practitioner can then observe numerical identifiability issues from whatever source in the same way (e.g. issues caused due to the structural model, and also issues caused due to the limitations in the data collection frequency or noise).

As it has been suggested here that SMC is a good tool for model inference, so it is now suggested that SMC can be applied to the numerical identifiability problem. Whereas numerical optimisation techniques usually only find one mode from each starting point (and it is not gen-

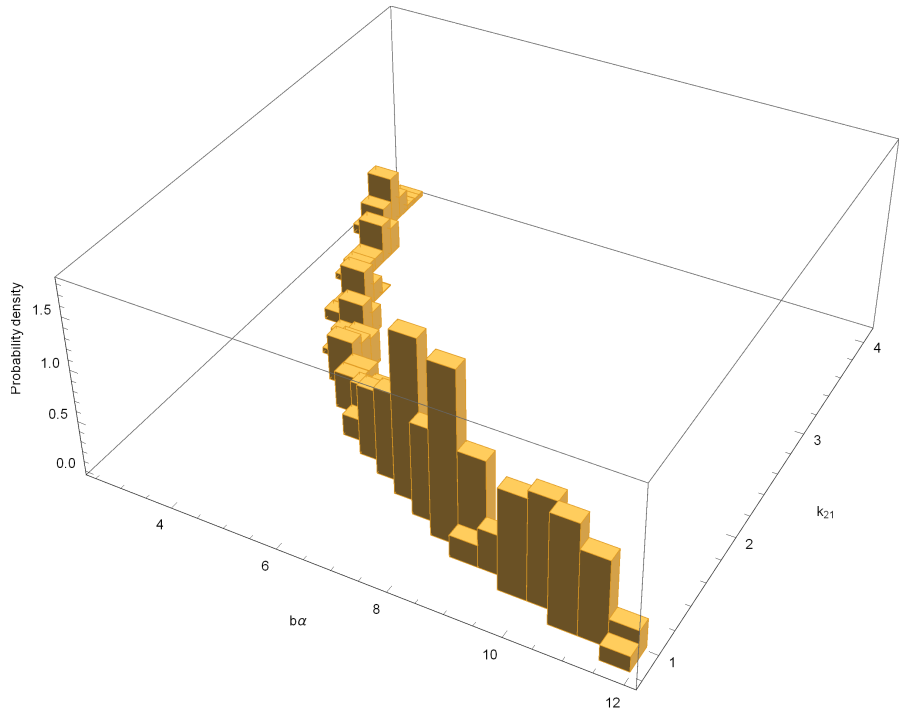
erally feasible to find all modes, or to confirm that all modes have been found), SMC estimates the entire posterior distribution and so if working correctly will find all non-negligible modes, a key property that makes it suitable for assessment of identifiability. However, it should be noted that while numerical optimisation always seeks to identify the mode to a given precision, SMC may only sample some particles in the vicinity of the mode, and not find the mode exactly. If desired, it is of course always possible to run a local optimisation method starting from the nearby particles to identify the modes more precisely.

### 3.1.1. SMC applied to structurally unidentifiable models

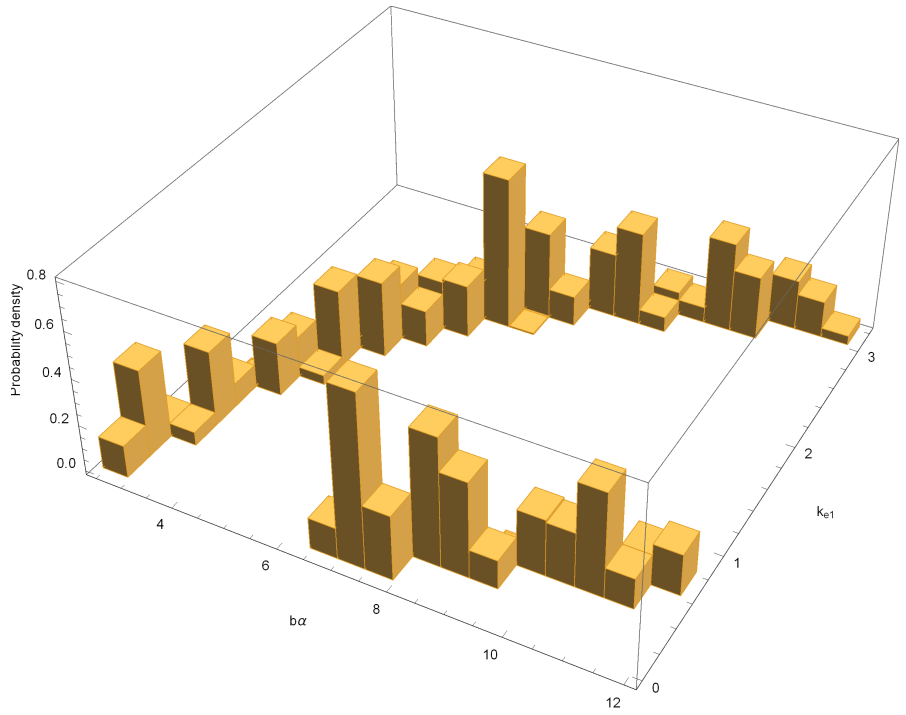
As illustrated in subsection 2.5.3, the 2 compartment model from Figure 2.4 is structurally unidentifiable when all the parameters are considered unknown (even considering $\alpha b$ as a combined parameter). For this linear model, the same structural identifiability findings carry over to the impulsive input case. The SMC sampler was therefore applied to this model in this setting to determine the behaviour of the sampler on a structurally unidentifiable model. The same parameter values as in subsubsection 2.7.6.7 were used. One resulting approximation of the posterior distribution is shown via the pairwise marginal density plots in Figures 3.1 through 3.6.

Each of the pairwise marginal densities clearly show a ridge in parameter space as in Figure 3.1, or multiple ridges as in Figures 3.2–3.6. The ridges arise from the fact that the parameters have one or more degrees of freedom while still describing the same input/output function. If it were not possible to conduct the structural identifiability analysis for this model, the shape of the ridges could be used to infer the functional relationship between the parameters while keeping the input/output function invariant. This information could then be used to assist the structural

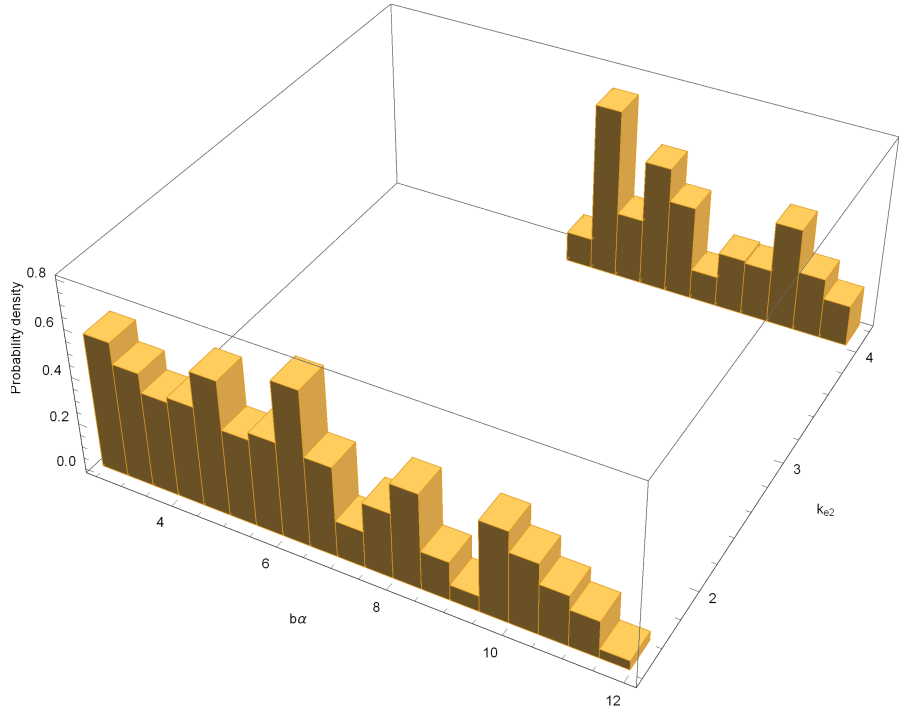
identifiability analysis.



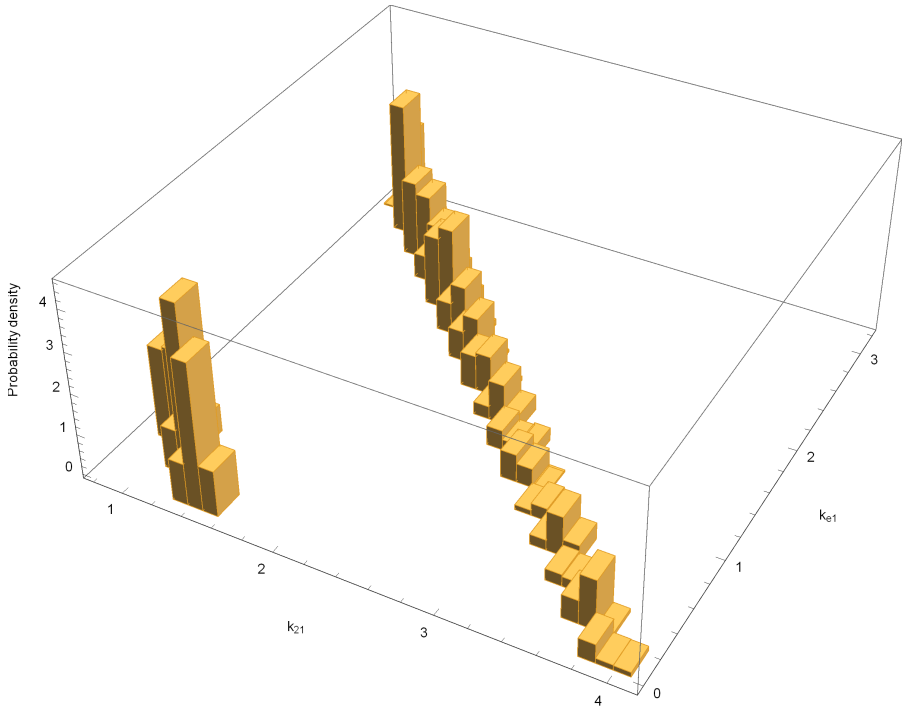

▲ Figure 3.1: SMC results on a structurally unidentifiable model: marginal histogram density plot of $b\alpha$ and $k_{21}$



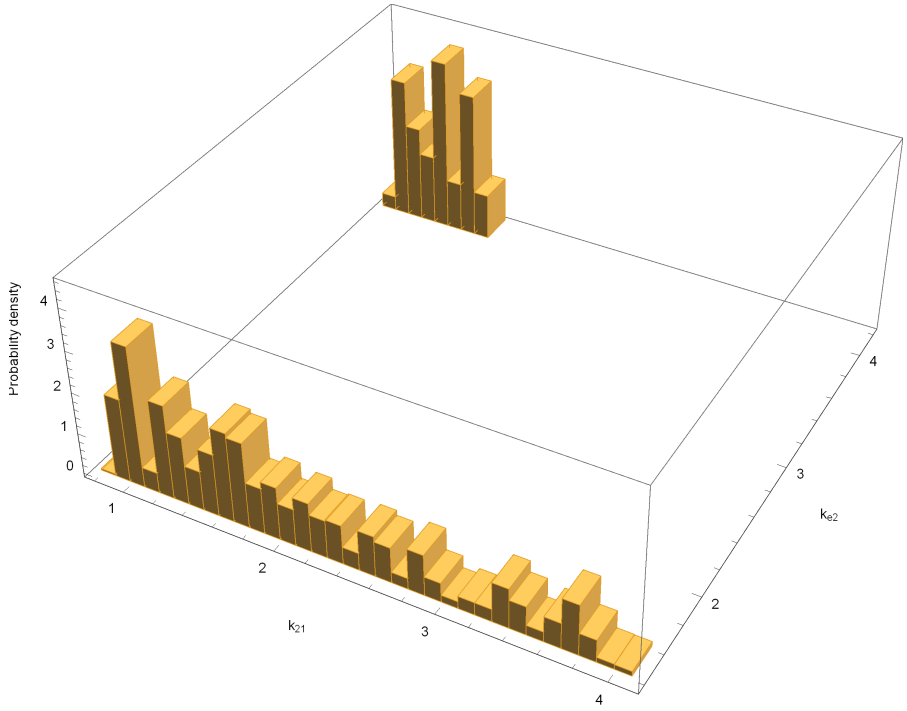

▲ Figure 3.2: SMC results on a structurally unidentifiable model: marginal histogram density plot of $b\alpha$, $k_{e1}$



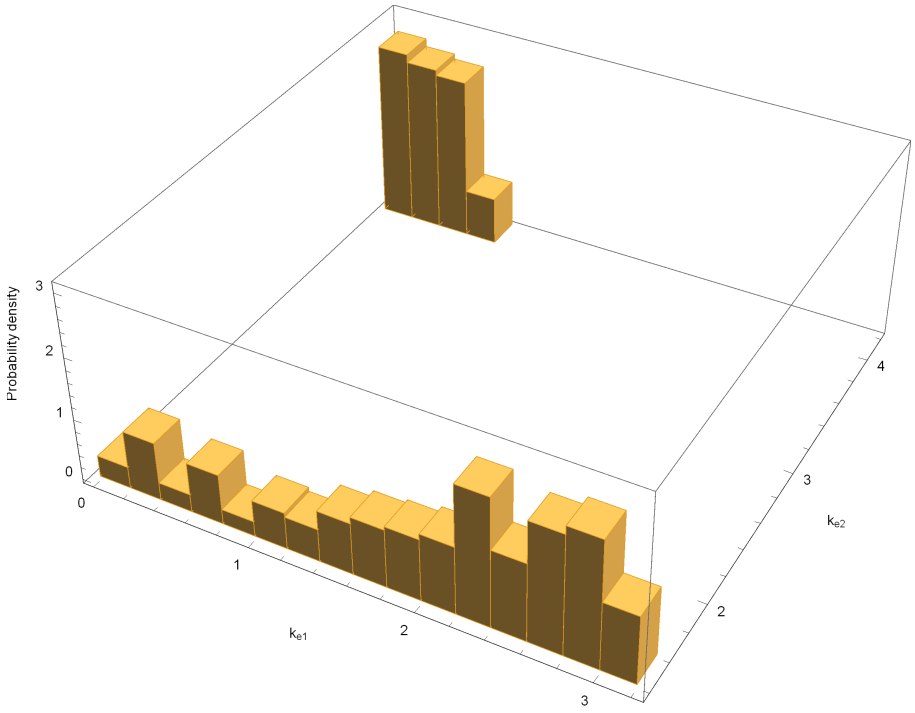

▲ Figure 3.3: SMC results on a structurally unidentifiable model: marginal histogram density plot of $b\alpha$, $k_{e2}$




▲ Figure 3.4: SMC results on a structurally unidentifiable model: marginal histogram density plot of $k_{21}$, $k_{e1}$




▲ Figure 3.5: SMC results on a structurally unidentifiable model: marginal histogram density plot of $k_{21}$, $k_{e2}$




▲ Figure 3.6: SMC results on a structurally unidentifiable model: marginal histogram density plot of $k_{e1}$, $k_{e2}$

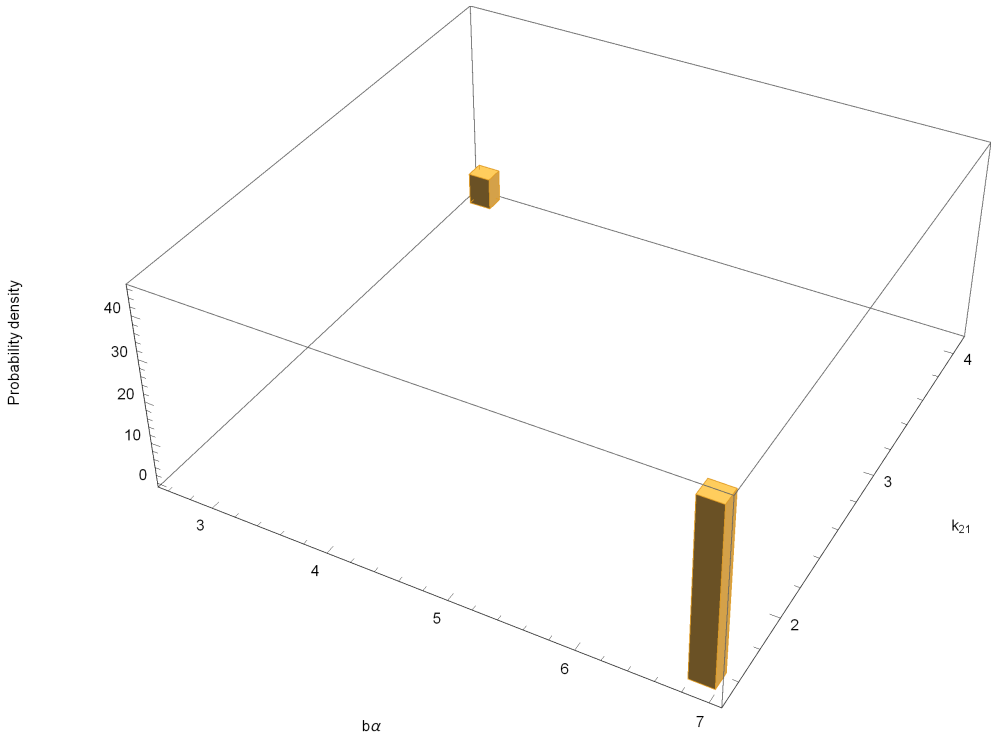
### 3.1.2. SMC applied to structurally locally identifiable models

As shown in subsection 2.5.3, the 2 compartment model presented earlier in Figure 2.4 is structurally locally identifiable (but not structurally globally identifiable) when $\alpha b$ is treated as a combined parameter and $k_{\mathrm{e},1} = 0$, while the remaining parameters are considered unknown.

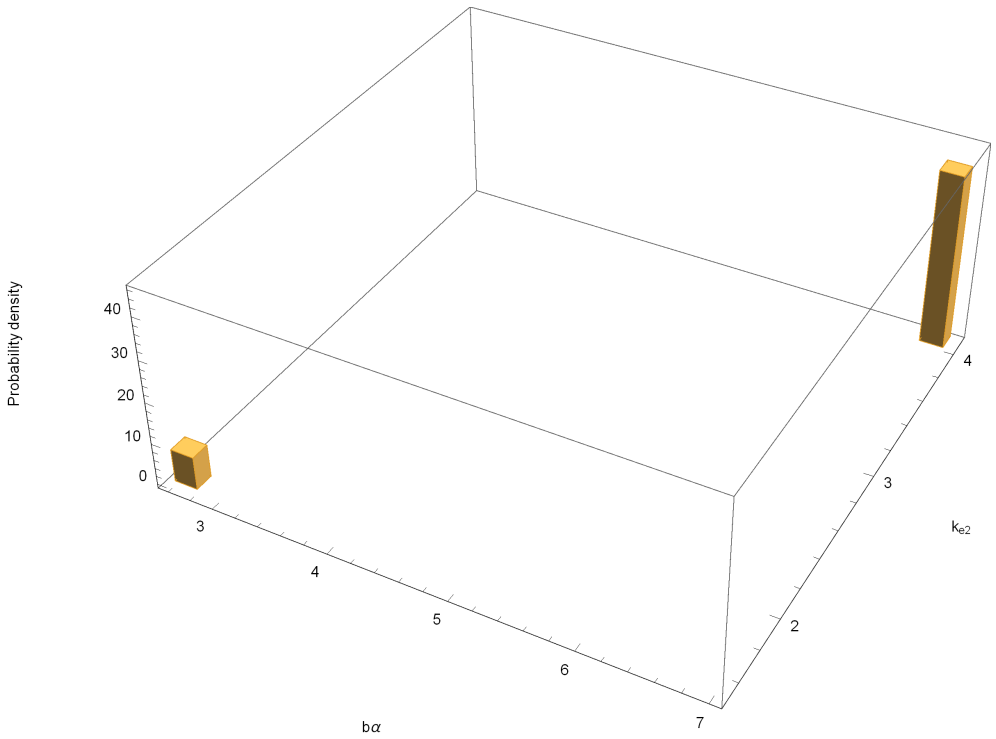
With the same simulated data as previously, there are two indistinguishable parameter combinations:

1. the original parameter set $\overline{\alpha b} = \alpha b = 2.6287$, $\overline{k}_{2,1} = k_{2,1} = 4.02488$ and $\overline{k}_{\mathrm{e},2} = k_{\mathrm{e},2} = 1.53968$; and
2. $\overline{\alpha b} = \frac{\alpha b k_{2,1}}{k_{\mathrm{e},2}} = 2.6287$, $\overline{k_{2,1}} = k_{\mathrm{e},2} = 4.02488$, and $\overline{k_{\mathrm{e},2}} = k_{2,1} = 1.53968$.

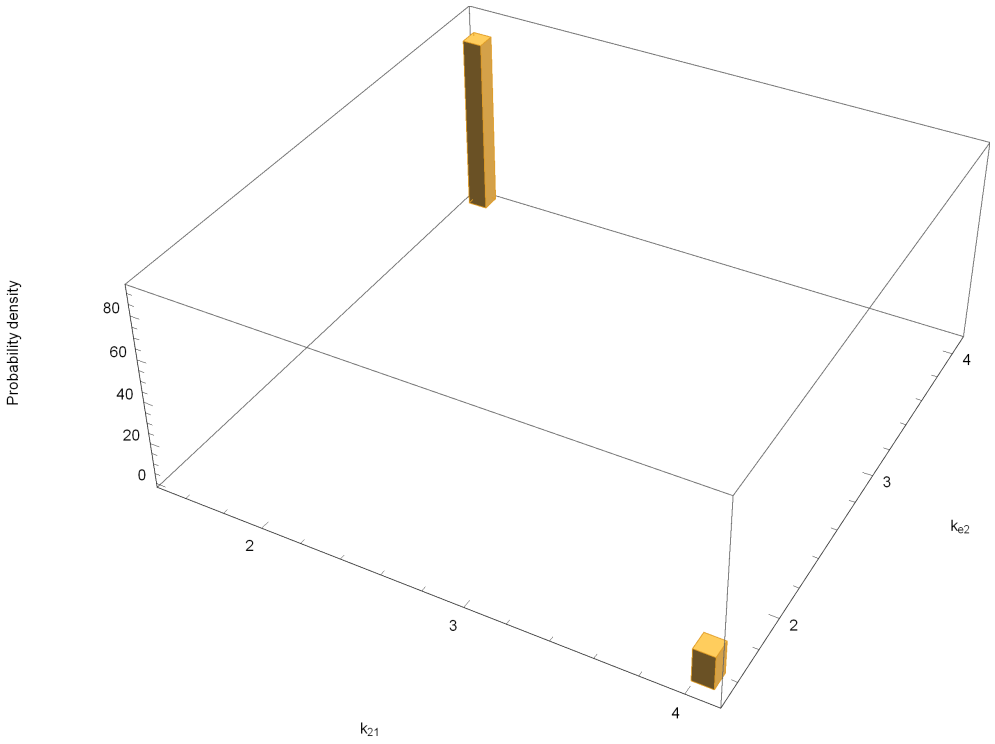
The SMC sampler was again applied to this model with $k_{\mathrm{e},1} = 0$ considered known and the remaining parameters considered unknown. A resultant approximated posterior distribution is shown via the pairwise marginal density plots in Figures 3.7 through 3.9. The plots clearly show that the marginal distributions have two separated and well-defined modes (though the large separation between them prevents display of more detail in the vicinity of these modes), and it can be seen that these modes occur exactly in the expected positions as revealed by the structural identifiability analysis. This demonstrates that one possible cause of having multiple but separated modes in the posterior distribution could be structural local identifiability of the model. If it were not already known, it would not be completely infeasible to identify the functional relation between the parameters from the positions of the modes and to use this information to confirm algebraically that the model is not structurally globally identifiable.




▲ Figure 3.7: SMC results on a structurally locally identifiable model: marginal histogram density plot of $b\alpha$, $k_{21}$




▲ Figure 3.8: SMC results on a structurally locally identifiable model: marginal histogram density plot of $b\alpha$, $k_{e2}$




▲ Figure 3.9: SMC results on a structurally locally identifiable model: marginal histogram density plot of $k_{21}$, $k_{e2}$

### 3.1.3. Using the posterior distribution to inform on identifiability

A much quicker way to inform on overall practical identifiability is to calculate the chi-squared distance between the target distribution and an exponentiated copy of the target distribution (like a "tempered" distribution but raised to a power larger than 1), essentially extracting the information directly from the computed posterior distribution. Raising the posterior distribution to a power larger than 1 amplifies its peaks and diminishes other regions, so if the posterior distribution has well-defined peaks then there will be a large difference between it and the exponentiated distribution. On the other hand, if the posterior distribution does not have well-defined peaks (e.g. because it is practically unidentifiable), then raising it to a larger power will not result in a significant difference, and so the chi-squared distance will be smaller.

Although this method requires significantly less computational resources, and does not even require estimating the exponentiated distribution, it does not inform on which parameters are problematic.

## 3.2. SMC for computing profile likelihoods/posteriors

It is possible to compute "profile posteriors" for each of the parameters, which would be useful for various purposes, including to examine those profiles to determine the practical identifiability of the corresponding parameters. Namely, if a profile is flat for large regions of the parameter space, then the parameter is practically unidentifiable. If the profile contains multiple peaks then there is a possibility of getting stuck in a local mode. Conversely, if the profile has a distinct maximal peak, and few smaller peaks, then that parameter is likely to be practically identifiable.

However, computing these "profile posteriors" would appear to involve solving a large quantity of optimisation problems, which would again be infeasible. A more efficient approach would be to borrow information from one optimisation problem to help with the next such problem, and this is indeed possible through SMC. Each iteration in such a setup corresponds to a particular value of the 'profiled' parameter, and the distribution at each iteration (or at least the particle with maximal posterior density) is of interest, as opposed to a standard SMC run where only the final iteration is of real interest. Successive distributions are similar because they only involve a small increment/decrement to the profiled parameter. This also requires standard SMC runs to reach the starting distribution for each parameter, before the profiling can proceed. However, calculating the profile posteriors through SMC somewhat prevents those profile posteriors being used to help verify the original approximation of the target distribution as produced by SMC.

A rough implementation of this was attempted by the author but was soon abandoned in favour of the preceding approach, as it was necessary to discard previous SMC runs when switching from exploration mode to profiling mode, and therefore was not considered efficient enough to

justify its use.

# Chapter 4

# Model development and analysis

## 4.1. Pharmacokinetic model development and analysis

The model from subsection 2.6.1 is revisited, and an extension is now presented to the analysis thereof.

### 4.1.1. Structural identifiability

The structural identifiability of this model was previously considered in Hall [1] and some of the following extensions to the results there were presented in Hall et al. [2].

Let $\boldsymbol{p}$ denote the vector of unknown parameters in the model. Take

$$\boldsymbol{p} = \begin{pmatrix} b & k_{21} & k_{31} & k_{42} & k_{43} & k_{\mathrm{e}2} & k_{\mathrm{e}3} & k_{\mathrm{e}4} & \alpha_1 & \alpha_2 \end{pmatrix}^{\mathrm{T}}, \tag{4.1}$$

where the feasible parameter space is $\Omega := (0, \infty)^n \ni \boldsymbol{p}$, with $n = 10$ denoting the number of unknown parameters.

The observation function $\boldsymbol{y}$ is now written $\boldsymbol{y}(\cdot, \boldsymbol{p})$ to emphasise its dependence on the unknown parameters.

The structural identifiability of this model was analysed using the Laplace transform approach [54] in Hall [1], one of the most commonly used methods for linear time-invariant systems, where it was found that $k_{e4}$ is SLI with either $\overline{k_{e4}} = k_{e4}$ or $\overline{k_{e4}} = k_{43} + k_{e3}$, and all other model parameters are SUI. Additional assumptions were then considered to see if they constrain the system model to be structurally identifiable. A slight modification of those assumptions is presented below:

(a) Other studies have reported apparent volumes of distribution for ARS and DHA following oral administration of ARS. In particular, Morris et al. [39] report the median volume of distribution for ARS at 6.8 l/kg and 1.55 l/kg for DHA in malaria patients (though these are noted to vary significantly relative to severity of infection). Such information can be used to treat the ratio of the observation gains as known; that is, $r := \alpha_2/\alpha_1$ is known ($\alpha_1 = \alpha$ and $\alpha_2 = r\alpha$, say). Using the above information from Morris et al. [39], this would give $r = 4.387$ (the observation gain for DHA is larger because it has the smaller volume of distribution);

(b) There is no known reason to suggest that the metabolism of the ARS occurs at significantly different rates before and after absorption, so it might be valid to consider the metabolism rate constants to be equal: $k_{31} = k_{42}$;

(c) There are no other major metabolites of DHA and negligible quantities of ARS are detected in urine. Hence, ARS is almost entirely converted to DHA, and so it may be reasonable to assume that the elimination rate parameter $k_{e2} = 0$;

(d) Absorption of the metabolite is rapid, thus its elimination may be negligible before it is absorbed, i.e. $k_{e3} = 0$.

Note that when constraints of this sort are imposed on parameters, the corresponding models are considered to be structurally distinct;

structural identifiability is concerned with the behaviour of *almost all* parameter values, and these assumptions may mean that previously null sets now have strictly positive measure.

Each combination of these four assumptions was assessed (or if considered previously, then re-assessed) using the same methods as previously, and the structural identifiability results are tabulated in Table 4.1. It can be seen that applying just one of the additional constraints does not improve the structural identifiability for the majority of the parameters. Applying any combination of two constraints except $\alpha_2/\alpha_1 = r$ and $k_{e2} = 0$ constrains all the parameters to be at least structurally locally identifiable. Applying any combination of three of the assumptions constrains the model to be structurally globally identifiable. The assumption (b) that $k_{31} = k_{42}$ seems to be the weakest in terms of improving structural identifiability.

## 4.1.2. Statistical analysis

### 4.1.2.1. Weight cap

Hall [1] did not use the described weighting matrix for the data, so the use of that matrix was evaluated here. However, as was anticipated, due to the wide range in concentrations reported for individual patients over the studied time interval, parameter fitting using the weighting matrix corresponding to the reported errors (see subsection 2.7.2) did not yield good fits. When using errors corresponding to predicted observations in contrast to weighting by actual observations, high concentrations were artificially predicted, corresponding to low weights. These points could thus be missed completely with little penalty on the likelihood. Prior to conducting the analysis, this outcome was anticipated and it was planned that the condition number of the weighting matrix might need to be controlled to resolve this. The singular values of the weighting matrix (to cater for the cases where the matrix was not diagonal due to the

| Assumptions | | | | Structural identifiability results | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\frac{\alpha_2}{\alpha_1} = r$ $r$ known | $k_{31} = k_{42}$ | $k_{e2} = 0$ | $k_{e3} = 0$ | $k_{21}$ | $k_{31}$ | $k_{42}$ | $k_{43}$ | $k_{e2}$ | $k_{e3}$ | $k_{e4}$ | $b\alpha_1$ | $b\alpha_2$ | $b\alpha$ |
| 0 | 0 | 0 | 0 | U | U | U | U | U | U | L | U | U | - |
| 0 | 0 | 0 | 1 | U | U | U | L | U | - | L | U | U | - |
| 0 | 0 | 1 | 0 | U | U | L | U | - | U | L | U | U | - |
| 0 | 0 | 1 | 1 | L | L | L | L | - | - | L | L | L | - |
| 0 | 1 | 0 | 0 | U | - | U | U | U | U | L | U | U | - |
| 0 | 1 | 0 | 1 | L | - | L | L | L | - | L | L | L | - |
| 0 | 1 | 1 | 0 | G | - | G | L | - | L | L | G | L | - |
| 0 | 1 | 1 | 1 | G | - | G | G | - | - | G | G | G | - |
| 1 | 0 | 0 | 0 | U | U | L | U | L | U | L | - | - | U |
| 1 | 0 | 0 | 1 | L | L | L | L | L | - | L | - | - | L |
| 1 | 0 | 1 | 0 | U | U | G | U | - | U | G | - | - | U |
| 1 | 0 | 1 | 1 | G | G | G | G | - | - | G | - | - | G |
| 1 | 1 | 0 | 0 | L | - | L | L | L | L | L | - | - | L |
| 1 | 1 | 0 | 1 | G | - | G | G | G | - | G | - | - | G |
| 1 | 1 | 1 | 0 | G | - | G | G | - | G | G | - | - | G |
| 1 | 1 | 1 | 1 | G | - | G | G | - | - | G | - | - | G |

The applicable parameters under any combination of the assumptions (1 if the assumption is applied and 0 if not) are either structurally unidentifiable (U), structurally locally identifiable (L) or structurally globally identifiable (G).

Table 4.1: Structural identifiability analysis results for the four compartment model under different combinations of constraints

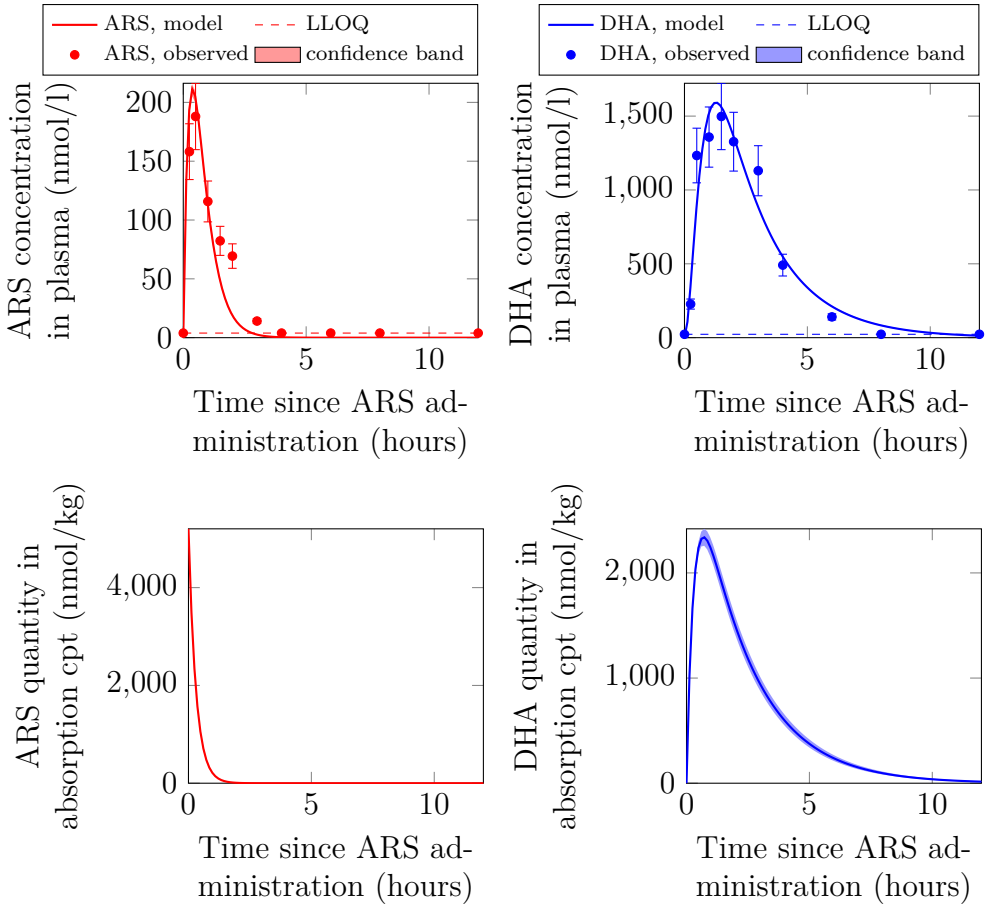
assumption of correlation between different measurement errors) were therefore capped so that no singular value exceeded 100 times the lowest singular value, resulting in the condition number of the weighting matrix becoming at most 100. Imposing this cap yielded much improved model fits, and so only these improved fits are presented here.

Even with this cap, the objective function had multiple local minima for many patients, and often had multiple local minima achieving similar objective function values but considerably different parameter estimates. In these cases, the global minimum was usually selected, except in a minority of cases where the fitted parameter values were extreme and a local minimum was identified that seemed more realistic. This highlights the fact that having a globally identifiable model structure is a necessary but not sufficient condition to ensure practical identifiability, especially in the presence of high model and observation errors.

#### 4.1.2.2. Maximum likelihood results

Observations and model fits for one of the patients ("Patient A") are shown in Figure 4.1 together with model predictions of the quantities in the absorption compartments. The corresponding parameter values and measures of their uncertainty are presented in Table 4.2 and Table 4.3. The confidence bands give an indication of the sensitivity of the fit.

It can be seen that the model fit for patient A appears to be satisfactory. Not all model fits were satisfactory for all patients. For brevity, results for the other patients are not presented here in the same way, but instead model fit results are summarised through the coefficient of determination and shown in Figure 4.2, and a summary of parameter estimates across all patients is provided in Table 4.4. The worst model fits correspond to patients whose observed ARS and DHA concentration–time profiles did not both reach peaks within 3 hours of dose administration, or those where at least one of the drugs exhibited multiple peaks (approximately half of the patients exhibited one of these issues, and are




▲ Figure 4.1: Example of model predicted ARS and DHA quantities/concentrations in each compartment for patient A using the 4-compartment pharmacokinetic model. Error bars are representative of assay error. Confidence bands are also shown but are very narrow.

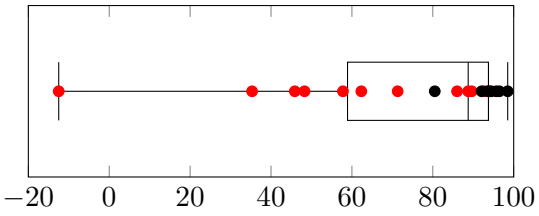
| Parameter | Fitted value | Standard error | Units |
|---|---|---|---|
| $b\alpha$ | 0.2330 | 0.0105 | kg/l |
| $r$ | 4.3870 | (fixed) | dimensionless |
| $k_{21}$ | 1.2518 | 0.0863 | $h^{-1}$ |
| $k_{42}$ | 2.0378 | 0.0199 | $h^{-1}$ |
| $k_{43}$ | 0.4604 | 0.0113 | $h^{-1}$ |
| $k_{e4}$ | 0.9975 | 0.0426 | $h^{-1}$ |
| $\rho$ | 0.0207 | (nuisance) | correlation |
| Objective function value | 3747.23 | | |
| Coefficient of determination | 96.46 | | % |

Table 4.2: Table of parameter estimates and their uncertainties for patient A using the 4-compartment pharmacokinetic model.

Parameter correlation matrix (darkness of black/red colour corresponds to strength of positive/negative value in each cell respectively):

| | $b\alpha$ | $k_{21}$ | $k_{42}$ | $k_{43}$ | $k_{e4}$ |
|---|---|---|---|---|---|
| $b\alpha$ | 1.000 | −0.977 | 0.199 | −0.767 | 0.994 |
| $k_{21}$ | −0.977 | 1.000 | −0.292 | 0.638 | −0.982 |
| $k_{42}$ | 0.199 | −0.292 | 1.000 | −0.021 | 0.187 |
| $k_{43}$ | −0.767 | 0.638 | −0.021 | 1.000 | −0.712 |
| $k_{e4}$ | 0.994 | −0.982 | 0.187 | −0.712 | 1.000 |

Table 4.3: Correlation matrix for the parameters in Table 4.2 for the 4-compartment pharmacokinetic model.




▲ Figure 4.2: Distribution of the coefficient of determination (%) over the Mahidol_PK dataset; marks in red correspond to patients with unexpected profiles. Recall that the objective was not to maximise the coefficient of determination, but this statistic allows easier comparison between subjects than the actual objective function values.

coloured in red in Figure 4.2). Note that the fit for one such patient has a negative coefficient of determination. This does not necessarily suggest that fitting the mean to the concentration–time profile of each drug would have performed better than fitting the model (although that is a natural interpretation), because the model still captures part of the absorption and elimination processes and therefore their shapes, though model predictions should not be relied upon in these circumstances. The coefficient of determination statistic was used to help quantify the goodness of the model fits, but it does not provide a good interpretation of the results: there is no direct correlation between this statistic and the likelihood function used. In particular, the objective criteria was not to maximise the coefficient of determination. These results suggest that it would be useful to compare the influence that the normal distribution had vs. Student's $t$-distribution or the Laplace distribution: see subsection 2.7.2 and section 5.1.

| Parameter | Mean (SD) | Units |
|---|---|---|
| $b\alpha$ | 0.472 (0.32) | kg/l |
| $r$ | 4.387 (fixed) | dimensionless |
| $k_{21}$ | 0.281 (0.30) | $h^{-1}$ |
| $k_{42}$ | 1.119 (0.74) | $h^{-1}$ |
| $k_{43}$ | 0.835 (0.59) | $h^{-1}$ |
| $k_{e4}$ | 1.612 (1.28) | $h^{-1}$ |
| CoD weighted | 74.25 (28.8) | % |
| ARS half-life | 0.93 (0.58) | h |
| DHA half-life | 0.72 (0.53) | h |

Table 4.4: Fitted parameter values, aggregated

The parameter estimation procedure was typically insensitive to the

correlation parameter $\rho$ (perhaps as a result of the weight cap) and this parameter was often fitted close to 0 even when not used as the initial value for the optimisation. Having preferred a local minimum over the global minimum in some cases, no individual parameter estimates were unreasonable in isolation. However, the parameter estimates were not always considered well determined and many varied significantly between patients.This was most marked for $k_{21}$, where the largest and smallest estimated values differed by a factor of 100, while the other parameters varied by roughly a factor of 10. The wide variability in the patient profiles makes it possible that (though unclear whether) this is plausible, and could be due to differences in the severity of the malaria, issues with the quality of the data, or other covariates (such covariates were not available for evaluation here).

Many people working in the field prefer to express elimination parameters in terms of half-lives. From the parameters in the parameter vector $\boldsymbol{p}$, the ARS half-life can be calculated as

$$t_{\frac{1}{2},\mathrm{ARS}} = \ln 2/(k_{42} + k_{\mathrm{e}2}), \tag{4.2}$$

and the DHA half-life as

$$t_{\frac{1}{2},\mathrm{DHA}} = \ln 2/k_{\mathrm{e}4}. \tag{4.3}$$

Estimates of these parameters obtained here (shown in Table 4.4) agree in range with those summarised in Morris et al. [39] (0.36–1.2 hours for ARS and 0.49–3.08 hours for DHA), but while Morris et al. [39] report that the DHA half-life is consistently longer than that of ARS, the same result was not found for all of the patients in this study; the reasons for this are unclear.

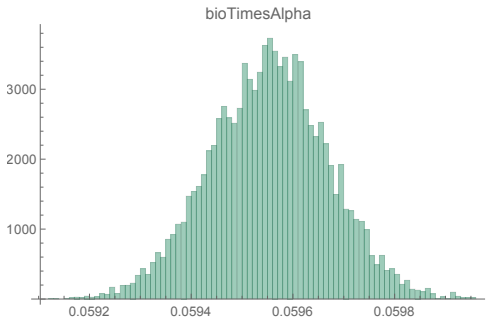
Model fitting was also conducted by relaxing one constraint at a time (still resulting in SGI model structures) to assess the effect on the para-

meter estimates. Doing so either did not significantly alter the parameter estimates, or otherwise did not generally improve fits visually (sometimes making them appear noticeably worse), and only marginally reduced the objective function values. The resulting estimates for some parameters were very close to their constrained values in some cases, while in others, the parameter estimates changed significantly and inconsistently, and their associated uncertainties increased also. When this occurred, the changes propagated to the other parameters too (due to the correlation), resulting in even wider variability of the parameters between patients. These results therefore provide evidence suggesting that the constraints imposed are as reasonable as could be hoped.

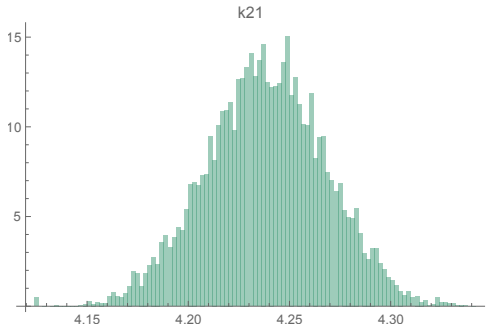
#### 4.1.2.3. Sequential Monte Carlo results

The Sequential Monte Carlo sampler described in subsubsection 2.7.6.1 was also applied to the four compartment pharmacokinetic model to explore both the model and data in more detail and to gain more experience with the SMC sampler. Independent uniform priors were assumed for all parameters, truncated to the plausible range to ensure the prior was proper and to improve the behaviour of the sampler compared to using a larger space.

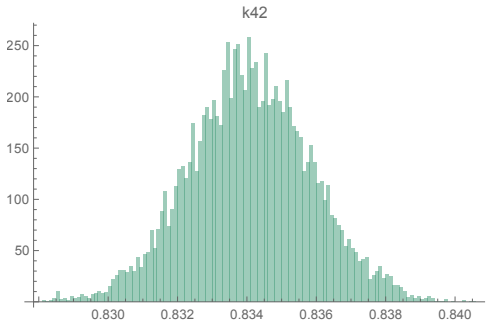
SMC estimated posterior marginal distributions for each of the parameters are presented in Figures 4.3 through to 4.7.



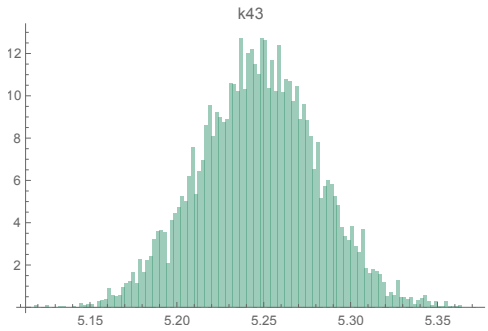

▲ Figure 4.3: SMC estimate of posterior marginal distribution for $b\alpha$ for patient A



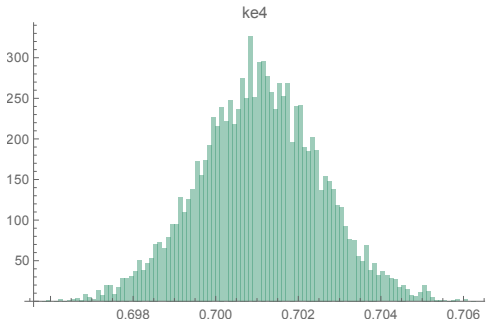

▲ Figure 4.4: SMC estimate of posterior marginal distribution for $k_{21}$ for patient A




▲ Figure 4.5: SMC estimate of posterior marginal distribution for $k_{42}$ for patient A

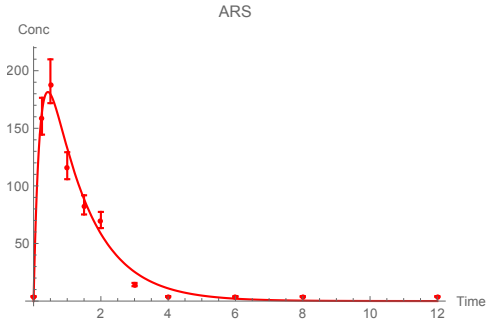



▲ Figure 4.6: SMC estimate of posterior marginal distribution for $k_{43}$ for patient A

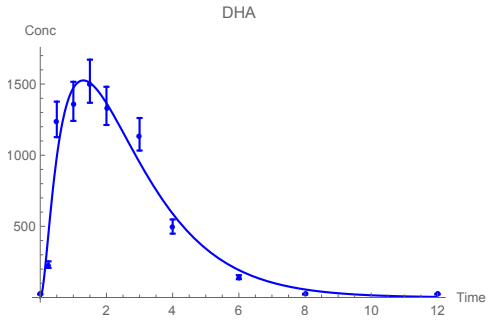



▲ Figure 4.7: SMC estimate of posterior marginal distribution for $k_{e4}$ for patient A

Pairwise marginal distributions have also been extracted and are shown in section A.1 to prevent cluttering this section. The model fits corresponding to the point with maximal posterior density are shown in Figures 4.8 and 4.9. It can be seen that the residuals for these plots are smaller than those in the previous fit (Figure 4.1).




▲ Figure 4.8: ARS model fit using point with maximal posterior density for patient A




▲ Figure 4.9: DHA model fit using point with maximal posterior density for patient A

## 4.2. Coupled parasite lifecycle model

### 4.2.1. Model development

As mentioned previously in section 2.3, a model must be designed with reference to the available data. A goal of this chapter is to model the data from subsection 2.3.3. Recall from that section that some patients were administered mefloquine in addition to artesunate on days 3 and 4. As the pharmacokinetic data for mefloquine are unavailable, it is difficult to model beyond day 3 for those patients and so this has not been done here.

It was previously concluded that there are changes in the stage-specific efficacy of the artemisinin derivatives between the two sites. It is therefore important that the model designed here is capable of capturing the same behaviour, and it will be interesting to observe whether the same effect is found.

The model is designed to be able to make use of *in vitro* data to help calibrate parameters that would otherwise be unidentifiable.

As the model is non-linear, unlike the pharmacokinetic model in subsection 2.6.1, it will be necessary to use the full quantity of drug administered as input to the model, rather than the per kg amount. Fortunately, the dataset contains the necessary data for this to be possible.

The most difficult part of the modelling was to ensure that the parasites' lifecycle is described continuously without having the effect of instantaneous ageing. It was initially thought that discrete-time models would be necessary for this, or models where each parasite replicate had an age parameter associated to it (which due to the large number of parasites initially present would be intractable), but a solution involving "shadow compartments" was found. This will be described in more detail in the relevant subsection.

### 4.2.2. Model

A novel model coupling the drug pharmacokinetics and the parasite life-cycle is proposed. The model structure has been tailored to correspond with the available data (see subsection 2.3.3), but no attempt is made to model the mefloquine kinetics or dynamics: data only for days prior to mefloquine administration will be used. Model predictions can be generated for the following days assuming mefloquine had not been administered, to examine the effect that the addition of mefloquine has had.

Although this model is an extension of the model in subsection 2.6.1, previous assumptions on those rate constants that are analogous will have to be revalidated or replaced.

The system model is not strictly compartmental so the system diagram has to be interpreted with care.

In introducing the model, it is helpful to first look at a subsection of it:

#### 4.2.2.1. Submodel

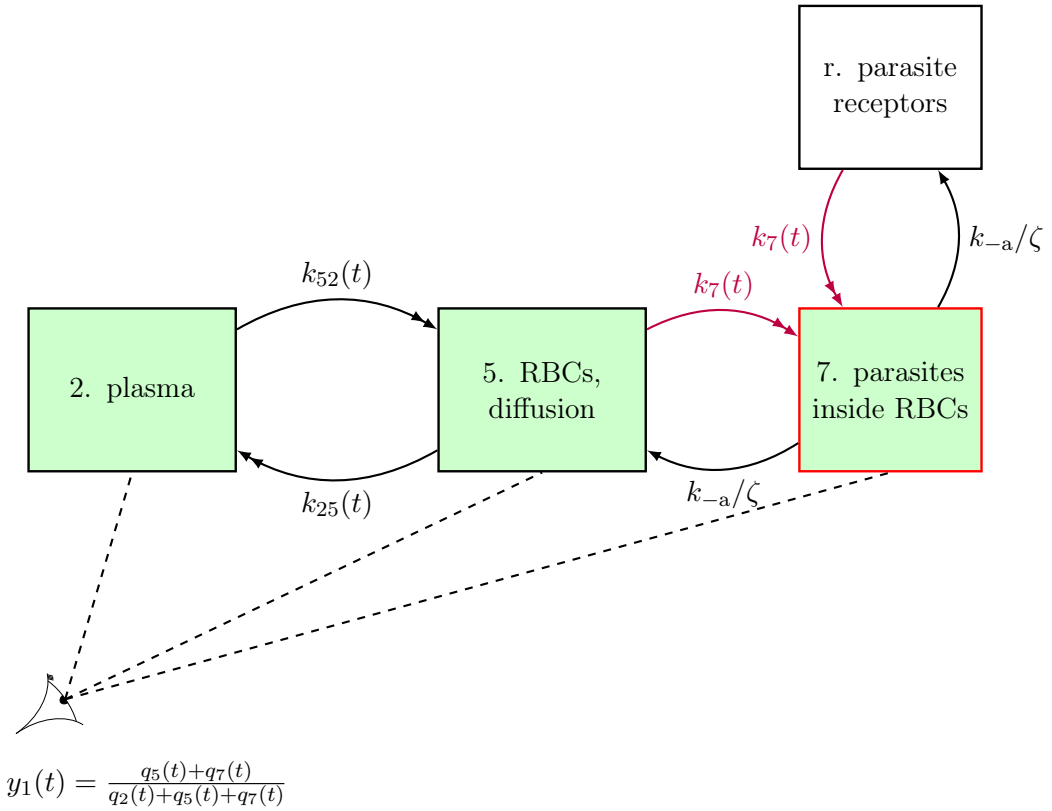
The proposed model essentially contains a submodel consisting of the plasma, red blood cell and parasite dynamics. It is simpler to present this submodel for a single drug first; see Figure 4.10. The *in vitro* data can be modelled by this submodel as an initial step, before considering the full model.

The submodel contains 4 compartments, of which 3 represent the location of the drug: a plasma compartment, a red blood cell (RBC) compartment, and a parasite compartment. The parasite compartment in fact represents the bound form of the drug, after binding with receptors in the parasites. Representing the pool of available receptors in the parasites as the 4th compartment therefore completes the selection of model states.

The drug in the plasma compartment is modelled as flowing by diffusion to and from the red blood cells. Ligand-receptor kinetics are then used to describe the binding of the drug inside the red blood cells to the parasitised portion of the red blood cells. Although radio-labelled studies have shown that drug bound to the parasites does not appear to be displaced, it is not known *a priori* whether it is possible for the drug to dissociate from the parasite receptors, albeit very slowly, or whether no dissociation can occur at all. The model therefore leaves in the possibility of dissociation, which is effectively removed when the dissociation rate constant is zero.

The *in vitro* study measures the proportion of uptake of the radio-labelled artemisinin in the "pellet" compared to the total (pellet plus supernatant) after centrifuging. The pellet therefore includes the red blood cells and the parasites within the red blood cells, while the supernatant corresponds to the plasma.

The parameters in this submodel and their meanings are listed in Table 4.5.




▲ Figure 4.10: *In vitro* submodel for drug uptake into red blood cells

The compartments within the model are numbered to correspond to those in the full model. The model is parameterised in terms of the haematocrit $\zeta$ (considered proportional to the number of RBCs), parasitaemia percentage $p$, and the total volume of blood $v_2$. These are considered static parameters for the *in vitro* experiments (rather than potentially changing with time). The initial concentration of the drug in the plasma is a known initial condition of the model, while the initial number of parasite receptors is only known up to a constant of proportionality. In the *in vitro* experiments, the volumes of blood and parasitaemia are

known, but the parasite lifecycle stage is not measured. Therefore, for simplicity and identifiability, the infected red blood cell compartment is not split by parasite stage.

The equations for the submodel are as follows:

$$\begin{cases} q_2'(t) &= \mu\left(\frac{q_5(t)}{v_5} - \frac{q_2(t)}{v_2}\right) \\ q_5'(t) &= -\mu\left(\frac{q_5(t)}{v_5} - \frac{q_2(t)}{v_2}\right) - \frac{k_{\mathrm{a}} q_5(t) q_{\mathrm{r}}(t)}{\zeta} + \frac{k_{\text{-a}} q_7(t)}{\zeta} \\ q_{\mathrm{r}}'(t) &= \frac{k_{\text{-a}} q_7(t)}{\zeta} - \frac{k_{\mathrm{a}} q_5(t) q_{\mathrm{r}}(t)}{\zeta} \\ q_7'(t) &= \frac{k_{\mathrm{a}} q_5(t) q_{\mathrm{r}}(t)}{\zeta} - \frac{k_{\text{-a}} q_7(t)}{\zeta} \\ q_2(0) &= c_0 v_2 \\ q_5(0) &= 0 \\ q_7(0) &= 0 \\ q_{\mathrm{r}}(0) &= \kappa\zeta p \\ y_1(t) &= \frac{q_5(t)+q_7(t)}{q_2(t)+q_5(t)+q_7(t)} \end{cases} \tag{4.4}$$

where $v_5 = \zeta\lambda_{\mathrm{c}}$, $v_2 = 0.7$ ml. i.e. in the above diagram,

$$\begin{aligned} k_{52}(t) &= \max\left\{\mu\left(\frac{q_5(t)}{v_5} - \frac{q_2(t)}{v_2}\right), 0\right\} \\ k_{25}(t) &= \max\left\{-\mu\left(\frac{q_5(t)}{v_5} - \frac{q_2(t)}{v_2}\right), 0\right\} \\ k_7(t) &= k_{\mathrm{a}} q_5(t) q_{\mathrm{r}}(t)/\zeta. \end{aligned} \tag{4.5}$$

The parameter $\lambda_{\mathrm{c}}$ provides an extra degree of freedom for the concentrations in red blood cells and whole blood to differ by a fixed ratio.

The *in vitro* data are only available for artemisinin, so the submodel will be initially calibrated based on this, though adjustments may have to be made for the different derivatives used in the full model (artesunate and dihydroartemisinin). The calibration of the submodel to the data is made under the assumption that the radio-tracer stayed bound to the artemisinin, which therefore includes the assumption that there was no metabolism of artemisinin, either into other artemisinin derivatives or into inactive metabolites. This assumption is necessary given the

limitations of the data provided, and will still produce useful insight. However, before incorporating parameters derived from these data into the full model, consideration should be given to necessary corrections, including the corrections already necessary for the different derivatives.

#### 4.2.2.2. Submodel structural identifiability

Before attempting to fit the *in vitro* data to the submodel, it is important to first check that the model is structurally identifiable. The model is an uncontrolled model with 5 unknown parameters, 3 known parameters, and no input. The model parameters are summarised in Table 4.5.

| Parameter | Description | Known/unknown |
|---|---|---|
| $\kappa$ | number of receptors per unit of infected hematocrit | unknown |
| $k_{\mathrm{a}}$ | drug-parasite receptor binding rate constant | unknown |
| $k_{\text{-a}}$ | drug-parasite receptor unbinding rate constant | unknown |
| $\mu$ | permeability of red blood cells | unknown |
| $\lambda_{\mathrm{c}}$ | concentration factor for diffusion equilibrium | unknown |
| $\zeta$ | hematocrit | known |
| $c_0$ | initial concentration of drug | known |
| $p$ | parasitaemia | known |
| $v_2$ | total volume of blood | known |

Table 4.5: Table of parameters in the artemisinin red blood cell uptake submodel

The Taylor series method is easily applied to this model. Calculating six Taylor series coefficients of the observation function shows the model to be structurally globally identifiable.

It may be of interest to also determine the situation when used with only uninfected red blood cell data, i.e. when $p = 0$. Re-applying the same method under these circumstances, only $\mu$ and $\lambda_{\mathrm{c}}$ are found to be structurally globally identifiable, with the remaining unknown paramet-

ers $\kappa$, $k_{\mathrm{a}}$ and $k_{\text{-a}}$ being structurally unidentifiable (in fact, not appearing in derivatives of the output function at all to any order checked, though it is noted that this method is technically inconclusive unless all Taylor series cofficients are checked), as would be expected.

#### 4.2.2.3. Submodel parameter estimates

Based on the above structural identifiability analysis, it would clearly be possible (under the assumption of continuous noise-free data) to estimate parameters $\mu$ and $\lambda_{\mathrm{c}}$ from the uninfected red blood cell data, and then to consider these parameters as known when estimating the remaining parameters from the infected red blood cell data. However, there are few data points available and it is probably more helpful to estimate all the parameters using all the data, so that is what was done here.

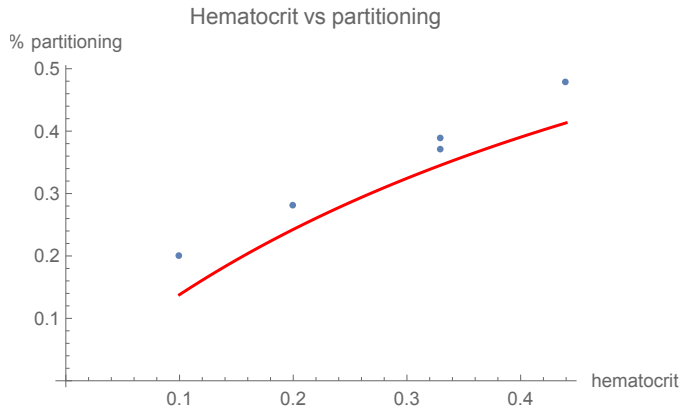
Parameter estimates were determined using the *in vitro* data for artemisinin (introduced in subsection 2.3.2) and are shown in Table 4.6, and the corresponding model predictions are shown in red in Figures 4.11 through to 4.17. Note that the same parameters are used to generate all plots (excepting the independent parameters $\zeta, p, c_0$), so that even though some plots do not look to have fitted perfectly, the reason for that is the model has been constrained by the data points shown in the other plots. The systematic error that appear consist of too-high model estimates in some plots, and too-low model estimates in others, suggesting that the fitting has performed as well as it can under the constraints. It was unknown whether the flow described by parameter $k_{\text{-a}}$ was in fact present in the system being modelled. The numerical optimiser was estimating parameter $k_{\text{-a}}$ to be close to 0 (around $10^{-9}$) so this parameter was instead forced to 0 to improve estimates of the remaining parameters.

As noted earlier, these data are for artemisinin and so these parameters cannot necessarily be used directly for ARS or DHA, but if the model behaves well for artemisinin then an appropriate change to the

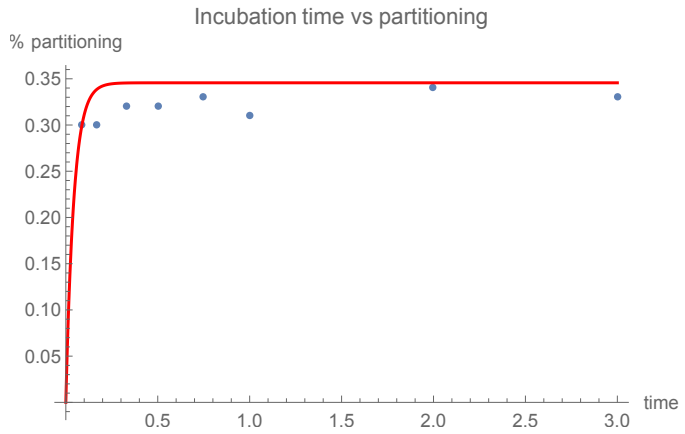
parameter values should render it suitable for those other derivatives. As this derivative is not of much interest here, a full statistical analysis is not carried out. It is sufficient to determine visually whether the model is behaving appropriately.

| Parameter | Description | Value |
|---|---|---|
| $\kappa$ | number of receptors per unit of infected hematocrit | 14,500 |
| $k_{\mathrm{a}}$ | drug-parasite receptor binding rate constant | 0.00225 |
| $k_{\text{-a}}$ | drug-parasite receptor unbinding rate constant | 0 (forced) |
| $\mu$ | permeability of red blood cells | 5.58 |
| $\lambda_{\mathrm{c}}$ | concentration factor for diffusion equilibrium | 1.12 |

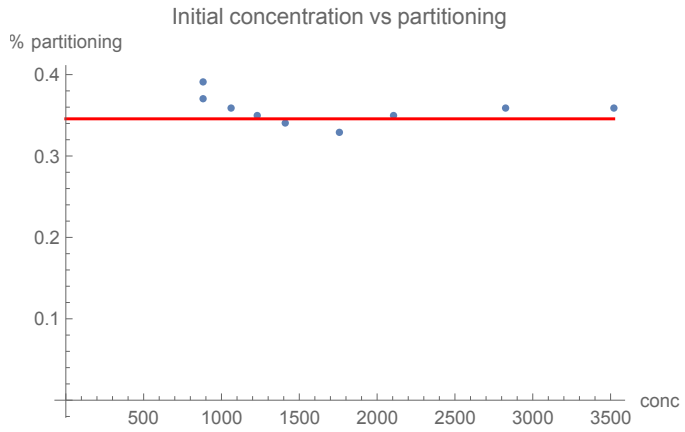
Table 4.6: Table of parameter estimates in the artemisinin red blood cell uptake submodel



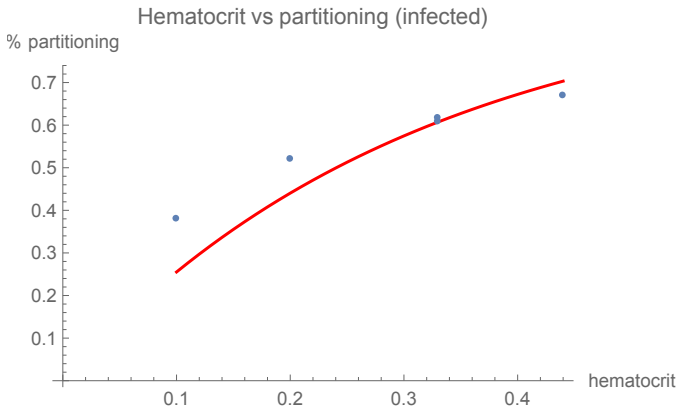

▲ Figure 4.11: Hematocrit against partitioning coefficient for uninfected cells (with $c_0 = 880$ nM, $t =$ 2 hours, $p = 0\%$), data and model



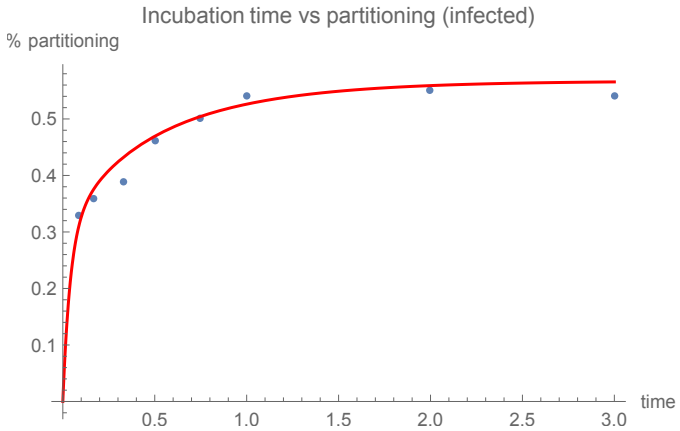

▲ Figure 4.12: Incubation time (hours) against partitioning coefficient for uninfected cells (with $c_0 = 1410$ nM, $\zeta = 33\%$, $p = 0\%$), data and model



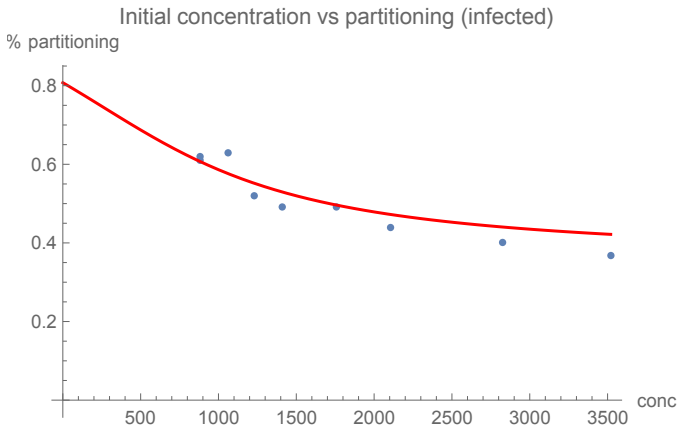

▲ Figure 4.13: Initial concentration against partitioning coefficient for uninfected cells (with $t = 2$ hours, $\zeta = 33\%$, $p = 0\%$), data and model



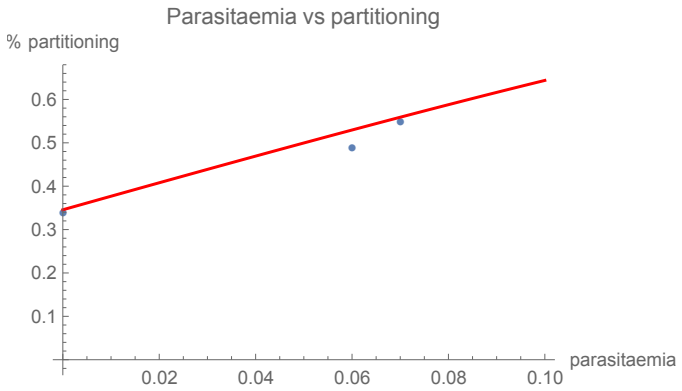

▲ Figure 4.14: Hematocrit against partitioning coefficient for infected cells (with $c_0 = 880$ nM, $t = 2$ hours, $p = 6\%$), data and model




▲ Figure 4.15: Incubation time (hours) against partitioning coefficient for infected cells (with $c_0 = 1410$ nM, $\zeta = 33\%$, $p = 7\%$), data and model




▲ Figure 4.16: Initial concentration against partitioning coefficient for infected cells (with $t = 2$ hours, $\zeta = 33\%$, $p = 6\%$), data and model




▲ Figure 4.17: Parasitaemia against partitioning coefficient (with $c_0 = 1410$ nM, $t = 2$ hours, $\zeta = 33\%$), data and model

It is clear that the model appears to capture the correct shape of the behaviour for each 'slice' of the available data. Some systematic errors are clearly seen in the uninfected haematocrit and uninfected incubation time slices, with under- and over-estimates of the data, respectively. It is also clear that the first data point for measuring uptake against time was too late to capture the initial uptake period (or rather, the uptake was so fast so as to prevent observation of this), so it seems that it should be hard to calibrate the uptake parameter $\mu$ exactly. Although no problems were reported by the numerical optimisation scheme used, no sensitivity analysis was conducted.

#### 4.2.2.4. Ageing model

As the lifecycle stage of the parasites plays a key role in both the time of replication of the parasites, and in the drug-parasite interactions, it is necessary to model the lifecycle stage of the parasites, or equivalently, their ages. There can be in excess of $10^{12}$ parasites in total in severe cases, corresponding to 200,000 parasites per microlitre of blood [93]. The parasites are also often synchronised or near-synchronised. It is therefore neither feasible nor particularly useful to keep track of the age of each individual parasite: keeping track of the lifecycle stage is a sufficient and simpler alternative.

To decide on the appropriate lifecycle stages to lump parasites into, first recall that there are 4 age classes into which malaria parasites are often distinguished: merozoites ($\sim$ 2 mins), ring forms ($\sim$ 12 hours), trophozoites ($\sim$ 24 hours) and schizonts ($\sim$ 12 hours). From a modelling perspective, the merozoites are transient; existing only on a negligible time scale relative to the remaining stages and time scales of the drug model, hence merozoites need not be incorporated as a distinct stage but parasites in this form can be lumped into either of the adjacent stages. The remaining three stages have different durations, and the susceptibility of the young and mature ring forms is thought to differ. Hence, it is helpful to reassign the stages to be more practical for the purposes of the model, and such that each has a similar duration. Eight age classes are therefore used, each of 6 hours duration: young ring forms, mature ring forms, young circulating trophozoites, mature circulating trophozoties, youngly sequestered trophozoites, mature sequestered trophozites, young schizonts and mature schizonts. There is thought to be little value in having age classes of shorter duration, due to the synchronous nature of the parasites and the forseeable difficulty in the practical identifiability of stage-specific parameters for shorter age classes.

Now that the age classes have been decided upon, the actual mech-

anics of the progression between age classes needs to be elucidated. Instantaneous progression of parasites from one age class to the next might be simple to model but, even if parasites are generally age-synchronised, this would be an undesirable property because there would be sudden discontinuous jumps in the numbers of parasites due to parasite death and replication. Hence a continuous-time model of ageing would be preferred. The simplest such model that occurred to the author consisted of each age class connected in series to its successor, in the standard first-order manner as is usual for many compartmental models. However, it is clear that first-order kinetics can only account for the mean behaviour of essentially exponentially distributed holding times (or extensions thereof with moderate complexity such as Erlang distributions) [94]. Such holding time distributions have non-zero probability densities for all positive times, and so are not really suitable to describe ageing. For instance, some non-zero proportion of the parasites age instantaneously in these models, and to make this proportion negligible requires a large number of intermediate compartments, which becomes infeasible.

It is possible to maintain continuous-time behaviour without this difficulty, however, by specifying that the flow rates between age classes undergo discrete-time step changes. Specifically, to describe a 48 hour lifecycle using eight age classes of equal duration, the eight age classes can be represented with two compartments each; one acting (at any given time) as a 'live' compartment and one as a 'shadow' compartment for each, giving 16 compartments in total. The terms 'live' and 'shadow' compartments are slightly misleading as their roles periodically alternate, but the 'live' compartment can be thought of as 'initially live' when $t = 0$ (i.e. when modelling begins) and the other as 'initially shadowing'. Live and shadow compartments are symmetrical in terms of their equations and hence equivalent in nature, just it is convenient to give them names that highlight the fact that they represent part of the same measured quantity (parasites in a certain age class). The distinction is

that the live compartment represents parasites prior to the transition to the next age class, and the shadow compartment represents those after the transition has completed (or vice versa, depending on the stage in the periodic cycle).

As the ageing is to happen continuously rather than instantaneously, a time window $\tau_{\text{change}}$ for ageing must be chosen, e.g. the advancing from one age class to the next could take place over a window of length $\tau_{\text{change}} =$ 1 hour, say. Suppose that at $t = 0$, the parasites were already $\delta t$ hours into their current age classes. i.e. in the age class of 0–6 hours, all the parasites are $\sim \delta t$ hours old, and similarly those in the 6–12 hour age class are all $\sim 6 + \delta t$ hours old, and so on. The advancement process then could be modelled as taking place between time $5.5 - \delta t$ and $6.5 - \delta t$ hours, every 6 hours. Generally, the switching begins $\tau_{\text{change}}/2$ time units before the lifecycle duration is reached, and completes $\tau_{\text{change}}/2$ time units after. This is achieved in the model by toggling on a flow during that time interval, such that the parasites in the 'live' compartments move to the age-advanced 'shadow' compartments, as shown in Figure 4.18. (In the next time interval 6 hours later, the flows are from these 'shadow' compartments into the 'live' compartments, as their roles have reversed.) Zeroth-order (constant) dynamics are chosen, but once all the parasites in the donor compartment have migrated (or died), that particular flow is toggled back to 0.

A background rate of parasite kill is specified here for illustrative purposes. This is replaced with a drug-dependent rate in the full coupled model.

At the beginning of the discrete event at time $t^*$ with $\cdot$ representing either the live compartments l or the shadow compartments s as appropriate, set $k_i^{(\cdot)} \leftarrow p_i^{(\cdot)}(t^*)/\tau_{\text{change}}$ where $\tau_{\text{change}}$ is the time interval over which the ageing process takes place. (To validate this, note that the time domain equation for the contribution of the zeroth-order ageing flow after $\tau_{\text{change}}$ hours is $-k_i\tau_{\text{change}}$, which with the above definition of

$k_i$ is equal to $-p_i^{(\cdot)}$. i.e. this corresponds to all parasites having left the compartment. Of course, with background/drug-induced death, it may not take $\tau_{\text{change}}$ for the compartment to become empty, but the ageing rates are still defined in this way.)

The differential equations governing the parasite lifecycle model are conceptually simple:

$$
\begin{aligned}
{p_1^{(\mathrm{l})}}'(t) &= -d_1 p_1^{(\mathrm{l})}(t) - k_1^{(\mathrm{l})} + M k_8^{(\mathrm{s})} \\
{p_2^{(\mathrm{l})}}'(t) &= -d_2 p_2^{(\mathrm{l})}(t) - k_2^{(\mathrm{l})} + k_1^{(\mathrm{s})} \\
{p_3^{(\mathrm{l})}}'(t) &= -d_3 p_3^{(\mathrm{l})}(t) - k_3^{(\mathrm{l})} + k_2^{(\mathrm{s})} \\
{p_4^{(\mathrm{l})}}'(t) &= -d_4 p_4^{(\mathrm{l})}(t) - k_4^{(\mathrm{l})} + k_3^{(\mathrm{s})} \\
{p_5^{(\mathrm{l})}}'(t) &= -d_5 p_5^{(\mathrm{l})}(t) - k_5^{(\mathrm{l})} + k_4^{(\mathrm{s})} \\
{p_6^{(\mathrm{l})}}'(t) &= -d_6 p_6^{(\mathrm{l})}(t) - k_6^{(\mathrm{l})} + k_5^{(\mathrm{s})} \\
{p_7^{(\mathrm{l})}}'(t) &= -d_7 p_7^{(\mathrm{l})}(t) - k_7^{(\mathrm{l})} + k_6^{(\mathrm{s})} \\
{p_8^{(\mathrm{l})}}'(t) &= -d_8 p_8^{(\mathrm{l})}(t) - k_8^{(\mathrm{l})} + k_7^{(\mathrm{s})} \\
{p_1^{(\mathrm{s})}}'(t) &= -d_1 p_1^{(\mathrm{s})}(t) + M k_8^{(\mathrm{l})} - k_1^{(\mathrm{s})} \\
{p_2^{(\mathrm{s})}}'(t) &= -d_2 p_2^{(\mathrm{s})}(t) + k_1^{(\mathrm{l})} - k_2^{(\mathrm{s})} \\
{p_3^{(\mathrm{s})}}'(t) &= -d_3 p_3^{(\mathrm{s})}(t) + k_2^{(\mathrm{l})} - k_3^{(\mathrm{s})} \\
{p_4^{(\mathrm{s})}}'(t) &= -d_4 p_4^{(\mathrm{s})}(t) + k_3^{(\mathrm{l})} - k_4^{(\mathrm{s})} \\
{p_5^{(\mathrm{s})}}'(t) &= -d_5 p_5^{(\mathrm{s})}(t) + k_4^{(\mathrm{l})} - k_5^{(\mathrm{s})} \\
{p_6^{(\mathrm{s})}}'(t) &= -d_6 p_6^{(\mathrm{s})}(t) + k_5^{(\mathrm{l})} - k_6^{(\mathrm{s})} \\
{p_7^{(\mathrm{s})}}'(t) &= -d_7 p_7^{(\mathrm{s})}(t) + k_6^{(\mathrm{l})} - k_7^{(\mathrm{s})} \\
{p_8^{(\mathrm{s})}}'(t) &= -d_8 p_8^{(\mathrm{s})}(t) + k_7^{(\mathrm{l})} - k_8^{(\mathrm{s})}
\end{aligned}
$$

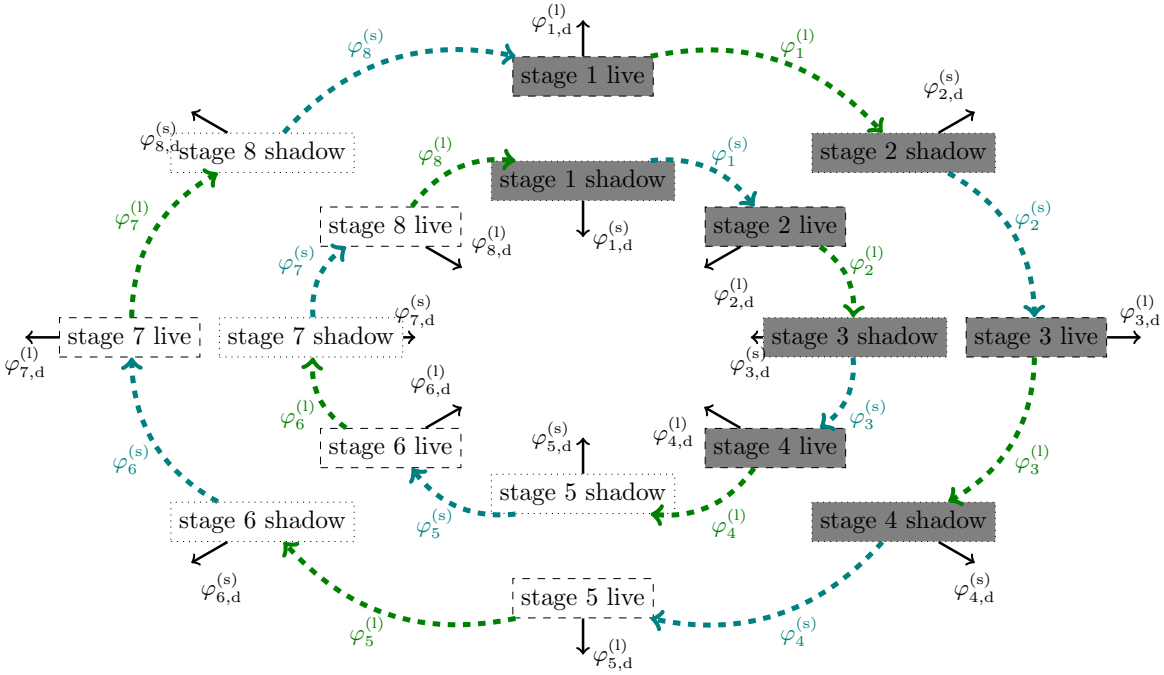
where the $k$ rate parameters are driven by discrete events. The discrete

events are:

$$\begin{aligned}
&\text{when } p_i^{(\cdot)} \text{ reaches } 0, && \text{set } k_i^{(\cdot)} \leftarrow 0 \\
&\text{when } t + t_{\text{offset}} \equiv 0 \mod 12, && \text{set } k_i^{(\text{s})} \leftarrow p_i^{(\text{s})}(t)/\tau_{\text{change}} \\
&\text{when } t + 6 + t_{\text{offset}} \equiv 0 \mod 12, && \text{set } k_i^{(\text{l})} \leftarrow p_i^{(\text{l})}(t)/\tau_{\text{change}}.
\end{aligned}$$

The $d_i$ terms represent first-order parasite death rates, which in the full model are coupled to the pharmacokinetics of the drugs, but for illustration here they consist solely of a constant background death rate. The model also incorporates a further discrete event to ensure if the density of the parasites falls below a predefined threshold, then the infection has effectively died out. Without this extra event, after enough cycles in the absence of a significant background/drug-induced death rate, the parasites might be shown to return to a substantial density under the model, which isn't always observed. The threshold can of course be configured to a plausible value.

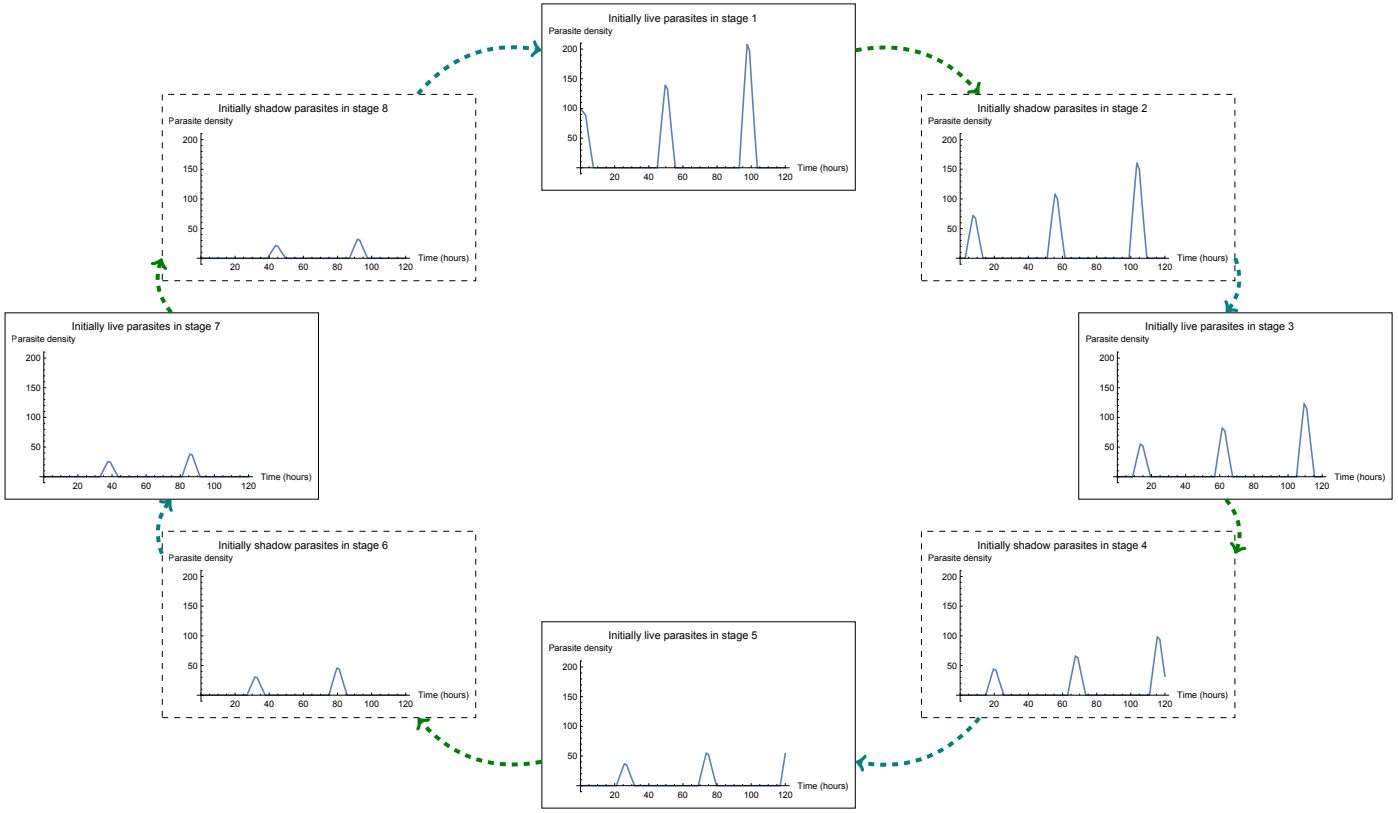
Figure 4.18 shows the compartments representing the age stages within the lifecycle and the connections between the live and shadow compartments.



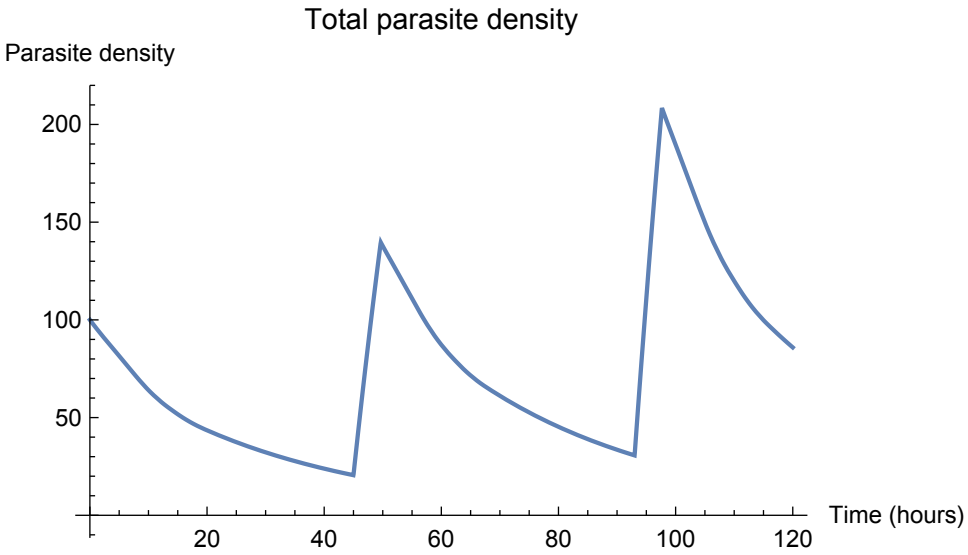

▲ Figure 4.18: Parasite lifecycle model showing connectivity between live and shadow compartments.
The solid flows are always enabled. Either the dashed green flows or the dashed teal flows are active (until the respective donor compartment runs out of parasites), both sets of dashed flows are never active simultaneously. The number of parasites in a given age class is always taken to be the sum of the quantities in each of the live and shadow compartments for that age class.
Compartments with a grey background are observable, while compartments with a white background are unobservable as they represent sequestered stages of the lifecycle.

An example of the behaviour of this parasite lifecycle model is shown in Figure 4.19. For this figure, the model was started with a single initial parasite population entirely within age class 1, and uses constant background death rate parameters. In the figure, it can be seen that the parasites that initially started in age class 1 move into the successive age classes after the expected time intervals. Figures 4.20 and 4.21 show the aggregated time courses of total and observable parasite densities across the age classes, respectively. The effect of the background parasite death rate can be seen, which stops the parasite population from exploding. It is also clear that the unobservability of the sequestered parasites has a significant effect on the observed parasitaemia, and that even though it may appear that all the parasites have died, this is not necessarily the case. Observations must therefore be made of parasitaemia at sufficient intervals to determine whether the infection is cured or whether the parasites are simply sequestered.

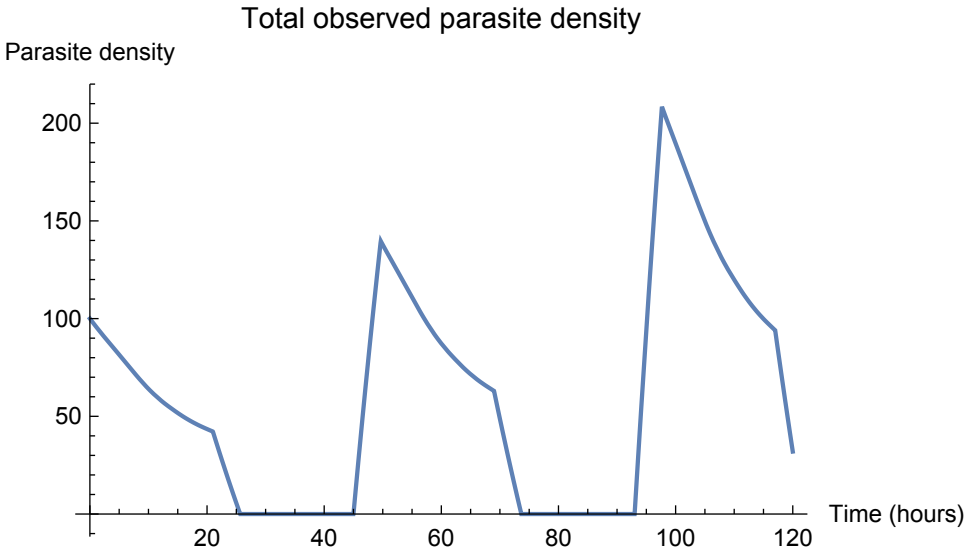
Note that the model is fully capable of having an initial parasite distribution that consists of multiple age classes, but to aid clarity, that situation has not been shown in the simulations here.




▲ Figure 4.19: Parasite lifecycle model simulation with all parasites initially in age class 1. The inner circle of age classes from Figure 4.18 are identically zero for the whole simulation duration and are omitted here.




▲ Figure 4.20: Total density of parasites across all classes in the simulation shown in Figure 4.19




▲ Figure 4.21: Total density of all observable (non-sequestered) parasites in the simulation shown in Figure 4.19

#### 4.2.2.5. Full model

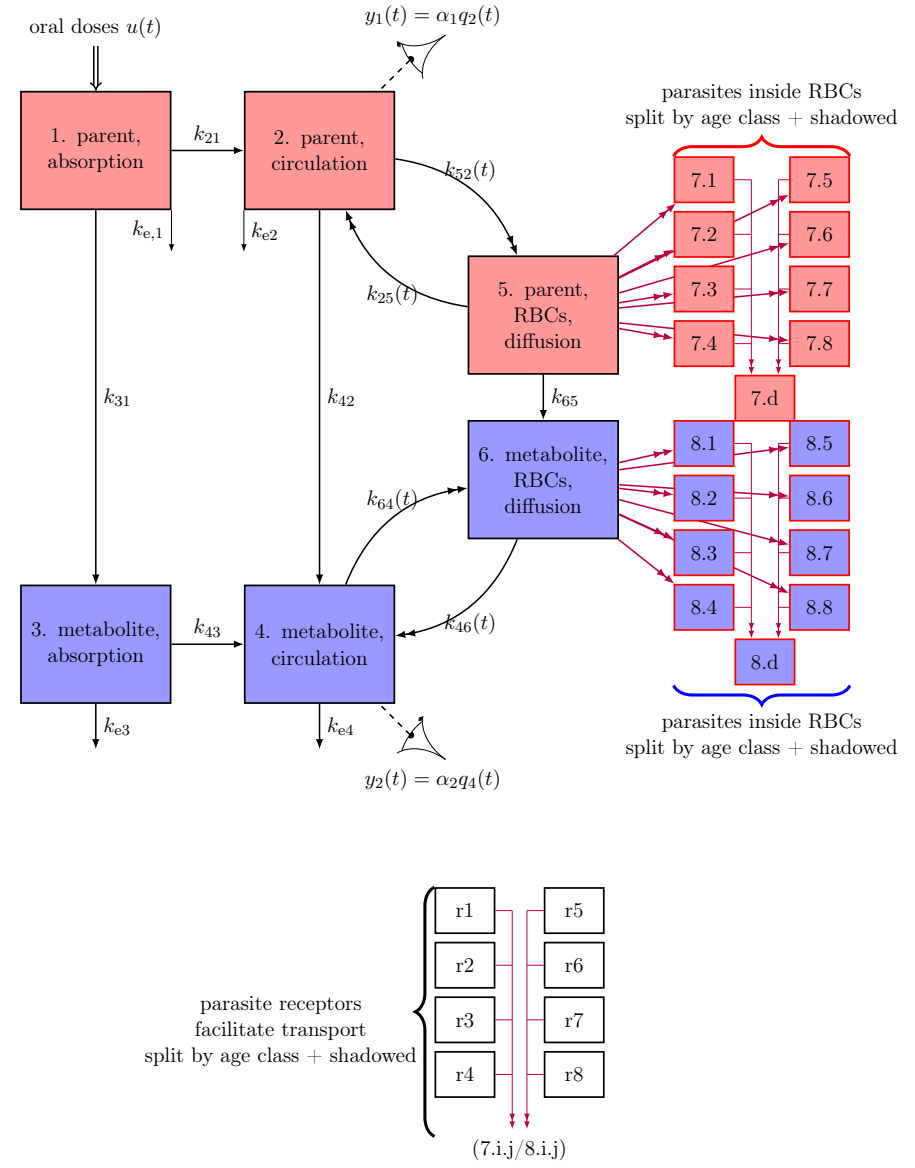
The full model is now presented, which couples a drug pharmacokinetic/pharmacodynamic model (Figure 4.22) with the parasite lifecycle/pharmacodynamic model introduced in the previous subsection (Figure 4.18), but now with death rates coupled to the pharmacokinetic model. Live and shadow compartments are also adopted in exactly the same way to keep track of the amount of the drugs inside the red blood cells.

The absorption compartments from the four compartment pharmacokinetic model, not relevant in the submodel above, are restored to the full pharmacokinetic/pharmacodynamic model for *in vivo* use. The parasite receptors and drugs inside parasites are now split into parasite age classes, and use shadow compartments to facilitate what is essentially the rotation of contents of these compartments when the parasites age.

Note that the parasitised red blood cell aspects of the drug model are coupled to the lifecycle model, and the lifecycle model is coupled to the drug concentration in parasitised red blood cells. The lifecycle model triggers discrete events every 6 hours to toggle on the relevant shadow flows, and discrete events are automatically triggered to toggle off the zeroth-order shadow flows once the donor compartments become empty. In between the discrete events, the model is described by a standard system of coupled first-order ordinary differential equations.

The drug model is as simple as possible while still capturing all the key features of the biological processes involved. As the study again relates to orally administered artesunate (ARS) and its principle metabolite dihydroartemisinin (DHA), the model is again divided into parent/metabolite compartments with input to the parent absorption/gut compartment.

As in the submodel, there are now compartments for red blood cells, and parasites within some of those red blood cells, in addition to the




Compartments 7.1–7.8, 8.1–8.8 and r1–r8 also have shadow copies but only the "initially live" set of compartments is shown to prevent obscuring the more interesting aspects of the model.

▲ Figure 4.22: Diagram of drug model component. Note that shadow compartments are also used here but are not shown to avoid obscuring the rest of the diagram.

plasma circulation compartments from the previous pharmacokinetic model. Obviously each subject has a large number of red blood cells, but for simplicity, the model only includes one red blood cell compartment for each drug and one parasite compartment for each drug, and thus captures only the average dynamics. The average quantities of drug inside the red blood cells may therefore change during parasite death and replication, but it is not felt that this affects the model utility.

Observations are made of the concentrations of ARS and DHA in plasma (rather than in whole blood), and of the non-sequestered parasite burden. While the drug concentrations are only measured in plasma in the available datasets, representing whole blood distribution in the model is more accurate than compensating for this through the observation gain parameters as in subsection 2.6.1, and concentrations in the blood cells are physiologically relevant. The remaining aspects of the model are analogous to the corresponding respective components of the submodel and previous pharmacokinetic model. Note that the model does not degenerate to the pharmacokinetic model in the case where there is no malaria, due to the inclusion of the red blood cells with non-linear dynamics.

Distribution of ARS and DHA into uninfected cells is reversible, but into infected cells is irreversible (due to conversion to free radical or inactive and unmeasured forms, as confirmed when fitting the submodel to the *in vitro* data); the red blood cells act as a sink for the drugs.

It is possible that metabolism of ARS to DHA will still occur to the ARS present inside red blood cells. The model does not allow ARS to metabolise once bound to the parasites, because this binding is irreversible and no further observations are made of either drug following such binding, and so the model would not be able to support nor benefit from such a metabolism parameter.

The initial conditions for the lifecycle component of the model, including the initial time offset, are not known and have to be estimated.

Artesunate is administered orally, and initially enters impulsively at

each dosing time into an absorption compartment. First-order irreversible absorption occurs into the systemic circulation (this compartment lumps together multiple organs and blood, and has an **apparent** volume of distribution). As in the submodel, the drug then undergoes passive diffusion into red blood cells; this diffusion is reversible and driven by the concentration gradient. At any of these stages, the drug can be metabolised (hydrolysed) into the more active derivative dihydroartemisinin, whence it can then undergo the remaining processes.

For subjects with malaria, both drugs are actively and irreversibly transported into the parasite component of infected red blood cells. For full model power, parasites (and hence drug in parasitised cells) are split into 8 age classes. Once either drug derivative is inside a parasite, its pharmacodynamic effect begins: parasites are killed, arrested or their growth stunted. Once the pharmacodynamic effect has taken place, the drug is no longer able to have any further effect — it is said to be denatured. In the model, this is represented by the drug flowing into its respective denatured compartment, labelled 7.d and 8.d in the diagram. These compartments are not strictly necessary as elimination from the system model would be an equally valid modelling choice, but it may be helpful to keep drug of the quantities of the drugs that have been denatured to evaluate how significant this effective route of elimination is.

Meanwhile, the parasites are progressing through their lifecycle: ageing, replicating, and showing affects from the interaction with the artemisinin derivatives. While a parasite matures (advances into the next age class) after having taken up some artemisinin, the respective amount of the artemisinin simultaneously moves into the respective age class also. When a parasite dies, the artemisinin it contains flows into the denatured/sink compartment. In order to prevent immediate further maturation through the next class and so on, shadow compartments are used. Essentially, there are two distinct copies of each compartment that

is linked to a specific parasite stage.

The volumes of the red blood cell and parasite-bound drug compartments are affected by the haematocrit and parasite burden at any particular time.

Uptake of drug into parasites of a particular age class requires availability of parasite receptors in that age class. This mechanism ensures that uptake of drug into parasites is saturable. The differential equations are therefore:

$$\begin{aligned}
\tfrac{\mathrm{d}q_1}{\mathrm{d}t} &= -(k_{21}+k_{31}+k_{\mathrm{e},1})q_1(t)+u(t)\\
\tfrac{\mathrm{d}q_2}{\mathrm{d}t} &= k_{21}q_1(t)-(k_{42}+k_{\mathrm{e}2})q_2(t)+k_{25}(t)-k_{52}(t)\\
\tfrac{\mathrm{d}q_3}{\mathrm{d}t} &= k_{31}q_1(t)-(k_{43}+k_{\mathrm{e}3})q_3(t)\\
\tfrac{\mathrm{d}q_4}{\mathrm{d}t} &= k_{42}q_2(t)+k_{43}q_3(t)-k_{\mathrm{e}4}q_4(t)+k_{46}(t)-k_{64}(t)\\
\tfrac{\mathrm{d}q_5}{\mathrm{d}t} &= -\textstyle\sum_{i=1}^{4}\sum_{x\in\{\mathrm{lead,shadow}\}}k_{7.x.i,5}(t)+k_{52}(t)-k_{25}(t)-k_{65}q_5(t)\\
\tfrac{\mathrm{d}q_6}{\mathrm{d}t} &= -\textstyle\sum_{i=1}^{4}\sum_{x\in\{\mathrm{lead,shadow}\}}k_{8.x.i,6}(t)+k_{64}(t)-k_{46}(t)+k_{65}q_5(t)\\
\tfrac{\mathrm{d}q_{r.x.i}}{\mathrm{d}t} &= -\tfrac{1}{\zeta(t)}(k_{\mathrm{a,ARS},i}q_5(t)+k_{\mathrm{a,DHA},i}q_6(t))q_{r.x.i}(t)+f_{\mathrm{shadow},r.x.i}(t)\\
\tfrac{\mathrm{d}q_{7.x.i}}{\mathrm{d}t} &= q_5(t)k_{\mathrm{a,ARS},i}q_{r.x.i}(t)+f_{\mathrm{shadow},7.x.i}(t)-d_{7.x.i}(t)\\
\tfrac{\mathrm{d}q_{8.x.i}}{\mathrm{d}t} &= q_6(t)k_{\mathrm{a,DHA},i}q_{r.x.i}(t)+f_{\mathrm{shadow},8.x.i}(t)-d_{8.x.i}(t)
\end{aligned} \tag{4.6}$$

where $q_{r.x.i}$ denotes the number of unbound parasite receptors for parasites in age class $i$, $x$ is either the lead or shadow compartment, $k_{\mathrm{a},j,i}$ denotes the uptake rate of drug $j$ into a parasite in age class $i$.

To ensure the above equations are not unwieldly lengthy, the non-

linear flows have been expressed as separate functions:

$$k_{52}(t) = \max\left\{\mu_{\text{ARS}}(c_2(t) - c_5(t)), 0\right\} \tag{4.7}$$

$$k_{25}(t) = \max\left\{\mu_{\text{ARS}}(c_5(t) - c_2(t)), 0\right\} \tag{4.8}$$

$$k_{64}(t) = \max\left\{\mu_{\text{DHA}}(c_4(t) - c_6(t)), 0\right\} \tag{4.9}$$

$$k_{46}(t) = \max\left\{\mu_{\text{DHA}}(c_6(t) - c_4(t)), 0\right\} \tag{4.10}$$

$$k_{7.x.i,5}(t) = \frac{k_{\text{a,ARS},i}}{\zeta(t)} q_5(t) q_{r.x.i}(t) \tag{4.11}$$

$$k_{8.x.i,6}(t) = \frac{k_{\text{a,DHA},i}}{\zeta(t)} q_6(t) q_{r.x.i}(t) \tag{4.12}$$

$$\tag{4.13}$$

The parasite lifecycle is divided into 8 age stages of equal lengths, reflecting the same divisions used in the drug model. The parasites mature into the next age class in continuous time. When progressing from age class 8 to age class 1 to renew the lifecycle, a number of daughter parasites are released, ready to infect more red blood cells. The model assumes instant invasion of new red blood cells by the daughter parasites (in reality, in non-anaemic patients, this re-invasion will take place within the order of a few seconds, and so neglecting this is a reasonable model simplification compared to the other time scales of interest). Where there are insufficient red blood cells to invade, this could be reflected in the model by reducing the number of daughter parasites produced, but all patients are known to be non-anaemic here.

It is necessary to implement a threshold at which parasites are considered to be fully killed, otherwise recrudesence will always be experienced due to the first-order kinetics.

Shadow compartments are used to ensure that parasites can only mature through one age class at a time. For purposes not connected to ageing, the number of parasites/drug in a particular state is the sum of

the two shadowing compartments for that state:

$$\begin{aligned} p_i &:= \textstyle\sum_{x=1}^{2} p_{x.i} \\ q_{\mathrm{ARS}.i} &:= \textstyle\sum_{x=1}^{2} q_{\mathrm{ARS}.x.i} \\ q_{\mathrm{DHA}.i} &:= \textstyle\sum_{x=1}^{2} q_{\mathrm{DHA}.x.i} \\ q_{r.i} &:= \textstyle\sum_{x=1}^{2} q_{r.x.i}. \end{aligned} \tag{4.14}$$

For stage $i$, the parasite death rate is

$$d_i(t) = \sum_{j \in \{\mathrm{ARS,DHA}\}} d_{j.i}(t) \tag{4.15}$$

where for convenience, the artemisinin derivatives are assumed to act independently, with

$$d_{j.i}(t) = \frac{E_{\max,j.i} c_{j.i}(t)}{\mathrm{EC}_{50,j.i} + c_{j.i}(t)}. \tag{4.16}$$

Here, $E_{\max,j.i}$ controls the speed of parasite death in stage $i$ caused by drug $j$, $c_{\mathrm{ARS}.i}$ is understood to mean $c_{7.i}$ and $c_{\mathrm{DHA}.i}$ to mean $c_{8.i}$. As this is non-linear with respect to the shadowing states, the shadowing states are assumed to have contributed proportionally to the effect:

$$d_{j.x.i} = -\kappa d_{j.i}(t) \frac{q_{j.x.i}(t)}{q_{j.1.i}(t) + q_{j.2.i}(t)}. \tag{4.17}$$

Parasite death is assumed to be caused when all the receptors belonging to that parasite have bound to the artemisinins (though this is not enforced by the model), so no further change needs to happen to the number of available parasite receptors upon parasite death.

Continuous-time parasite ageing is realised through use of discrete event toggles. It is assumed that the parasites within each age class are roughly synchronised. The parasites' lifecycle is known to have a duration of approximately 48 hours. Therefore, the ageing flows are toggled every 6 hours. After 8 such toggles, the parasites will have matured

through all 8 age classes and thus completed one full cycle. The timing and duration of these discrete event toggles are deterministic and are not influenced by other states of the model (though in reality it is hypothesised that the artemisinins can delay parasite growth). After each toggle, the shadow compartments and the lead compartments are swapped. In the following, $\overline{x}$ denotes the complementary lead/shadow state to $x$. The discrete events to toggle off the shadow flows are also deterministic but their times are not known in advance, they are triggered when the relevant state variables reaching zero.

While parasite ageing flows are activated, so too are the corresponding flows for drugs inside the parasites and the number of parasite receptors, with proportionate rates designed to ensure that the ageing process takes the same duration to complete for each component. Parasite ageing is described simply by linear zeroth-order flows during the activation window.

$$\frac{\mathrm{d}p_{i.x}}{\mathrm{d}t} = f_{\mathrm{shadow},p.i.x}(t) - d_{i.x}(t) \tag{4.18}$$

$$d_{i.x} = d_i(t)\frac{p_{i.x}(t)}{p_{i.1}(t) + p_{i.2}(t)}. \tag{4.19}$$

$\mathbb{1}_{\mathrm{S},x}$ is an indicator flag for the shadow flows for $x$ being activated. When the model first runs, both $\mathbb{1}_{\mathrm{S},1}$ and $\mathbb{1}_{\mathrm{S},2}$ are zero. Once $t_{\mathrm{lifecycle}}$ reaches 12 hours, $\mathbb{1}_{\mathrm{S},1}$ is set to 1. After the next 12 hours, $\mathbb{1}_{\mathrm{S},1}$ is set to 0 and $\mathbb{1}_{\mathrm{S},2}$ is set to 1. At this point, it should be the case that the states with $x = 1$ will be virtually empty (otherwise the shadow flow rates are set too small). This toggling of $\mathbb{1}_{\mathrm{S},1}$ and $\mathbb{1}_{\mathrm{S},2}$ then happens every 6 hours.

Most of the shadow flows have a similar structure, though may differ in behaviour at the renewal of the cycle. During the reproduction phase of the lifecycle, new parasite receptors are generated in proportion to the

number of new parasites generated.

$$
\begin{aligned}
f_{\text{shadow},p.x.i}(t) &= \begin{cases} \mathbb{1}_{\text{S},\overline{x}} M k_{\text{S},1,8} p_{\overline{x}.8}(t) - \mathbb{1}_{\text{S},x} k_{\text{S},i+1,i} p_{x.i}(t) & i = 1 \\ \mathbb{1}_{\text{S},\overline{x}} k_{\text{S},i,i-1} p_{\overline{x}.i-1}(t) - \mathbb{1}_{\text{S},x} k_{\text{S},i+1,i} p_{x.i}(t) & i \in \{2,\ldots,7\} \\ \mathbb{1}_{\text{S},\overline{x}} k_{\text{S},i,i-1} p_{\overline{x}.i-1}(t) - \mathbb{1}_{\text{S},x} k_{\text{S},1,8} p_{x.i}(t) & i = 8 \end{cases} \\
f_{\text{shadow},r.x.i}(t) &= \begin{cases} \mathbb{1}_{\text{S},x} \kappa M k_{\text{S},8,1} p_{\overline{x}.8}(t) & i = 1 \\ \mathbb{1}_{\text{S},\overline{x}} \kappa k_{\text{S},i,i-1} q_{r.\overline{x}.i-1}(t) - \mathbb{1}_{\text{S},x} \kappa k_{\text{S},i+1,i} q_{r.x.i}(t) & i \in \{2,\ldots,7\} \\ \mathbb{1}_{\text{S},\overline{x}} \kappa k_{\text{S},i,i-1} q_{r.\overline{x}.i-1}(t) - \mathbb{1}_{\text{S},x} \kappa k_{\text{S},1,8} q_{r.x.i}(t) & i = 8 \end{cases} \\
f_{\text{shadow},7.x.i}(t) &= \begin{cases} -\mathbb{1}_{\text{S},x} \kappa k_{\text{S},i+1,i} q_{7.x.i}(t) & i = 1 \\ \mathbb{1}_{\text{S},\overline{x}} \kappa k_{\text{S},i,i-1} q_{7.\overline{x}.i-1}(t) - \mathbb{1}_{\text{S},x} \kappa k_{\text{S},i+1,i} q_{7.x.i}(t) & i \in \{2,\ldots,7\} \\ \mathbb{1}_{\text{S},\overline{x}} \kappa k_{\text{S},i,i-1} q_{7.\overline{x}.i-1}(t) - \mathbb{1}_{\text{S},x} \kappa k_{\text{S},1,8} q_{7.x.i}(t) & i = 8 \end{cases} \\
\frac{\mathrm{d}q_{7.d}}{\mathrm{d}t} &= \sum_{i=1}^{8} \sum_{x=1}^{2} d_{i.x}(t) q_{7.x.i}(t) + \sum_{x=1}^{2} \mathbb{1}_{\text{S},x} \kappa k_{\text{S},1,8} q_{7.x.8}(t) \\
f_{\text{shadow},8.x.i}(t) &= \begin{cases} -\mathbb{1}_{\text{S},x} \kappa k_{\text{S},i+1,i} q_{8.x.i}(t) & i = 1 \\ \mathbb{1}_{\text{S},\overline{x}} \kappa k_{\text{S},i,i-1} q_{8.\overline{x}.i-1}(t) - \mathbb{1}_{\text{S},x} \kappa k_{\text{S},i+1,i} q_{8.x.i}(t) & i \in \{2,\ldots,7\} \\ \mathbb{1}_{\text{S},\overline{x}} \kappa k_{\text{S},i,i-1} q_{8.\overline{x}.i-1}(t) - \mathbb{1}_{\text{S},x} \kappa k_{\text{S},1,8} q_{8.x.i}(t) & i = 8 \end{cases} \\
\frac{\mathrm{d}q_{8.d}}{\mathrm{d}t} &= \sum_{i=1}^{8} \sum_{x=1}^{2} d_{i.x}(t) q_{8.x.i}(t) + \sum_{x=1}^{2} \mathbb{1}_{\text{S},x} \kappa k_{\text{S},1,8} q_{8.x.8}(t)
\end{aligned}
\tag{4.20}
$$

The shadowing rate parameters control the speed and synchronicity of parasite ageing, and in particular, replication. In the above, $M$ is the parasite replication factor, which is considered constant.

Also during this phase, any drug that was previously inside the parasite prior to replication is assumed to have no further effect on the daughter parasites (and is not released in active form into the bloodstream), and so moves into the sink compartment.

#### 4.2.2.6. Model parameters

For convenience, a table of all the parameters in the model is presented in Table 4.7, while Table 4.8 provides a table of all the state variables and

Table 4.9 lists all the intermediate functions used in the model equations.

Values for $IC_{50}$ concentrations for ARS and DHA are available from the literature [95] and may assist with identification and estimation of them from the available data, though these are plasma concentrations and not "within parasite" concentrations as used in this model.

| Parameter | Description |
|---|---|
| $k_{21}$ | ARS absorption rate |
| $k_{31}$ | ARS to DHA metabolism rate prior to absorption |
| $k_{42}$ | ARS to DHA metabolism rate after absorption |
| $k_{43}$ | DHA absorption rate |
| $k_{\mathrm{e},1}$ | ARS elimination rate before absorption |
| $k_{\mathrm{e}2}$ | ARS elimination rate from circulation |
| $k_{\mathrm{e}3}$ | DHA elimination rate before absorption |
| $k_{\mathrm{e}4}$ | DHA elimination rate from circulation |
| $k_{65}$ | ARS to DHA metabolism rate inside RBCs |
| $\eta_{j,i}$ | uptake rate of drug $j$ into a parasite in age class $i$ |
| $\mu_j$ | RBC permeability (diffusion rate) for drug $j$ |
| $k_{\mathrm{a},j,i}$ | parasite-drug binding rate for drug $j$ in age class $i$ |
| $\alpha_1$ | ARS quantity-concentration factor |
| $\alpha_2$ | DHA quantity-concentration factor |
| $k_{7.x.(j+1),7.y.j}$, $j \in \{1,2,3\}$ | shadow flow rate (toggled) |
| $k_{7.x.\mathrm{d},7.y.4}$, $j \in \{1,2,3\}$ | shadow flow rate (toggled) |
| $k_{8.x.(j+1),8.y.j}$, $j \in \{1,2,3\}$ | shadow flow rate (toggled) |
| $k_{8.x.\mathrm{d},8.y.4}$, $j \in \{1,2,3\}$ | shadow flow rate (toggled) |
| $k_{\mathrm{r}.x.(j+1),\mathrm{r}.y.j}$, $j \in \{1,2,3\}$ | shadow flow rate (toggled) |
| $k_{\mathrm{e,r}.y.j}$, $j \in \{1,2,3\}$ | shadow flow rate (toggled) |
| $E_{\max,i,j}$ | maximum rate of parasite death in stage $i$ from drug $j$ |

Table 4.7: Table of parameters for coupled model

| State | Description |
|---|---|
| $q_1$ | quantity of ARS not yet absorbed into circulation |
| $q_2$ | quantity of ARS in circulation |
| $q_3$ | quantity of DHA not yet absorbed into circulation |
| $q_4$ | quantity of DHA in circulation |
| $q_5$ | quantity of ARS inside red blood cells (unbound) |
| $q_6$ | quantity of DHA inside red blood cells (unbound) |
| $q_{7.x.i}$ | quantity of ARS bound to parasites in age stage $i$ and shadow state $x$ |
| $q_{8.x.i}$ | quantity of DHA bound to parasites in age stage $i$ and shadow state $x$ |
| $q_{7.d}$ | quantity of ARS that has denatured and is no longer associated to a living parasite |
| $q_{8.d}$ | quantity of DHA that has denatured and is no longer associated to a living parasite |
| $q_{r.x.i}$ | quantity of parasite receptors that is available for binding with the drugs, for parasites in age stage $i$ and shadow state $x$ |
| $p_{x.i}$ | quantity of living parasites that are in age state $i$ and shadow state $x$ |

Table 4.8: Table of state variables used in the coupled model

| Function | Description |
|---|---|
| $c_1$ | parent concentration prior to systemic circulation |
| $c_2$ | parent concentration in systemic circulation |
| $c_3$ | metabolite concentration prior to systemic circulation |
| $c_4$ | metabolite concentration in systemic circulation |
| $c_5$ | parent concentration in RBCs from passive diffusion |
| $c_6$ | metabolite concentration in RBCs from passive diffusion |
| $c_{7.i.j}$ | parent concentration in infected RBCs from facilitated uptake |
| $c_{8.i.j}$ | metabolite concentration in infected RBCs from facilitated uptake |
| $t \mapsto k_{52}(t)$ | parent drug passive duffusion into RBCs |
| $t \mapsto k_{25}(t)$ | parent drug passive duffusion from RBCs |
| $t \mapsto k_{64}(t)$ | metabolite passive duffision into RBCs |
| $t \mapsto k_{46}(t)$ | metabolite passive duffision from RBCs |
| $t \mapsto k_{7.i.j,5}(t)$ | parent drug uptake into infected RBCs in age stage $j$, shadow state $i$ |
| $t \mapsto k_{8.i.j,6}(t)$ | metabolite uptake into infected RBCs in age stage $j$, shadow state $i$ |
| $t \mapsto k_{7.\mathrm{d},7.i.j}(t)$ | parent drug flow when parasites die |
| $t \mapsto k_{8.\mathrm{d},8.i.j}(t)$ | metabolite drug flow when parasites die |
| $t \mapsto k_{\mathrm{e},\mathrm{r}.i.j}(t)$ | parasite receptor flow when parasites die |
| $y_1$ | parent drug observation |
| $y_2$ | metabolite observation |
| $p_i$ | total quantity of parasites in age class $i$ |
| $d_i$ | rate of parasite death for parasites in age class $i$ |

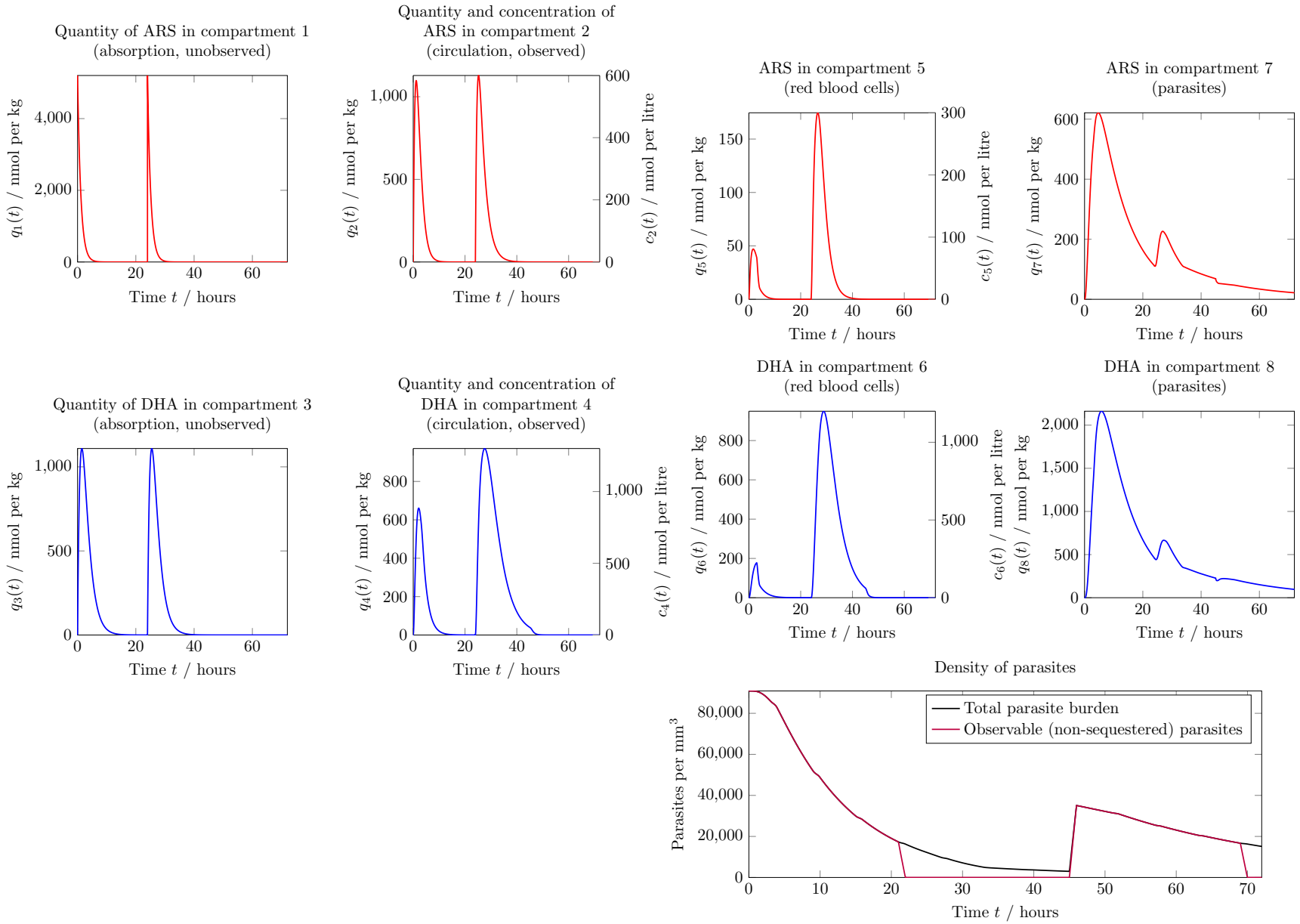
Table 4.9: Table of intermediate functions used in the coupled model

### 4.2.3. Model simulation

As part of the model development process, simulations were obtained to verify that the model has the intended behaviour, capturing the key features it was intended to, and to verify that it describes the data to the correct order of magnitude, without having any unanticipated or unrealistic dynamics.

Model parameters were chosen so that the simulated curves peak at roughly the same values as in Figure 4.1, and with an initial parasitaemia of a typical order of magnitude as in the MORU_ARC3_PD dataset. Naturally the model has many different parameters and input conditions that can be manipulated (and were manipulated by the author during the creation of this work), but a single simulation corresponding to one particular combination of parameters is shown here in Figure 4.23 to avoid confusing the reader or overloading this thesis with largely repetitive figures and descriptions. Some of these other possible model configurations/conditions are discussed in conjunction with an evaluation of the presented simulation. It should also be noted that for many of the features observed in the simulation, different parameters may substantially alter or eliminate those features: this is not a shortcoming of the model but instead demonstrates its flexibility.

Plots of drug quantities in Figure 4.23 are shown in the same layout as the respective compartments from Figure 4.22, except that only the totals of the shadowed compartments are shown (as the shadowing is only a modelling device and hinders rather than assists with a visualisation of the underlying system). For compartments where there is a sensible interpretation of drug concentration, these are also indicated on the same plots as the corresponding quantities, via use of dual $y$-axes: the left-hand $y$ axes represent quantities and the right-hand ones represent concentrations. The haematocrit was taken to be constant so there is in fact a constant volume for each of these compartments in this simulation.




▲ Figure 4.23: Illustrative full coupled parasite-drug model simulation, for identical doses of ARS administered at $t = 0$ and $t = 24$ hours

It is natural to take time $t = 0$ to be the timing of the first dose as this is usually also when the observations of the parasitaemia begin and the model would obviously not show any interesting dynamics before any drugs have been administered. In Figure 4.23, it can be seen that the pharmacodynamic effect of the drugs starts to kill the parasites almost immediately upon dosing of the artesunate at time 0, but does not suddenly stop when the plasma concentrations of artesunate and dihydroartemisinin reach low levels. Indeed, a significant proportion of the drugs enter into the parasites and remain there until either the parasites die or the corresponding infected red blood cells burst. This is seen to be more significant for DHA than it is for ARS, as expected due to the fast conversion of ARS to DHA. By quantity, in this example, over 50% of the ARS that reaches the systemic circulation is irreversibly taken up by parasitised red blood cells. For DHA, the quantity in the parasitised cells reaches more than double the peak quantity in the general circulation.

It is also readily seen that the model suggests that a single dose of ARS may not be sufficient for full pharmacodynamic effect; the peak ARS and DHA concentrations inside red blood cells are significantly lower following the first dose than the second. Of course, the parasitaemia is also lower by the time of the second dose, but the quantity of each drug inside the parasitised cells is saturable and still increases again following the second dose.

It is important to note that the volumes of distribution for ARS and DHA represented in this model (corresponding to the circulation compartments) are not comparable with volumes of distribution obtained for malaria patients according to a model without a red blood cell or parasitised cell uptake mechanism.

It is also apparent from the time course of observed DHA concentrations that the pharmacokinetics of at least DHA are significantly affected by the severity and stage of the malarial infection, because the second peak is much higher than the first. Hence in any multiple dosing study involving DHA, capturing this effect requires a model such as this, with coupling of the drug dynamics to the parasite burden. It is not reasonable to conclude that ARS does not require such a complicated model, because with other parameter values, similar behaviour is observed as for DHA.

In producing this simulation, a single synchronous group of parasites was specified, assumed to be 3 hours into their lifecycle at time $t = 0$. Hence, the parasites sequestered between $t = 21$ and 22 hours and replicated between $t = 45$ and 46 hours. Parameters were set such that the ageing and replication each took place over a period of one hour duration. These transitions are also apparent from the model simulations but are smooth gradual changes, not sharp sudden ones. Again the initial population of parasites and the ageing durations/speed can be specified differently in the model. For example, if the parasite population never appears to dip and regrow, this may be specified with either parasites in each of the age classes initially, with slower ageing transition times, or with a faster rate of parasite kill. At time $t = 45$ hours, there is a small but noticeable dip in the quantity of each drug that is present inside the parasitised cells where the drug from ruptured red blood cells is modelled as being effectively eliminated from the system, while simultaneously, the drugs are being taken up into the newly parasitised red blood cells as the parasites re-invade.

As has been reported for oral administration of ARS and noted in chapter 2, the parasite burden is reduced 10,000-fold within one parasite lifecycle period as seen in the plot of $p(t)$. Though somewhat difficult to observe from the figure due to the smooth transitions and confounding from other factors, the rate of parasite kill also varies with time due to

the stage-specific effects, which were set to have a susceptibility weight of 0.5 for the first 6 hours of the parasite lifecycle, 1 for the next 30 hours (until the parasites sequester) and then 0.1 for the remaining 12 hours for illustrative purposes.

As mentioned in subsection 2.3.3, this model was designed with studies in mind that included a third dose of ARS on day three, but the model behaviour following this third dose was essentially identical to that of the second, and so a 72 hour time window is shown excluding the third dose to better illustrate the model behaviour upon parasite replication once treatment has finished. It is therefore clear that if insufficient doses of ARS are administered, then the illness may recrudesce.

In summary, the model demonstrates all the expected and desired features as intended and has many parameters available for tuning the behaviour as applicable to each particular subject/patient.

# Chapter 5

# Closing

## 5.1. Conclusions and future work

It was not possible to meet all the aims and objectives set out in section 1.2, but the main aim was to develop a novel pharmacokinetic and pharmacodynamic model to capture important properties of artemisinin-based treatments, and indeed such a model has been proposed and shown to be viable.

New methods have been developed to help analyse this model, though not all these methods were able to successfully applied within the time constraints of the project, and some of the new methods designed were ultimately not applicable to the final version of the model structure and were therefore not presented here as it would have disrupted the focus on the overall aim.

The *in vitro* model has been shown to be structurally globally identifiable and describes the radio-labelled data reasonably well, although the parameters are not able to simultaneously describe all the *in vitro* data to the same level of accuracy, but the overall shape is generally captured.

Due to the discrete events within the model, it is not possible to analyse it from a structural identifiability perspective as a complete unit, but it is possible to consider the behaviour of the model up until the

first discrete event (though this clearly will not be able to inform on the identifiability of the discrete event times, which will require practical identifiability analyses). The author has applied structural identifiability methods to the model in this setting but these have so far proved intractable, probably due to the complexity of the model in terms of the number of parameters, number of state variables, and the nature of the non-linearities. It might be possible to proceed under the assumption that more of the parameters are known, or by deriving necessary and sufficient conditions for some of the parameters by analysing submodels of the full model. Further or alternatively, consideration of numerical and practical identifiability should be more straight-forward (but require forward simulations of the full model). Though these are also computationally demanding, which is why no results have been successfully obtained so far, numerical and practical identifiability have the advantage that sufficient computation time is really the only ingredient needed for such an analysis, whereas structural identifiability using algebraic methods is in some sense a "black box" because the symbolic algebra package does the calculations and the progress cannot be inferred, nor the intermediate results displayed or used. Eventually, often after a period of a few days, the symbolic algebra system will run out of memory (even if the problem is assigned to a highly specified and otherwise capable server) and the entire calculation will be terminated, leaving no usable evidence of any progress towards a solution.

The SMC method described in this work has been demonstrated on a small selection of models, including low dimensional models to explore and illustrate the typical behaviour in various situations, and also on the four compartment model introduced in previous work. Though not entirely straightforward, the method can be adjusted to run in parallel on modern computing clusters [85, 96, 97, 98], which could speed up the use of SMC considerably.

It is seen that the assumption of normally distributed errors is perhaps

not the most appropriate, since there are many outliers in the model fits which are heavily punished under this error model. A distribution with heavier tails such as Laplace or Student's $t$-distribution would be worth considering.

Finally, a simulation from the full coupled pharmacokinetic and pharmacodynamic drug-parasite model was presented and discussed, and it was demonstrated that the model met the overall aims (with the arguable exception of simplicity in an absolute sense, but necessary given the complexity of the real world system it attempts to represent). Implementation of the model in a more user-friendly package that the reader could interact with would potentially alleviate some of the negative effects of the model complexity. Simulations generated according to the full coupled model could be produced within a few seconds for most parameter combinations (but took up to 2 hours for others), though no attempt was yet made to optimise this. The time that would therefore be required to generate forward model evaluations for use in SMC likelihood calculations may therefore be concerning, but hopefully with optimisation and improved computational facilities that have been developed since the author started this work, would be well within the limits of present feasibility. This might permit eventual fulfilment of the remaining objective of using the model and methods described herein to perform inference for clinically-obtained data.

# Bibliography

[1] A.J. Hall. Pharmacokinetic modelling of derivatives of the anti-malarial drug artemisinin. Master's thesis. Coventry: University of Warwick, August 2011.

[2] Adam J. Hall, Michael J. Chappell, John A.D. Aston and Stephen A. Ward. Pharmacokinetic modelling of the anti-malarial drug artesunate and its active metabolite dihydroartemisinin. *Computer methods and programs in biomedicine*, 112(1):1–15, 2013. ISSN: 0169-2607. DOI: 10.1016/j.cmpb.2013.05.010.

[3] Adam J. Hall, Michael J. Chappell, John A.D. Aston and Stephen A. Ward. Reprint of "pharmacokinetic modelling of the anti-malarial drug artesunate and its active metabolite dihydroartemisinin". *Computer methods and programs in biomedicine*, 114(3):e14–e28, 2014. ISSN: 0169-2607. DOI: 10.1016/j.cmpb.2013.12.001.

[4] World Health Organisation. Malaria fact sheet. Website (Retrieved 29-Sept-2015). 2015. URL: http://www.who.int/mediacentre/factsheets/fs094/en/index.html.

[5] Centers for Disease Control and Prevention. Impact of malaria. Website (Retrieved 29-Sept-2015). 2015. URL: http://www.cdc.gov/malaria/malaria_worldwide/impact.html.

[6] John Luke Gallup and Jeffrey D. Sachs. The economic burden of malaria. *The american journal of tropical medicine and hygiene*, 64(1 suppl):85–96, 2001.

[7] World Health Organisation. 10 facts on malaria. Website (Retrieved 29-Sept-2015). 2015. URL: http://www.who.int/features/factfiles/malaria/en/.

[8] RTSS Clinical Trials Partnership. Efficacy and safety of rts, s/as01 malaria vaccine with or without a booster dose in infants and children in africa: final results of a phase 3, individually randomised, controlled trial. *The lancet*, 386(9988):31–45, 2015. ISSN: 0140-6736. DOI: 10.1016/S0140-6736(15)60721-8.

[9] World Health Organisation. *Guidelines for the treatment of malaria*. World Health Organization, 3rd edition, 2015. ISBN: 9789241549127.

[10] Aidan J. O'Donnell, Petra Schneider, Harriet G. McWatters and Sarah E. Reece. Fitness costs of disrupting circadian rhythms in malaria parasites. *Proceedings of the royal society of london b: biological sciences*, 278(1717):2429–2436, 2011. ISSN: 0962-8452. DOI: 10.1098/rspb.2010.2457.

[11] Jordi Ferrer, Jaume Vidal, Clara Prats, Joaquim Valls, Esperanza Herreros, Daniel Lopez, Antoni Giró and Domingo Gargallo. Individual-based model and simulation of plasmodium falciparum infected erythrocyte in vitro cultures. *Journal of theoretical biology*, 248(3):448–459, 2007. ISSN: 0022-5193. DOI: 10.1016/j.jtbi.2007.05.030.

[12] Renu Tuteja. Malaria - an overview. *Febs journal*, 274(18):4670–4679, 2007. ISSN: 1742-4658. DOI: 10.1111/j.1742-4658.2007.05997.x.

[13] F. Ellis McKenzie and William H. Bossert. An integrated model of plasmodium falciparum dynamics. *Journal of theoretical biology*, 232(3):411–426, 2005. ISSN: 0022-5193. DOI: 10.1016/j.jtbi.2004.08.021.

[14] Hernando A. del Portillo, Mireia Ferrer, Thibaut Brugat, Lorena Martin-Jaular, Jean Langhorne and Marcus V. G. Lacerda. The role of the spleen in malaria. *Cellular microbiology*, 14(3):343–355, 2012. ISSN: 1462-5822. DOI: 10.1111/j.1462-5822.2011.01741.x.

[15] D. Kwiatkowski and B.M. Greenwood. Why is malaria fever periodic? a hypothesis. *Parasitology today*, 5(8):264–266, 1989. ISSN: 0169-4758. DOI: 10.1016/0169-4758(89)90261-5.

[16] Charles R.J.C. Newton, Tran Tinh Hien and Nicholas White. Cerebral malaria. *Journal of neurology, neurosurgery & psychiatry*, 69(4):433–441, 2000. DOI: 10.1136/jnnp.69.4.433.

[17] Louis H. Miller, Michael F. Good and Genevieve Milon. Malaria pathogenesis. *Science*, 264(5167):1878–1883, 1994. DOI: 10.1126/science.8009217.

[18] Jamie T. Griffin, T. Déirdre Hollingsworth, Hugh Reyburn, Chris J. Drakeley, Eleanor M. Riley and Azra C. Ghani. Gradual acquisition of immunity to severe malaria with increasing exposure. *Proceedings of the royal society of london b: biological sciences*, 282(1801):20142657, 2015. ISSN: 0962-8452. DOI: 10.1098/rspb.2014.2657.

[19] M.P. Grobusch and P.G. Kremsner. Uncomplicated malaria. In R.W. Compans, M.D. Cooper, T. Honjo, H. Koprowski, F. Melchers, M.B.A. Oldstone, S. Olsnes, M. Potter, P.K. Vogt, H. Wagner, DavidJ. Sullivan and Sanjeev Krishna, editors, *Malaria: drugs, disease and post-genomic biology*. Volume 295, in Current Topics in Microbiology and Immunology, pages 81–104. Springer, 2005. ISBN: 978-3-540-25363-1. DOI: 10.1007/3-540-29088-5_4.

[20] Bee San Tan. Population pharmacokinetics of artesunate and its active metabolite dihydroartemisinin. PhD thesis. University of Iowa, 2009. URL: http://ir.uiowa.edu/etd/442/.

[21] P. Newton, Y. Suputtamongkol, P. Teja-Isavadharm, S. Pukrittayakamee, V. Navaratnam, I. Bates and N. White. Antimalarial bioavailability and disposition of artesunate in acute falciparum malaria. *Antimicrobial agents and chemotherapy*, 44(4):972–977, 2000. DOI: 10.1128/AAC.44.4.972-977.2000.

[22] World Health Organisation. Overview of malaria treatment. Website (Retrieved 29-Sept-2015). 2015. URL: http://www.who.int/malaria/areas/treatment/overview/en/.

[23] Rozanne Harmse, Ho Ning Wong, Frans Smit, Richard K. Haynes and David D. N'Da. The case for development of 11-aza-artemisinins for malaria. *Current medicinal chemistry*, 22(31):3607–3630, 2015. ISSN: 0929-8673/1875-533X. DOI: 10.2174/0929867322666150729115752.

[24] Falgun Shah, Shuang-Qing Zhang, Shilpa Prakash Kandhari, Prasenjit Mukherjee, Amar Chittiboyina, Mitchell A Avery and Bonnie A Avery. In vitro erythrocytic uptake studies of artemisinin and selected derivatives using lc–ms and 2d-qsar analysis of uptake in parasitized erythrocytes. *Bioorganic & medicinal chemistry*, 17(14):5325–5331, 2009.

[25] N.J. White. Preventing antimalarial drug resistance through combinations. *Drug resistance updates*, 1(1):3–9, 1998. ISSN: 1368-7646. DOI: 10.1016/S1368-7646(98)80208-2.

[26] Debashish Das, Rupam Tripura, Aung Pyae Phyo, Khin Maung Lwin, Joel Tarning, Sue J. Lee, Warunee Hanpithakpong, Kasia Stepniewska, Didier Menard, Pascal Ringwald, Kamolrat Silamut, Mallika Imwong, Kesinee Chotivanich, Poravuth Yi, Nicholas P. J. Day,

Niklas Lindegardh, Duong Socheat, Chea Nguon, Nicholas J. White, François Nosten and Arjen M. Dondorp. Effect of high-dose or split-dose artesunate on parasite clearance in artemisinin-resistant falciparum malaria. *Clinical infectious diseases*, 56(5):e48–e58, 2013. DOI: 10.1093/cid/cis958.

[27] Nectarios Klonis, Maria P. Crespo-Ortiz, Iveta Bottova, Nurhidanatasha Abu-Bakar, Shannon Kenny, Philip J. Rosenthal and Leann Tilley. Artemisinin activity against Plasmodium falciparum requires hemoglobin uptake and digestion. *Proceedings of the national academy of sciences*, 108(28):11405–11410, 2011. DOI: 10.1073/pnas.1104063108.

[28] Paul M. O'Neill, Victoria E. Barton and Stephen A. Ward. The molecular mechanism of action of artemisinin — the debate continues. *Molecules*, 15(3):1705–1721, 2010. ISSN: 1420-3049. DOI: 10.3390/molecules15031705.

[29] Thomas Antoine, Nicholas Fisher, Richard Amewu, Paul M. O'Neill, Stephen A. Ward and Giancarlo A. Biagini. Rapid kill of malaria parasites by artemisinin and semi-synthetic endoperoxides involves ros-dependent depolarization of the membrane potential. *Journal of antimicrobial chemotherapy*, 69(4):1005–1016, 2014. DOI: 10.1093/jac/dkt486.

[30] Steven R. Meshnick. The mode of action of antimalarial endoperoxides. *Transactions of the royal society of tropical medicine and hygiene*, 88(Supplement 1):31–32, 1994. ISSN: 0035-9203. DOI: 10.1016/0035-9203(94)90468-5.

[31] Nectarios Klonis, Darren J. Creek and Leann Tilley. Iron and heme metabolism in plasmodium falciparum and the mechanism of action of artemisinins. *Current opinion in microbiology*, 16(6):722–727, 2013. ISSN: 1369-5274. DOI: 10.1016/j.mib.2013.07.005.

[32] P. Olliaro, S. Ramanathan, M. Vaillant, S.E. Reuter, A.M. Evans, S. Krudsood, S. Looareesuwan, J.R. Kiechel, W.R.J. Taylor and V. Navaratnam. Pharmacokinetics and comparative bioavailability of artesunate and mefloquine administered separately or as a fixed combination product to healthy volunteers and patients with uncomplicated plasmodium falciparum. *Journal of bioequivalence & bioavailability*, 2(3):059–066, 2010. ISSN: 0975-0851. DOI: 10.4172/jbb.1000032. URL: http://omicsonline.org/ArchiveJBB/2010/May/03/JBB-02-059.php.

[33] Timothy M.E. Davis, Hoang Lan Phuong, Kenneth F. Ilett, Nguyen Canh Hung, Kevin T. Batty, Vu Duong Bich Phuong, Shane M. Powell, Huynh Van Thien and Tran Quang Binh. Pharmacokinetics and pharmacodynamics of intravenous artesunate in severe falciparum malaria. *Antimicrobial agents and chemotherapy*, 45(1):181–186, 2001. DOI: 10.1128/AAC.45.1.181-186.2001.

[34] J.A. Simpson, T. Agbenyega, K.I. Barnes, G. Di Perri, P. Folb, M. Gomes, S. Krishna, S. Krudsood, S. Looareesuwan, S. Mansor et al. Population pharmacokinetics of artesunate and dihydroartemisinin following intra-rectal dosing of artesunate in malaria patients. *Plos medicine*, 3(11):e444, 2006. DOI: 10.1371/journal.pmed.0030444.

[35] P. Teja-Isavadharm, G. Watt, C. Eamsila, K. Jongsakul, Q. Li, G. Keeratithakul, N. Sirisopana, L. Luesutthiviboon, T.G. Brewer and D.E. Kyle. Comparative pharmacokinetics and effect kinetics of orally administered artesunate in healthy volunteers and patients with uncomplicated falciparum malaria. *The american journal of tropical medicine and hygiene*, 65(6):717, 2001.

[36] T.S. Skinner, L.S. Manning, W.A. Johnston and T.M.E. Davis. In vitro stage-specific sensitivity of Plasmodium falciparum to quinine and artemisinin drugs. *International journal for parasitology*, 26(5):519–525, 1996. ISSN: 0020-7519. DOI: 10.1016/0020-7519(96)89380-5.

[37] Sompob Saralamba, Wirichada Pan-Ngum, Richard J. Maude, Sue J. Lee, Joel Tarning, Niklas Lindegårdh, Kesinee Chotivanich, F. Nosten, Nicholas P. J. Day, Duong Socheat, Nicholas J. White, Arjen M. Dondorp and Lisa J. White. Intrahost modeling of artemisinin resistance in Plasmodium falciparum. *Proceedings of the national academy of sciences*, 108(1):397–402, 2011. DOI: 10.1073/pnas.1006113108.

[38] Katherine Kay and Ian M. Hastings. Improving pharmacokinetic-pharmacodynamic modeling to investigate anti-infective chemotherapy with application to the current generation of antimalarial drugs. *Plos comput biol*, 9(7):e1003151, July 2013. DOI: 10.1371/journal.pcbi.1003151.

[39] C.A. Morris, S. Duparc, I. Borghini-Fuhrer, D. Jung, C.S. Shin and L. Fleckenstein. Review of the clinical pharmacokinetics of ARS and its active metabolite DHA following intravenous, intramuscular, oral or rectal administration. *Malaria journal*, 10(1):263, 2011. ISSN: 1475-2875. DOI: 10.1186/1475-2875-10-263.

[40] Herman J. Woerdenbag, Niesko Pras, Wim van Uden, T. Elco Wallaart, Aäron C. Beekman and Charles B. Lugt. Progress in the research of artemisinin-related antimalarials: an update. *Pharmacy world & science*, 16(4):169–180, 1994. ISSN: 0928-1231. DOI: 10.1007/BF01872865.

[41] X.Q. Li, A. Björkman, T.B. Andersson, L.L. Gustafsson and C.M. Masimirembwa. Identification of human cytochrome P450s

that metabolise anti-parasitic drugs and predictions of in vivo drug hepatic clearance from in vitro data. *European journal of clinical pharmacology*, 59(5):429–442, 2003. ISSN: 0031-6970. DOI: 10.1007/s00228-003-0636-9.

[42] N. Lindegårdh, A.M. Dondorp, P. Singhasivanon, N.J. White and N.P.J. Day. Validation and application of a liquid chromatographic-mass spectrometric method for determination of artesunate in pharmaceutical samples. *Journal of pharmaceutical and biomedical analysis*, 45(1):149–153, 2007. ISSN: 0731-7085. DOI: 10.1016/j.jpba.2007.04.030.

[43] Qigui Li and Peter Weina. Artesunate: the best drug in the treatment of severe and complicated malaria. *Pharmaceuticals*, 3(7):2322–2332, 2010. ISSN: 1424-8247. DOI: 10.3390/ph3072322.

[44] K.T. Batty, T.M.E. Davis, L.T.A. Thu, T. Quang Binh, T. Kim Anh and K.F. Ilett. Selective high-performance liquid chromatographic determination of artesunate and $\alpha$-and $\beta$-dihydroartemisinin in patients with falciparum malaria. *Journal of chromatography b: biomedical sciences and applications*, 677(2):345–350, 1996. ISSN: 0378-4347. DOI: 10.1016/0378-4347(95)00428-9.

[45] Kenneth F. Ilett, Brian T. Ethell, James L. Maggs, Timothy M. E. Davis, Kevin T. Batty, Brian Burchell, Tran Quang Binh, Le Thi Anh Thu, Nguyen Canh Hung, Munir Pirmohamed, B. Kevin Park and Geoffrey Edwards. Glucuronidation of dihydroartemisinin in vivo and by human liver microsomes and expressed udp-glucuronosyltransferases. *Drug metabolism and disposition*, 30(9):1005–1012, 2002. DOI: 10.1124/dmd.30.9.1005.

[46] E.M. Scholar and W.B. Pratt. *The antimicrobial drugs*. Oxford University Press, 2000. ISBN: 9780195125283.

[47] Ulrika S. Svensson, Mohamed Alin, Mats O. Karlsson, Yngve Bergqvist and Michael Ashton. Population pharmacokinetic and pharmacodynamic modelling of artemisinin and mefloquine enantiomers in patients with falciparum malaria. *European journal of clinical pharmacology*, 58(5):339–351, 2002. ISSN: 0031-6970. DOI: 10.1007/s00228-002-0485-y.

[48] Nehal Vyas, Bonnie A. Avery, Mitchell A. Avery and Christy M. Wyandt. Carrier-mediated partitioning of artemisinin into Plasmodium falciparum-infected erythrocytes. *Antimicrobial agents and chemotherapy*, 46(1):105–109, 2002.

[49] W. Hanpithakpong, B. Kamanikom, A.M. andrp, P. Singhasivanon, N.J. White, N.P.J. Day and N. Lindegardh. A liquid chromatographic-tandem mass spectrometric method for determination of artesunate and its metabolite dihydroartemisinin in human plasma. *Journal of chromatography b*, 876(1):61–68, 2008. ISSN: 1570-0232. DOI: 10.1016/j.jchromb.2008.10.018.

[50] M.A. Van Agtmael, C.A.A. Van der Graaf, T.K. Dien, R.P. Koopmans and C.J. Van Boxtel. The contribution of the enzymes CYP2D6 and CYP2C19 in the demethylation of artemether in healthy subjects. *European journal of drug metabolism and pharmacokinetics*, 23(3):429–436, 1998. ISSN: 0378-7966. DOI: 10.1007/BF03192305.

[51] Arjen M. Dondorp, François Nosten, Poravuth Yi, Debashish Das, Aung Phae Phyo, Joel Tarning, Khin Maung Lwin, Frederic Ariey, Warunee Hanpithakpong, Sue J. Lee, Pascal Ringwald, Kamolrat Silamut, Mallika Imwong, Kesinee Chotivanich, Pharath Lim, Trent Herdman, Sen Sam An, Shunmay Yeung, Pratap Singhasivanon, Nicholas P.J. Day, Niklas Lindegardh, Duong Socheat and Nicholas J. White. Artemisinin resistance in Plasmodium falciparum malaria. *New*

*england journal of medicine*, 361(5):455–467, 2009. DOI: 10.1056/NEJMoa0808859.

[52] Patricia Schlagenhauf. Mefloquine for malaria chemoprophylaxis 1992–1998: a review. *Journal of travel medicine*, 6(2):122–133, 1999. ISSN: 1708-8305. DOI: 10.1111/j.1708-8305.1999.tb00843.x.

[53] John A. Jacquez. *Compartmental analysis in biology and medicine*. BioMedware, Ann Arbor, MI, third edition, 1996.

[54] Keith R. Godfrey. *Compartmental models and their application*. Academic Press, 1983.

[55] P. Doucet and P.B. Sloep. *Mathematical modeling in the life sciences*. E. Horwood, 1992.

[56] Gordon L. Brownell, Mones Berman and James S. Robertson. Nomenclature for tracer kinetics. *The international journal of applied radiation and isotopes*, 19(3):249–262, 1968. ISSN: 0020-708X. DOI: 10.1016/0020-708X(68)90022-7.

[57] É. Walter and L. Pronzato. *Identification of parametric models from experimental data*. Of *Communications and control engineering*. Springer, 1997. ISBN: 9783540761198.

[58] Keith R. Godfrey and William R. Fitch. The deterministic identifiability of nonlinear pharmacokinetic models. *Journal of pharmacokinetics and pharmacodynamics*, 12(2):177–191, 1984. ISSN: 1567-567X. DOI: 10.1007/BF01059277.

[59] D.J. Cole, B.J.T. Morgan and D.M. Titterington. Determining the parametric structure of models. *Mathematical biosciences*, 228(1):16–30, 2010. ISSN: 0025-5564. DOI: 10.1016/j.mbs.2010.08.004.

[60] Maria Pia Saccomani, Stefania Audoly and Leontina D'Angiò. Parameter identifiability of nonlinear systems: the role of initial conditions. *Automatica*, 39(4):619–632, 2003.

[61] Sándor Vajda. Structural equivalence of linear systems and compartmental models. *Mathematical biosciences*, 55(1):39–64, 1981. ISSN: 0025-5564. DOI: 10.1016/0025-5564(81)90012-2.

[62] Gabriella Margaria, Eva Riccomagno, Michael J. Chappell and Henry P. Wynn. Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences. *Mathematical biosciences*, 174(1):1–26, 2001. ISSN: 0025-5564. DOI: 10.1016/S0025-5564(01)00079-7.

[63] Mats Jirstrand. Algebraic methods for modelling and design in control. PhD thesis. Linköping University, Department of Electrical Engineering, Division of Automatic Control, 1996. ISBN: 91-7871-676-4.

[64] Nicolette Meshkat, Chris Anderson and Joseph J. DiStefano III. Alternative to ritt's pseudodivision for finding the input-output equations of multi-output models. *Mathematical biosciences*, 239(1):117–123, 2012. ISSN: 0025-5564. DOI: 10.1016/j.mbs.2012.04.008.

[65] N.D. Evans, H.A.J. Moyse, D. Lowe, D. Briggs, R. Higgins, D. Mitchell, Daniel Zehnder and Michael J. Chappell. Structural identifiability of surface binding reactions involving heterogeneous analyte: application to surface plasmon resonance experiments. *Automatica*, 49(1):48–57, 2013. ISSN: 0005-1098. DOI: 10.1016/j.automatica.2012.09.015.

[66] Sandor Vajda, Keith R. Godfrey and Herschel Rabitz. Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Mathematical biosciences*, 93(2):217–248, 1989. ISSN: 0025-5564. DOI: 10.1016/0025-5564(89)90024-2.

[67] H. Pohjanpalo. System identifiability based on the power series expansion of the solution. *Mathematical biosciences*, 41(1-2):21–33, 1978. ISSN: 0025-5564. DOI: 10.1016/0025-5564(78)90063-9.

[68] Hannu Pohjanpalo. Identifiability of deterministic differential models in state space. Technical report. Research Report. Technical Research Centre of Finland, 1982.

[69] Johan Karlsson, Milena Anguelova and Mats Jirstrand. An efficient method for structural identifiability analysis of large dynamic systems. In *16th ifac symposium on system identification*. Volume 16. (1). International Federation of Automatic Control, 2012, pages 941–946. ISBN: 978-3-902823-06-9.

[70] Maria Pia Saccomani. Structural vs practical identifiability in system biology. In *International work-conference on bioinformatics and biomedical engineering, IWBBIO 2013, granada, spain, march 18-20, 2013. proceedings*. Copicentro Editorial, 2013, pages 305–313. ISBN: 978-84-15814-13-9. URL: http://iwbbio.ugr.es/papers/iwbbio_055.pdf.

[71] Andreas Raue, Clemens Kreutz, Tomas Maiwald, Julie Bachmann, Marcel Schilling, Ursula Klingmüller and Jeans Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, 2009. DOI: 10.1093/bioinformatics/btp358.

[72] T. Gordi, R. Xie, N.V. Huong, D.X. Huong, M.O. Karlsson and M. Ashton. A semi-physiological pharmacokinetic model for artemisinin in healthy subjects incorporating autoinduction of metabolism and saturable first-pass hepatic extraction. *British journal of clinical pharmacology*, 59(2):189–198, 2005. ISSN: 1365-2125. DOI: 10.1111/j.1365-2125.2004.02321.x.

[73] Beesan Tan, Himanshu Naik, In-Jin Jang, Kyung-Sang Yu, Lee Kirsch, Chang-Sik Shin, J. Craft and Lawrence Fleckenstein. Population pharmacokinetics of artesunate and dihydroartemisinin following single- and multiple-dosing of oral artesunate in healthy subjects. *Malaria journal*, 8(1):304, 2009. ISSN: 1475-2875. DOI: 10.1186/1475-2875-8-304.

[74] Mike B. Gravenor, Alun L. Lloyd, Peter G. Kremsner, Michel A. Missinou, Mike English, Kevin Marsh and Dominic Kwiatkowski. A model for estimating total parasite load in falciparum malaria patients. *Journal of theoretical biology*, 217(2):137–148, 2002. ISSN: 0022-5193. DOI: 10.1006/jtbi.2002.3030.

[75] Julie A. Simpson, Sophie Zaloumis, Alysha M. DeLivera, Ric N. Price and James M. McCaw. Making the most of clinical data: reviewing the role of pharmacokinetic-pharmacodynamic models of anti-malarial drugs. *The aaps journal*, 16(5):962–974, 2014. DOI: 10.1208/s12248-014-9647-y.

[76] G.A.F. Seber and C.J. Wild. *Nonlinear regression*. Volume 503. Wiley, 1989.

[77] Pierre Del Moral, Arnaud Doucet and Ajay Jasra. Sequential monte carlo samplers. *Journal of the royal statistical society: series b (statistical methodology)*, 68(3):411–436, 2006. ISSN: 1467-9868. DOI: 10.1111/j.1467-9868.2006.00553.x.

[78] Yan Zhou, Adam M. Johansen and John A.D. Aston. Towards automatic model comparison: an adaptive sequential monte carlo approach. *Journal of computational and graphical statistics*:to appear, 2015. ISSN: 1061-8600. DOI: 10.1080/10618600.2015.1060885.

[79] Christopher F.H. Nam, John A.D. Aston and Adam M. Johansen. Parallel sequential monte carlo samplers and estimation of the number of states in a hidden markov model. *Annals of the institute of statistical mathematics*, 66(3):553–575, 2014. ISSN: 0020-3157. DOI: 10.1007/s10463-014-0450-4.

[80] Thi Le Thu Nguyen, Francois Septier, Gareth W. Peters and Yves Delignon. Efficient sequential monte-carlo samplers for bayesian inference. ArXiv e-print 1504.05753. April 2015.

[81] John MacCormick and Michael Isard. Partitioned sampling, articulated objects, and interface-quality hand tracking. In David Vernon, editor, *Computer vision - eccv 2000*. Volume 1843, in Lecture Notes in Computer Science, pages 3–19. Springer, 2000. ISBN: 978-3-540-67686-7. DOI: 10.1007/3-540-45053-X_1.

[82] J.S. Liu. *Monte carlo strategies in scientific computing*. Of *Springer Series in Statistics*. Springer, 2008. ISBN: 9780387763699.

[83] Randal Douc and Olivier Cappé. Comparison of resampling schemes for particle filtering. In *Image and signal processing and analysis, 2005. ispa 2005. proceedings of the 4th international symposium on*. IEEE. IEEE, September 2005, pages 64–69. DOI: 10.1109/ISPA.2005.195385.

[84] Adam M. Johansen. SMCTC: Sequential Monte Carlo in C++. *Journal of statistical software*, 30(6):1–41, April 2009. ISSN: 1548-7660. URL: http://www.jstatsoft.org/v30/i06.

[85] Yan Zhou. vSMC: Parallel Sequential Monte Carlo in C++. *Journal of statistical software*, 62(9):1–49, 2014. ISSN: 1548-7660. URL: http://www.jstatsoft.org/v62/i09.

[86] Lawrence M. Murray. Bayesian state-space modelling on high-performance hardware using LibBi. ArXiv e-print 1306.3277. June 2013.

[87] Adrien Todeschini, François Caron, Marc Fuentes, Pierrick Legrand and Pierre Del Moral. Biips: software for Bayesian inference with interacting particle systems. ArXiv e-print 1412.3779. December 2014.

[88] Gareth O. Roberts and Jeffrey S. Rosenthal. Examples of adaptive mcmc. *Journal of computational and graphical statistics*, 18(2):349–367, 2009.

[89] James Carpenter, Peter Clifford and Paul Fearnhead. Improved particle filter for nonlinear problems. *IEE Proceedings - Radar, Sonar and Navigation*, 146(1):2–7, 1999.

[90] Yuguo Chen. Another look at rejection sampling through importance sampling. *Statistics & probability letters*, 72(4):277–283, 2005. ISSN: 0167-7152. DOI: 10.1016/j.spl.2005.01.002.

[91] D.J. Sheskin. *Handbook of parametric and nonparametric statistical procedures: third edition*. CRC Press, 2003. ISBN: 9781420036268.

[92] Ana Justel, Daniel Peña and Rubén Zamar. A multivariate kolmogorov-smirnov test of goodness of fit. *Statistics & probability letters*, 35(3):251–259, 1997. ISSN: 0167-7152. DOI: 10.1016/S0167-7152(97)00020-5.

[93] Jeremy N. Burrows and David Waterson. Discovering new medicines to control and eradicate malaria. In Richard Elliott, editor, *Third world diseases*, pages 125–180. Springer, 2011. ISBN: 978-3-642-23487-3. DOI: 10.1007/7355_2011_14.

[94] James H. Matis and Thomas E. Wehrly. Generalized stochastic compartmental models with erlang transit times. *Journal of pharmacokinetics and biopharmaceutics*, 18(6):589–607, 1990. ISSN: 0090-466X. DOI: 10.1007/BF01073940.

[95] C. Wongsrichanalai, T. Wimonwattrawatee, P. Sookto, A. Laoboonchai, D.G. Heppner, D.E. Kyle and W.H. Wernsdorfer. In vitro sensitivity, of plasmodium falciparum to artesunate in thailand. *Bulletin of the world health organization*, 77(5):392–398, 1999. ISSN: 0042-9686.

[96] Garland Durham and John Geweke. Massively parallel sequential monte carlo for bayesian inference. 2011.

[97] Anthony Lee, Christopher Yau, Michael B. Giles, Arnaud Doucet and Christopher C. Holmes. On the utility of graphics cards to perform massively parallel simulation of advanced monte carlo methods. *Journal of computational and graphical statistics*, 19(4):769–789, 2010. DOI: 10.1198/jcgs.2010.10039.

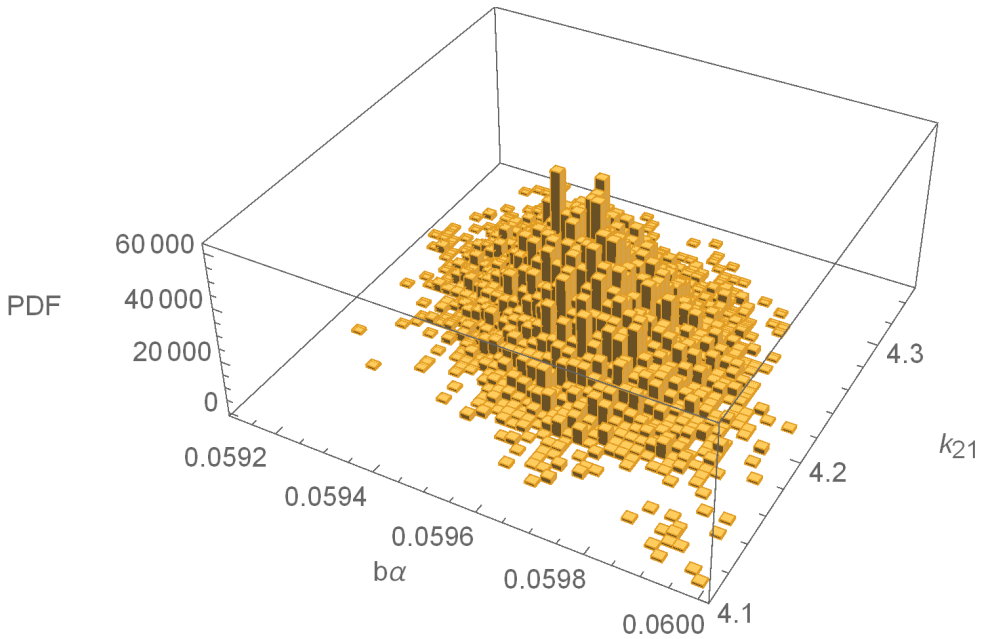
[98] Brooks Paige, Frank Wood, Arnaud Doucet and Yee Whye Teh. Asynchronous anytime sequential monte carlo. In *Advances in neural information processing systems 27*. Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence and K. Q. Weinberger, editors. Curran Associates, Inc., 2014, pages 3410–3418.

# Appendices

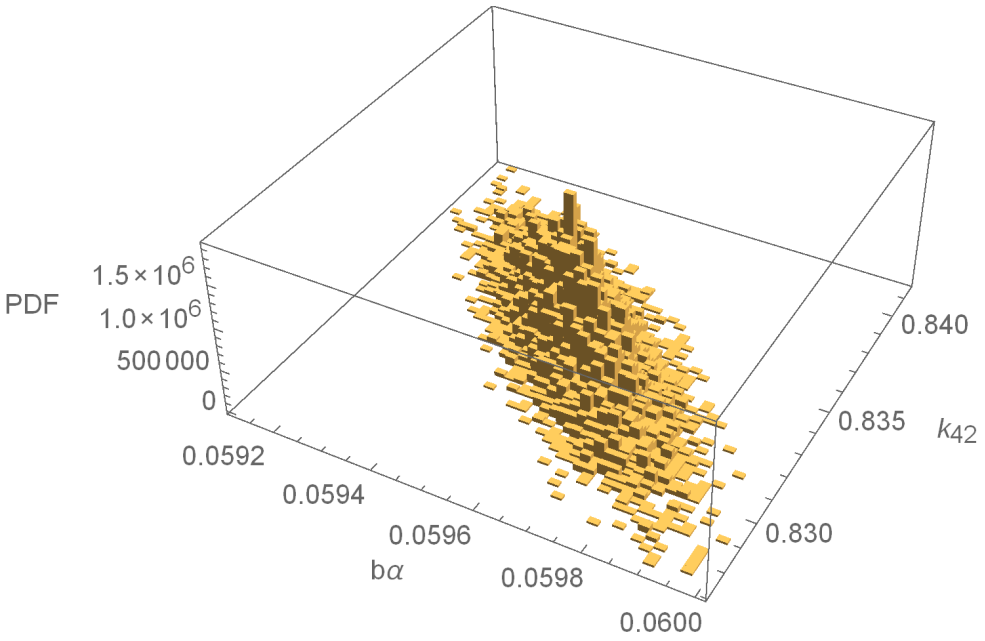
# Appendix A

# Plots

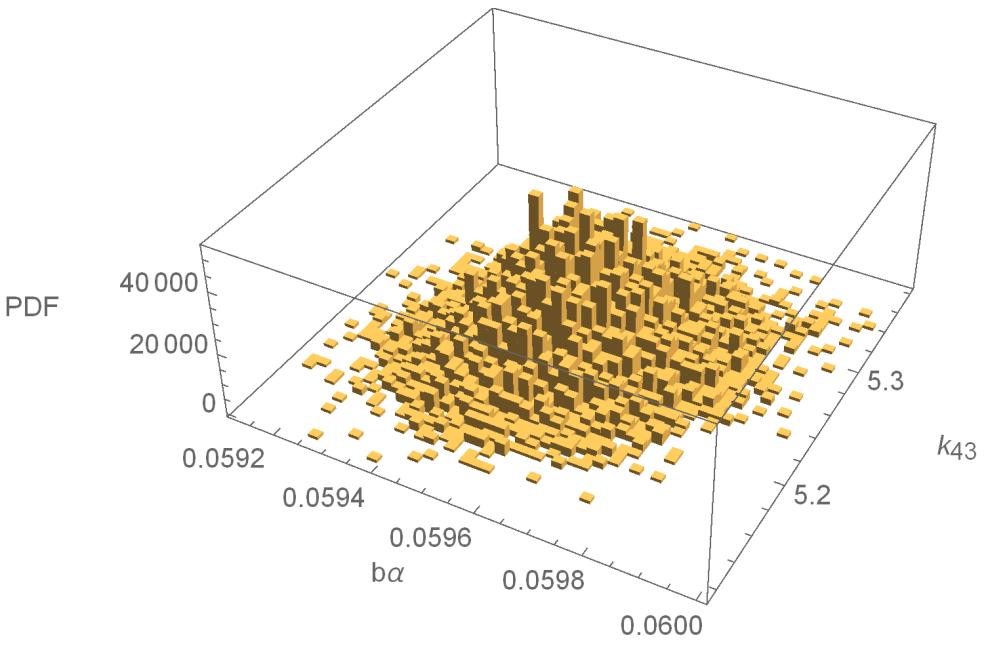
## A.1. Four compartment model pairwise posterior marginal distributions



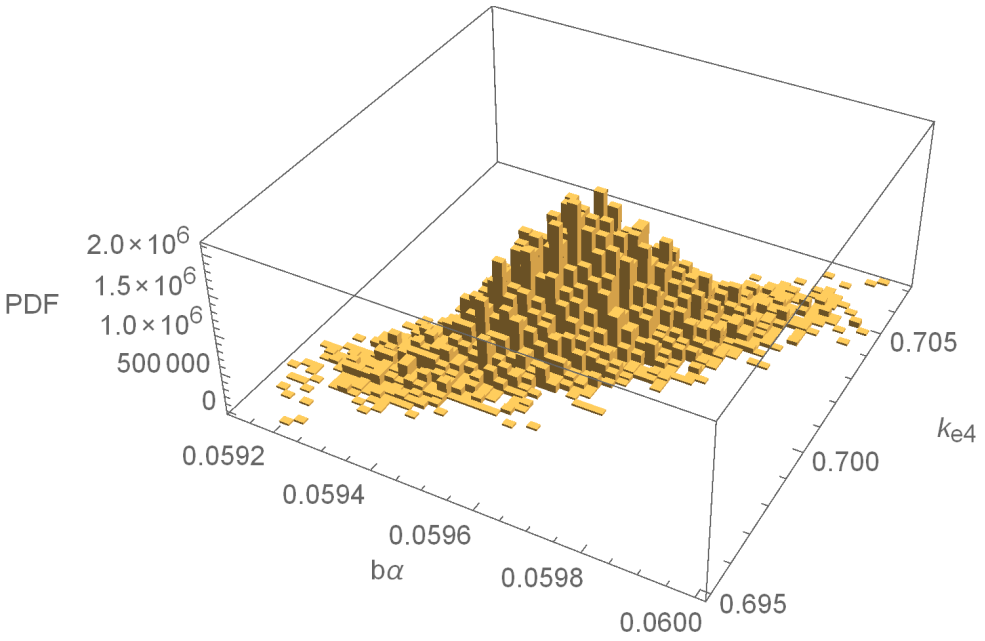

▲ Figure A.1: SMC estimate of posterior pairwise marginal distribution for $b\alpha$ and $k_{21}$ for patient A



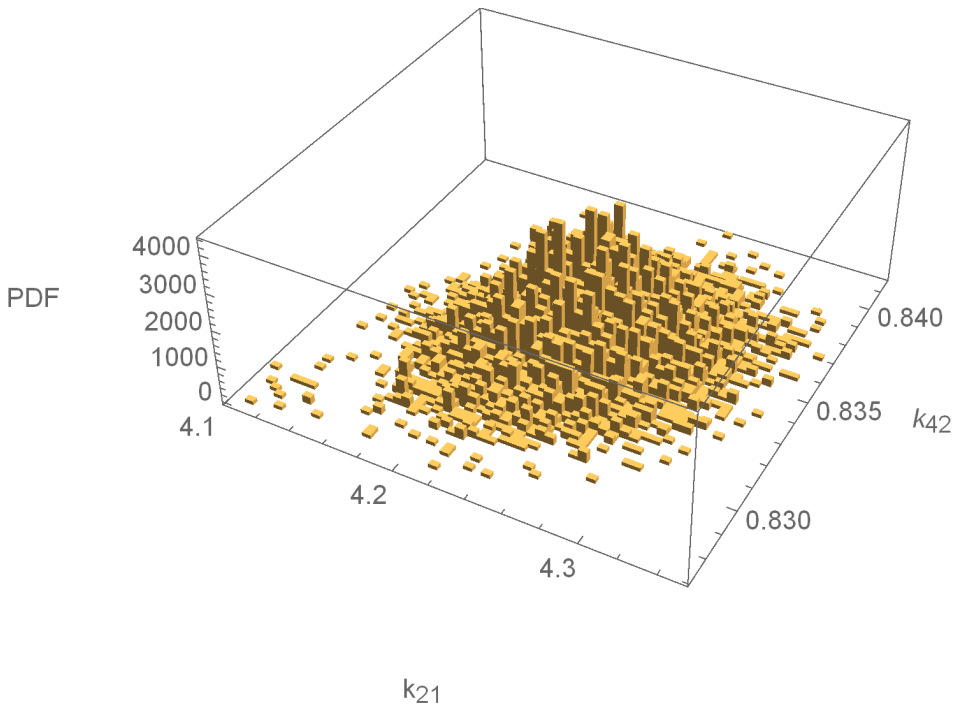

▲ Figure A.2: SMC estimate of posterior pairwise marginal distribution for $b\alpha$ and $k_{42}$ for patient A



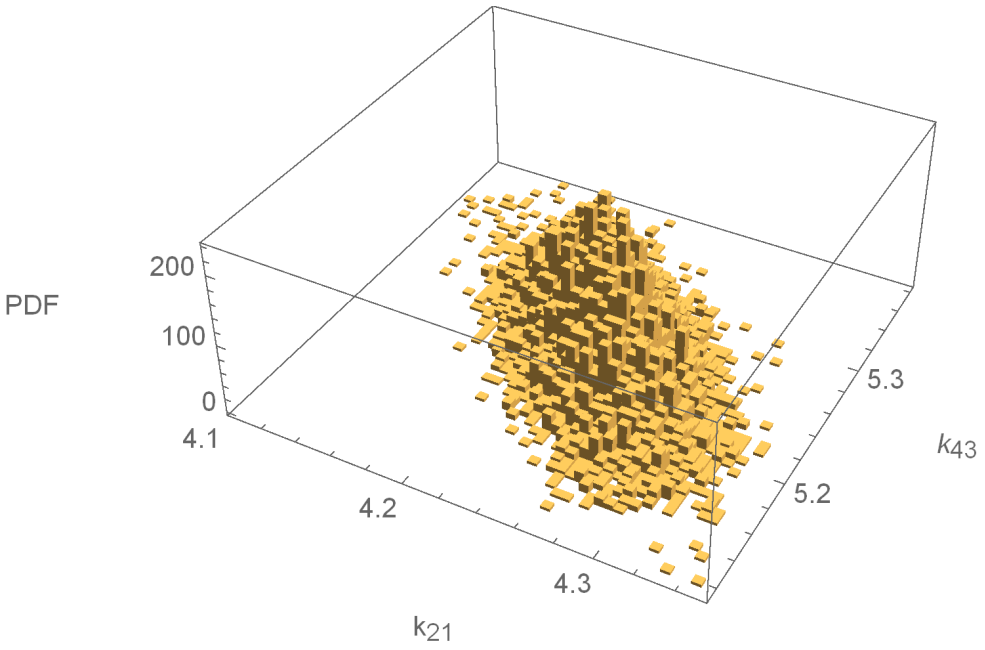

▲ Figure A.3: SMC estimate of posterior pairwise marginal distribution for $b\alpha$ and $k_{43}$ for patient A



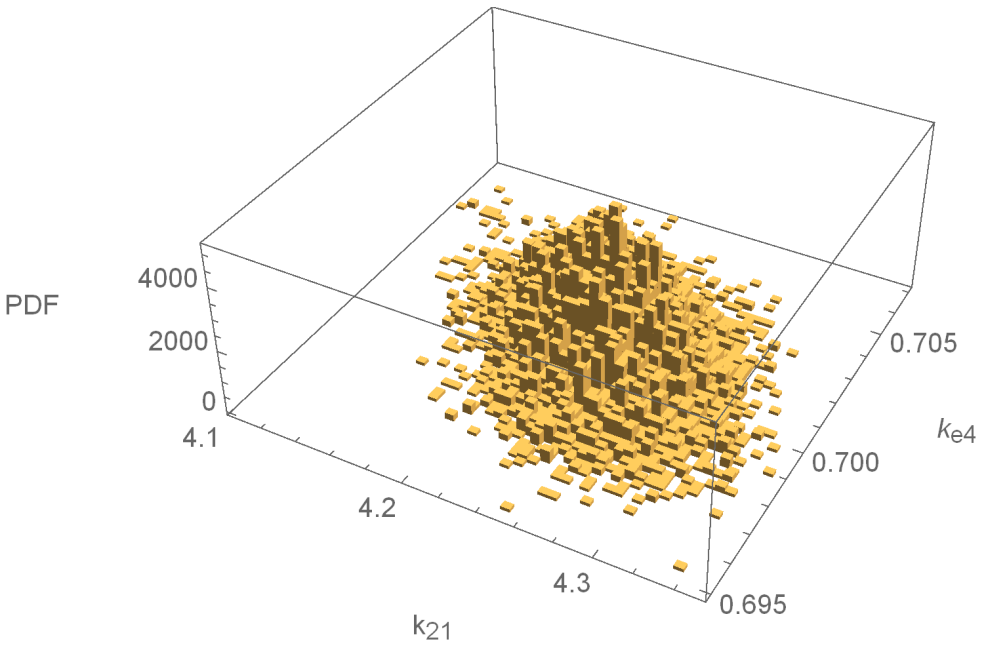

▲ Figure A.4: SMC estimate of posterior pairwise marginal distribution for $b\alpha$ and $k_{e4}$ for patient A



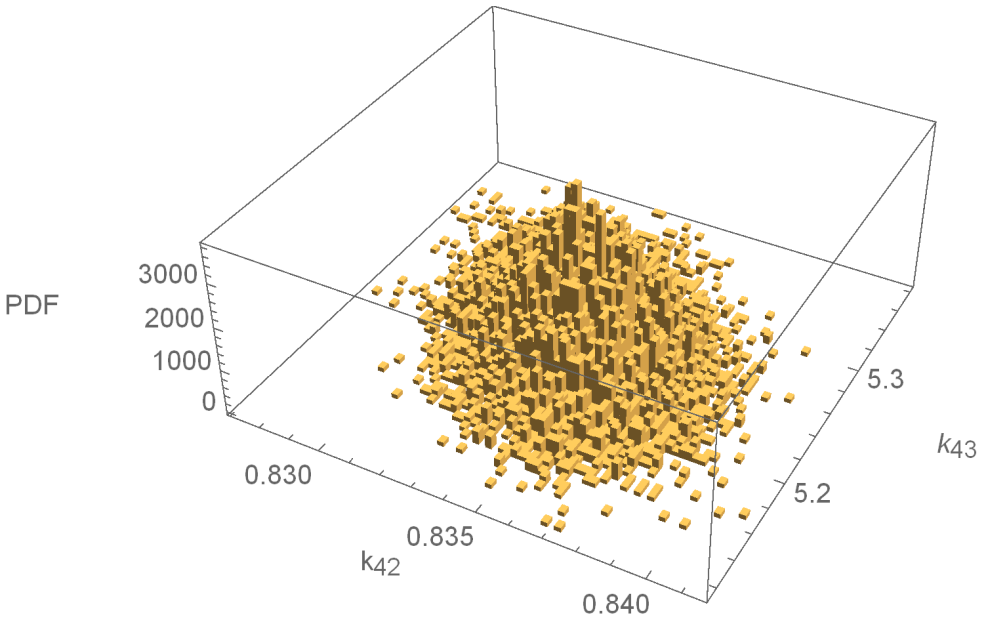

▲ Figure A.5: SMC estimate of posterior pairwise marginal distribution for $k_{21}$ and $k_{42}$ for patient A



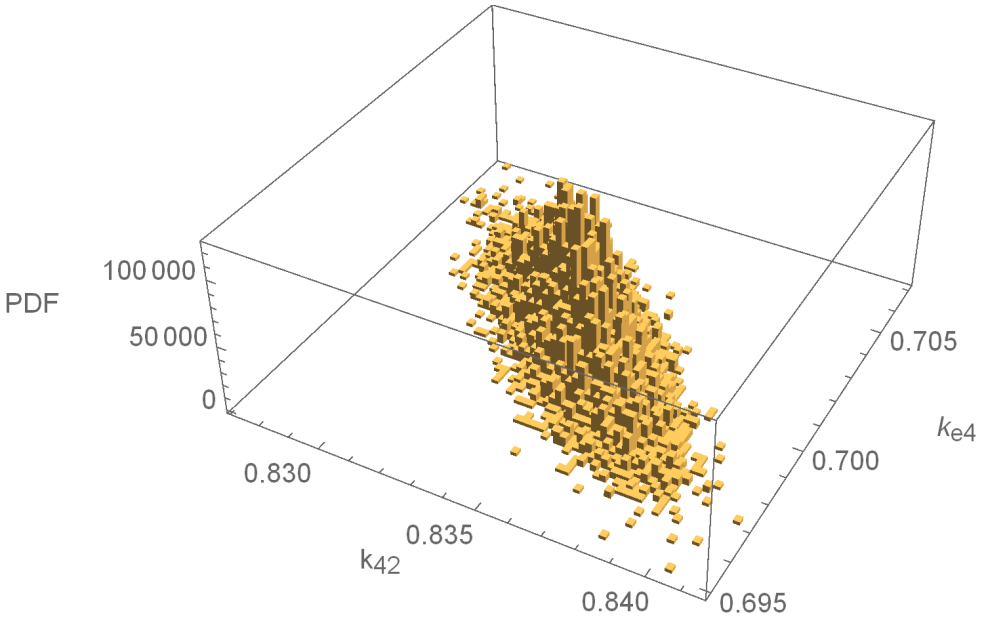

▲ Figure A.6: SMC estimate of posterior pairwise marginal distribution for $k_{21}$ and $k_{43}$ for patient A



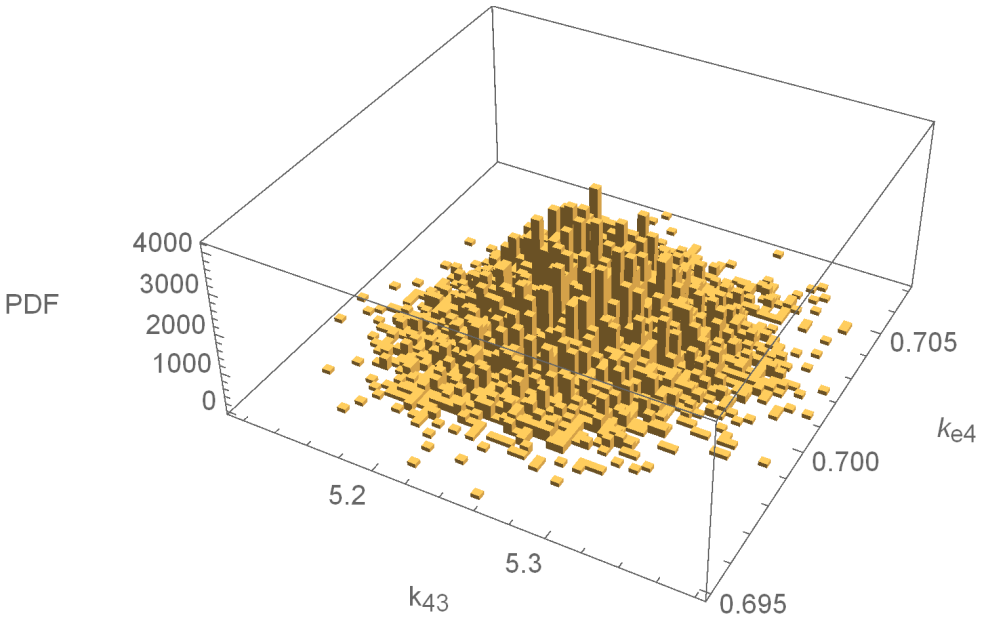

▲ Figure A.7: SMC estimate of posterior pairwise marginal distribution for $k_{21}$ and $k_{e4}$ for patient A




▲ Figure A.8: SMC estimate of posterior pairwise marginal distribution for $k_{42}$ and $k_{43}$ for patient A




▲ Figure A.9: SMC estimate of posterior pairwise marginal distribution for $k_{42}$ and $k_{e4}$ for patient A




▲ Figure A.10: SMC estimate of posterior pairwise marginal distribution for $k_{43}$ and $k_{e4}$ for patient A