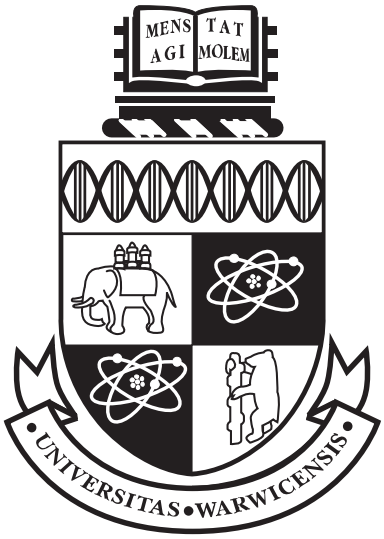


Mathematical modelling of immune condition dynamics: a clinical perspective.

by

J.G. Hattersley

Thesis submitted to the University of Warwick

in partial fulfilment of the requirements

for admission to the degree of

Doctor of Philosophy

School of Engineering

November 2009

THE UNIVERSITY OF
WARWICK

“My magnificent octopus.”

Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisors, Drs M.J. Chappell and N.D. Evans, for their support and guidance throughout this work. They have both been extremely patient in guiding me through the pitfalls of mathematical analysis. Secondly, I would like to thank my clinical collaborators for providing data and more important discussing medical complexities with a simple grease monkey: Drs Bradwell and Meade, at the Binding Site Ltd, Dr Hutchison at University Hospital Birmingham, and Drs Higgins and Zehnder at the University Hospital Coventry and Warwickshire. Finally, I wish to thank my wife, Sara, without her understanding, encouragement and diligence in deciphering my inane babble, none of this would be possible.

Declarations

This thesis is the original work of the author, with the following publications describing parts of the work:

Chapter 4

G. Mead, C. Beardsmore, S. Reid, J. Hattersley, P. Moss, G. Pratt, A. Macwhannal, A. Jacob, S. Handa, C Craddock, M Cook and S Basu. Kinetics of tumour kill during induction chemotherapy for multiple myeloma using frequent free light chain measurements. *In Proceedings of the 11th Congress of the European Hematology Association Amsterdam*, the Netherlands, Ferrata Storti Foundation, Pavia, Italy, June 15 - 18 2006, pp. 440–441.

C.A. Hutchison, M.Cook, S. Harding, G. Mead,J. Hattersley, N.D. Evans, M.J. Chappell, P. Cockwell and A. Bradwell. Mathematical modelling of free light chain removal by plasma exchange and extended hemodialysis in patients with cast nephropathy. *American Journal of Kidney Disease*, 49(4), 2007 pp. A46–A46.

C.A. Hutchison, P. Cockwell, S. Reid, K. Chandler, G.P. Mead, J. Harrison, J. Hattersley, N.D. Evans, M.J. Chappell, M. Cook, H. Goehl, M. Storr and A.R. Bradwell. Efficient Removal of Immunoglobulin Free Light Chains by Hemodialysis in Multiple Myeloma: In-vitro and in-vivo studies. *Journal of the American Society of Nephrology*, (18),2007, pp. 886-895.

C.A. Hutchison, A.R. Bradwell, M. Cook, K. Basnayake, S. Basu, S. Harding, J. Hattersley, N.D. Evans, M.J. Chappell, P. Sampson, L. Foggensteiner, D. Adu and P. Cock-

well. Treatment of Acute Renal Failure Secondary to Multiple Myeloma with Chemotherapy and Extended High Cut-Off Hemodialysis. *Clinical Journal of the American Society of Nephrology*, 4(4), 2009, pp. 745-754.

J. Hattersley, N.D. Evans, M.J. Chappell, M. Cook, P. Cockwell, A.R. Bradwell and C.A. Hutchison. Relative importance of chemotherapy versus direct removal of immunoglobulin free light chains when managing myeloma kidney: a kinetic model. *Journal of the American Society of Nephrology*, submitted Nov 2009.

Chapter 5

J. Hattersley, N.D. Evans, M.J. Chappell, G. Mead and A.R. Bradwell. A mathematical model describing the metabolism of IgG-Kappa and Lambda. *Mathematical Bioscience*, submitted March 2009.

Chapter 7

J. Hattersley, N.D. Evans, M.J. Chappell, G. Mead and C.A. Hutchison. Nonparametric Prediction of Free-Light Chain Generation in Multiple Myeloma Patients. *In Proceedings of the 17th World Congress The International Federation of Automatic Control*, Seoul, Korea, July 6-11, 2008.

V. Stovin, I. Guymer, M. Chappell and J. Hattersley. The use of deconvolution techniques to identify the fundamental mixing characteristics of urban drainage structures. *In Proceedings of the 8th International Conference on Urban Drainage Modelling*, Tokyo, Japan. Sept 7-11, 2009.

Summary

This thesis describes the use of mathematical modelling to analyse the treatment of patients with immune disorders; namely, Multiple Myeloma, a cancer of plasma cells that create excess monoclonal antibody; and kidney transplants, where the immune system produces polygonal antibodies against the implanted organ. Linear and nonlinear compartmental models play an important role in the analysis of biomedical systems; in this thesis several models are developed to describe the *in vivo* kinetics of the antibodies that are prevalent for the two disorders studied. These models are validated against patient data supplied by clinical collaborators. Through this validation process important information regarding the dynamic properties of the clinical treatment can be gathered. In order to treat patients with excess immune antibodies the clinical staff wish to reduce these high levels in the patient to near healthy concentrations. To achieve this they have two possible treatment modalities: either using artificial methods to clear the material, a process known as apheresis, or drug therapy to reduce the production of the antibody in question. Apheresis techniques differ in their ability to clear different immune complexes; the effectiveness of a range of apheresis techniques is categorised for several antibody types and antibody fragments. The models developed are then used to predict the patient response to alternative treatment methods, and schedules, to find optimal combinations. In addition, improved measurement techniques that may offer an improved diagnosis are suggested. Whilst the overall effect of drug therapy is known, through measuring the concentration of antibodies in the patient's blood, the short-term relationship between drug application and reduction in antibody synthesis is still not well defined; therefore, methods to estimate the generation rate of the immune complex, without the need for invasive procedures, are also presented.

Chapter 1

Introduction

1.1 Problem statement

The human immune system is a diverse collection of organisms and processes that enable an individual to fight a huge range of infections. It has redundancy and flexibility beyond the scope of current scientific capabilities and continues to function even within the most arduous environments. In everyday life the immune system responds to infection and invading bodies in an efficient manner, often without the individual being aware of its presence. However, as with all natural processes the immune system itself is susceptible to disease and malfunction, either through infection or genetic disorders.

Immune disorders can be broadly categorised into two types: immunocomprised, where the disease leaves the patient open to infection through lack of an immune response; or over-active, in which the immune system's reaction to the invading antigens is excessive. In this work the focus shall be on the latter condition, more specifically gammopathic immune dysfunctions. These are medical conditions that result in an increase in the synthesis of antibodies, which may be due to disease (e.g. Multiple Myeloma, Lupus, Multiple Sclerosis, or Rheumatoid Arthritis) or a natural but unwanted reaction to a medical treatment that would otherwise be beneficial to the patient (e.g. organ transplants or implanted medical devices). This imbalance allows an incredibly complex system to be modelled mathematically due to the predominance of one type of immune response; in addition, through observation of this imbalance light can also be shed on the normal operation of biological processes.

For such conditions it is possible to investigate the chemical interactions *in vitro* to yield greater understanding of the base processes involved. However the purpose of this work is to take a more holistic approach, viewing the patient as a simple system. The reason for this is to assist clinicians directly, and where possible in real-time, with ongoing patient treatment to improve the medical efficacy and the patient's experience, rather than to fuel further investigation or drug/device development. As such a control system approach is the most appropriate method for investigation, with the aim of yielding practical solutions that can be applied to patient treatment.

In order to treat patients with gammopathic diseases clinicians have two methods of treatment available:

- To reduce the production of the antibodies that are in excess. This is achieved through the administration of drugs (e.g. immunosuppression or chemotherapy).
- To remove additional antibodies through the use of artificial clearance techniques (e.g. haemodialysis).

Initially it was envisaged that with minimal knowledge of the patient, pre-treatment, control theory could be used to indicate the most appropriate treatment combinations; either in terms of the most efficient use of resources, minimising patient discomfort or achieving the best results in the shortest time. However, after preliminary investigations it was discovered that this would not be feasible due to a lack of knowledge around the underlying biological systems. Therefore fundamental work was required prior to addressing the optimality of treatment combinations, as current models do not describe the effects of either of these mechanisms in sufficient detail; neither in terms of the antibody dynamics nor the effects of the treatment in terms of the synthesis control or synthetic clearance. Therefore the focus of this work is to provide fundamental techniques and models to assist clinicians in patient prognosis and to predict, with greater confidence and accuracy, the state of the patient's medical condition and response to treatment. It is envisaged that the models and algorithms presented will help uncover the relationships between individual patients and the treatment's effect, with the hope of designing personalised patient treatment regimes in the future.

The aims of the research can, then, be simply stated as the development of novel analytical techniques to assist clinicians in diagnosis, prediction of current patient

status and description of patient prognosis, with different treatments, through the use of mathematical modelling. In scope, it is concerned with antibodies, and associated antibody fragments, and treatments which enhance the natural clearance or suppress *in vivo* synthesis. The objectives required to achieve these aims are:

- To create, or extend, where a previous model is applicable, mathematical models to effectively describe the *in vivo* dynamics of the major antibody proteins (IgG, IgA and IgM) and antibody fragments associated with Multiple Myeloma and kidney transplant.
- To validate these models analytically and against patient data, where possible.
- To generate estimates for key parameters relating to the kinetics and synthesis of antibodies.
- To use the validated models and estimated parameters to:
  - retrospectively determine the effectiveness of current treatments with provided patient data.
  - simulate patient response to alternative treatments or schedules, to determine effectiveness.
- To develop new methods to determine unmeasurable indicators of patient response to treatment.
- To investigate new analytical techniques to categorise patient response to treatments.

The primary contribution to the body of knowledge are the methods and descriptions outlined in translating medical experience and knowledge into mathematical models. However, there are several key features of this research that require highlighting, these can be broken down into contributions to mathematical modelling of biomedical systems or clinical diagnosis and treatment:

- Biomedical Modelling

  - Identification of issues surrounding current models, and extend them allow applicability to patient treatment.
  - Creation and validation of new models to describe antibody kinetics, specifically:
    * A linear kinetics model for the removal of antibody fragments.
    * A multiple filter model to investigate the impact on clearance of using dual dialysis filters.
    * A nonlinear compartmental model describing the recycling of antibodies *in vivo* and the competitive binding between different antibody types.
  - The development of deconvolution algorithms to determine the immeasurable antibody synthesis to assist in discovering the effects of treatment on the underlying antibody production. The methods developed were adaptable enough to allow their use in non-immunological problems (e.g. diabetes glucose intolerance testing and residual distribution of particles in manhole drainage).

- Clinical Practice
  - To show the beneifit of using mathematical modelling as a diagnostic tool that is applicable to ongoing patient treatment.
  - New techniques to assist in the categorisation of patients according to their response to treatment:
  - Identification of issues surrounding current best-practices regarding patients treatment and allow hypothesis testing of alternative therapies.
  - The identification and validation of a potential new markers in patients with Intact Secreting Multiple Myeloma.
  - The use of cluster analysis, in conjunction with dimension reduction methods, to classify patient response through time series data alone.

1.2 Physiological modelling

As the purpose of this thesis is to provide valid models that describe physiological processes it is beneficial to explore the methods of modelling available and explain the logic behind choosing one method over another.

A model, irrespective of domain or application, is an abstract construct that explains a real-world situation through the use of a set of simplified rules and constraints. Once a model has been created it can be used to investigate the response to differing perturbations or alternative scenarios. From these theoretical experiments it is possible to predict the outcome of real-world scenarios or improve explanations of unknown problems. In the following study the focus is upon mathematical models in which equations and formulae are used to define the model of the real process. In order to create such a model a process of conversion must take place from real-world measurements into a mathematical structure; numerous excellent texts are available on the methodology of mathematical modelling [Bender, 1978; Bonate, 2006; Carson et al., 1983; Close et al., 2002; Rohatagi et al., 2004].

The natural question that arises is: why create mathematical models at all? The most obvious answer is that of cost. Using a mathematical model can significantly reduce the resources required to investigate complex problems. It may be possible to produce the necessary experiment using real-world facilities but to repeat or modify the procedure would incur a significant cost. However, with a mathematical model, especially with the wide-spread availability of powerful computing resources, it is extremely quick and cost effective to re-run tests or multiple scenarios with significant modifications without addition cost. Alternatively, it may not always be feasible to recreate the experiment in the real-world, whether this is due to danger from the outside environs or implausible sets of constraints; or in the treatment of patients, due to the recovery being paramount over investigation of biological processes.

Mathematical modelling can also address issues surrounding the use of test subjects during the development of medical drugs or procedures [Bross, 1989]. The use of live subjects (animal or human) in testing is an ethical problem that is better discussed in medical literature; however, the use of mathematical modelling to validate clinical methods prior to moving into latter stages of drugs trials to reduce the effects on living

organisms to a minimum is an import step. The move towards mathematical modelling has been encouraged by the Food and Drug Administration (FDA) in the United States, which insists on 'significant' pharmacokinetics modelling during its review processes for new drug development; requiring that stringent criteria are met in terms of modelling results prior to moving onto Phase II[1] of a drugs development [F.D.A, 2004].

A problem that the biomedical domain often encounters is the inability to physically access key features of a system. In most engineering or scientific applications the device or system in question can be analysed and scrutinised completely, through invasive measurement techniques, e.g. gaining physical access, if required. However, with living tissue this is not always possible, or ethical, due the delicate nature of its structure; although invasive and non-invasive measurement techniques can be employed it is not always possible to gather the required data. As shall been seen in subsequent chapters it is often feasible using appropriate methods to create data that simulate the effects of hidden artifices that are immeasurable but vital to a patient's response.

System models can be divided into two main types, based on the technique employed to construct the model: they may be either data or process driven models. In data driven models the underlying mechanics are not considered and mathematical models are generated that describe statistics of the measurements and observations taken, with the hope of identifying relationships or trends that will enable analysis of the problem. However, without understanding of the processes underpinning the system in question it is more difficult to address issues of optimisation and improvement. For this reason the predominance of modelling, with the exception of some categorisation work carried out in chapter 6, is process driven. From a patient centred perspective, these models are constructed from knowledge gained from *in vitro* research and discussion with medical collaborators to construct simplified models describing the key *in vivo* dynamics of a single patient. The same basic model can then be extrapolated either 'intra', or 'inter' patient.

For pharmacokinetic modelling two common types of models are used: physiologically based pharmacokinetic modelling (PBPK) and compartmental [Droz, 1992; Levitt, 2004; Ritschel and Kearns, 2004]. PBPK modelling uses realistic anatomical

[1]FDA Phase II - phase of study where large groups of human subjects (20-300) are incorporated into the trial.

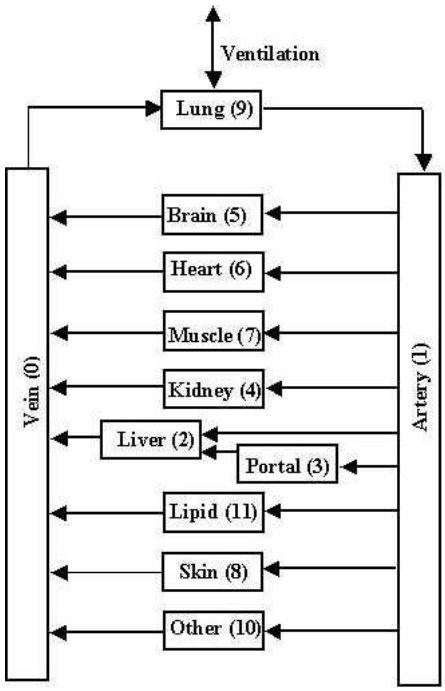


Figure 1.1: *Physiological model [Levitt, 2002].*

models to describe a drug's effect on a patient. In such models this is achieved by defining the system's organs (e.g. liver, kidney, lungs) and the interconnection and flows of material between the organs in a network flow schematic. An example of a physiological schematic can be seen in Figure 1.1. After the model is created a series of equations can be derived relating the flow-rates and drug kinetics. Physiological models require detailed knowledge and data of how material flows between the organs, and the potential for each organ to store and maintain the substance in question. This level of information can be gathered but requires invasive imaging and tissue sampling and as such is not available during standard treatment procedures. However, an interesting example of a PBPK model describing the distribution and clearance of injected monoclonal antibody was presented by Ferl et al. [2005], the model used mice data for evaluation and validation; this is discussed in more detail in Chapter 5. For the immune conditions studied within this thesis the measurement data available are limited to measurements in plasma only, infrequently taken and at intervals that are convenient for the patient, given the ongoing treatment. Therefore, given the data available, compartmental modelling is a far more suitable approach.

1.3 Structure of thesis

This thesis is divided in to eight chapters. In the second chapter the biological background of the immune complexes and their *in vivo* catabolism shall be described. Whilst the background is not in great detail it should be sufficient for the reader to understand the medical treatments under consideration during the subsequent chapters and assist in the evaluation of the boundaries of the modelling problem, e.g. what is necessary to model, what may be omitted or assumed. Chapter three contains preliminary material on compartmental analysis and model validation which are fundamental in understanding the models, results and conclusions in the subsequent four chapters. The remaining chapters focus on the models developed to describe human immune conditions kinetics. Each chapter builds on the work described in the preceding chapter to include greater granularity of antibody kinetics.

The mathematical modelling begins in chapter four, with models that describe the kinetics of antibodies fragments, known as Light Chains. These are studied through data gained from the Birmingham University Hospital Renal Unit, from patients with a condition known as Free Light Chain secreting Multiple Myeloma. Patients with this condition are treated with apheresis techniques and chemotherapy; therefore, the models described are extended to accommodate these treatments. Recommendations are drawn from the simulation analysis regarding the optimal apheresis modality.

In chapter five the models are extended to incorporate whole antibody kinetics. These are constructed of the Light Chains described in the previous chapter but they are catabolised and metabolised differently *in vivo*. As such the model is extended to incorporate this. In this form of Multiple Myeloma the cancerous tumours secret intact, or whole, antibodies. Intact Secreting Myeloma patients produce monoclonal antibodies, the most common of which is IgG; in fact all patient data available for this study is IgG. Therefore the models produced focus solely upon this antibody. However, IgG, in terms of the kinetic behaviour, is the most complex of the antibodies with other antibody subclasses (e.g. IgA and IgM) following the same metabolic pathways. Although the models are developed to assist in Multiple Myeloma patients they are applicable for more generic applications and this is exploited in the following chapter to investigate the three main types of antibody produced against medical implants:

IgG, IgA and IgM. Whilst each type behaves slightly differently *in vivo*, the systems dynamics can be derived from the base models presented in the previous chapter. In conjunction with Coventry and Warwickshire University Hospitals, the models are then applied to the problem of estimating synthesis of antibody in kidney transplant patients. In addition to providing standard antibody measurements, clinicians were able to provide information regarding the patient's production of antibodies specifically against the donated organ. However, due to anomalies in the data, modelling techniques previously presented were not applicable and promising preliminary work is presented regarding the categorising of patient responses to transplants using cluster analysis of time series data.

A key feature of any system is its input. Whilst in classical engineering problems these are often known phenomena, this is rarely the case in the biomedical domain when considering real patients. In the previous chapters the study was limited to cases where patients were assumed to be producing antibodies at a constant rate or the synthesis could be described by a known functional form, e.g. exponential or logistic decay. In Chapter 7 a technique known as deconvolution is used to estimate the antibody synthesis under the influence of chemotherapy treatment. This method extends a technique from astrophysics called Maximum Entropy Deconvolution to provide optimal estimates with respect to the amount of information present in the data. The results allow several important conclusions to be drawn regarding the effect of chemotherapy with respect to the suppression of tumour producing cells. The versatility of this deconvolution method is then demonstrated by using the Maximum Entropy method to estimate, non-parametrically, other nonlinear features of several other models, including a patient's volume changes due to hypoproteinemia and the synthesis of insulin in response to glucose intolerance tests in diabetes. Finally, digressing slightly from the pharmacokinetics theme, the technique has proved successful, without modification, in a non-biomedical domain to estimate system impulse responses of man-holes in drainage systems; this enable the residual distribution time of particles within a drain to be estimated. This problem has implication in the long term accumulation of particulate matter from drainage systems, e.g. disposal of drugs after medical treatment.

To assist in comprehension each chapter will close with a discussion of the key findings and important features developed in it. A summary of conclusions drawn

by all the work will be succinctly outlined in the closing chapter. This chapter will also address possible avenues for future research that may be used to build upon the theories presented in this thesis.

Chapter 2

Medical background

This chapter presents the basics of immunology and the treatments used on patients with gammopathic immune dysfunction. Due to the complex nature of the immune system it is not possible to provide a complete description within this thesis; however, it is anticipated that sufficient detail will been given to allow the analysis performed in later chapters to be understood without reference to external sources.

The following chapter is separated into four sections: the first discusses the immune system and natural immune responses, the second, focuses upon immune conditions that may occur naturally or due to disease. The final two sections discuss the treatments available to patients with problematic immune systems and the measurement techniques used by clinicians to assist in diagnosis and monitoring.

2.1 The immune system

The human immune system is a collection of complex biological processes that combine to protect the host against infection and disease through external agents known as antigens. An antigen is a foreign body that invokes a reaction by the immune system. This response is separated into two branches [Davies, 1997; Janeway et al., 1999; Manning and Turner, 1976]:

- Humoral - produces antibodies that detect and neutralise antigens. This is often called the 'specific' response.

- Cell mediated - the immune response that does not use antibodies to remove infection, referred to as 'non-specific' response.

The work presented here focuses on the humoral response, as the immune conditions investigated produce measurable excess antibodies which have a detrimental impact upon the patient. To model cell mediated response would require extensive laboratory work, which would detract from the primary direction of this study, namely, practical recommendations for patient treatment.

2.1.1 Humoral immune response

The process of antibody synthesis, or production, begins with B lymphocytes (more commonly referred to as 'B cells'), that are produced in bone marrow. After a series of maturation stages B cells differentiate into 2 types: memory cells, to allow information to be retained relating to an antigen for fast response on subsequent re-infection; or plasma cell, which exist mainly in bone marrow, and secrete antibodies directly into the blood [Davies, 1997]. The structure of a generic antibody can be seen in Figure 2.1.

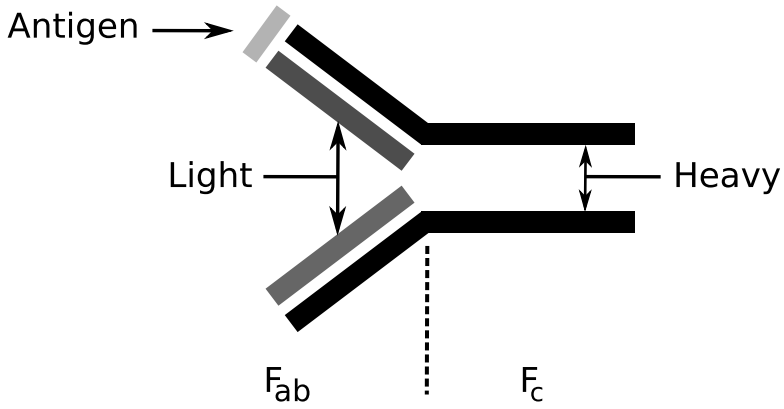


Figure 2.1: *Simple antibody structure*

The antibody is made of two polypeptide sub-units: two larger units, called heavy-chains, and two smaller ones, referred to as light-chains. There are different types of heavy-chains which combine to form monomers, dimers and polymers; a simple diagram of these can be seen in Figure 2.2. There are five classes (isotypes) of antibody created

in response to infection; a list of the types, their location, function and that percentage of immune response each isotype contributes to the total, can be seen in Table 2.1.

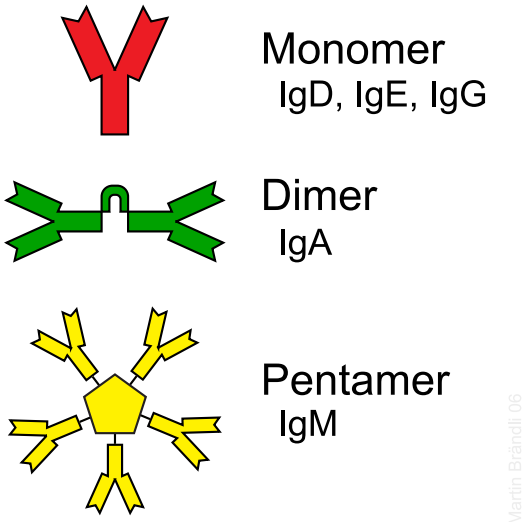


Figure 2.2: *Structure of antibodies [Brandli, 2008].*

An antibody binds to a specific antigen through the F_{ab} portion of its structure, whilst the opposite end, referred to as the F_c region, is responsible for signalling other cells to assist in removing the antigen. Whilst it is popularly viewed that antibodies are responsible for the destruction of antigens they are in fact unable to penetrate cell walls and are only used to inform other cells (T lymphocytes) of the presence of an antigen, that actually perform the destruction of the antigen. The F_c region is also used by the body to prevent IgG isotypes from being catabolised (see Section 2.1.2), this process shall be explored in detail in Chapter 5.

For the medical conditions considered in this thesis only IgA, IgM and IgG are relevant and as such the focus of the modelling will be around these proteins. It can also been seen from Table 2.1 that these account for the majority of the immune response, although, due to the structural similarity of the immune complexes, the models developed are applicable to IgD and IgE, if required in further investigations.

2.1.2 Natural clearance

Due to physical size, whole antibody and antibody fragments are cleared through different clearance pathways.

Table 2.1: *Classes of antibody [King, 1998, p. 9-11].*

Class	Function	%
IgA	Predominant in mucous and is responsible for protecting exposed areas of the body, e.g. nose, breathing passages, etc.	$10\% - 15\%$
IgG	The most abundant antibody, it is found in all bodily fluids to detect plasma infections.	$75\% - 85\%$
IgM	Largest of the antibody, similar to IgG but with multiple F_{ab} regions.	$5\% - 10\%$
IgD	Found in the abdomen, it's function is currently unclear.	$< 1\%$
IgE	Located in lungs, skin, airways - reacts against foreign bodies and responsible for allergic response.	$< 1\%$

Whole antibody

All the antibodies and constituent elements discussed above are in an elaborate metabolic cycle, constantly being synthesised and cleared by natural processes. Whilst the mechanics of these processes are often complex and poorly understood, qualitative studies have been conducted using radiolabelling to determine the half-life and catabolic pathways of the five Ab groups [Barth et al., 1964; Rogentine et al., 1966; Strober et al., 1968; Waldman and Strober, 1969; Waldmann et al., 1976]. The results of these studies are summarised in Table 2.2. In addition to the half-lives, the table also shows the Fractional Catabolic Rate (FCR) for each antibody. This a common clearance measure used in immunology and is the percentage of antibody cleared from the body per day.

Table 2.2 shows the different rates at which antibodies are cleared, the most notable of these is IgG, with a very low clearance rate, and conversely, extended half-life; approximately 5 times greater than an other antibody type. It is believed that all other antibodies with the exception of IgG are removed at a constant rate from the plasma

Table 2.2: *Antibody clearance and Half-life, summarised from [Mariani and Strober, 1990].*

Class	FCR (% cleared/day)	$\frac{1}{2}$-life (days)
IgG	6.7	23
IgA	25.2	5.8
IgM	18.6	5.1
IgD	37	2.8
IgE	62.0	1.3

pool irrespective of the concentration of that antibody. The clearance of IgG, however, has been shown to vary greatly depending on the *in vivo* concentration [Waldman and Strober, 1969]. At greater concentrations the clearance increases, whilst at low concentrations it decreases. This is believed to occur due to a receptor, referred to as the neo-natal receptor or FcRn, that binds to circulating IgG in the intravascular pool and prevents it from being catabolised [Brambell et al., 1964]. The overall effect of this is to increase the half-life of IgG significantly. Figure 2.3 shows estimates for the fractional catabolic rate of IgG. The estimates were calculated by Waldman and Strober [1969] from patients with different serum concentrations using radiolabelling. Even given the patient variability, the trend of high clearance at higher serum concentration can clearly be seen. This will be discussed further in Chapter 5.

In patients without this receptor [Waldmann and Terry, 1990] the half-life is in the range of other antibodies shown in Table 2.2, directly indicating recycling capabilities of the FcRn.

Light-chains

In a healthy subject light-chain fragments are naturally produced in abundance of heavy-chain fragments (see Figure 2.1), resulting in an excess of light-chains, known as 'Free Light-chains', or FLCs. It is therefore necessary for the body to clear these from plasma. There are currently two known pathways of clearance for FLC [Bradwell

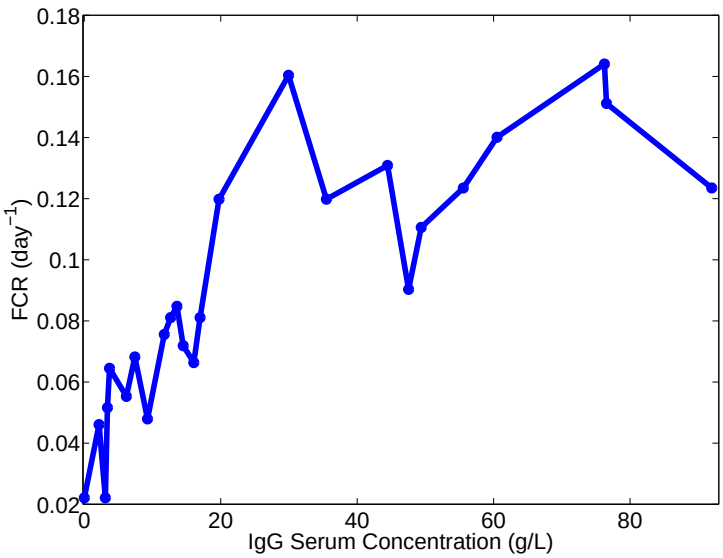


Figure 2.3: *IgG Fractional Catabolic Rate (FCR) clearance of IgG, data from Waldman and Strober [1969].*

et al., 2005]:

- Light-chains, unlike intact immunoglobulins, are small enough to be cleared via glomeruli in the kidneys.
- The Reticuloendothelial System (RES) is a system that is responsible for the removal of 'worn out' cells within the body. This system provides an underlying clearance of light-chains.

Removal via the kidneys, under normal immune conditions, is perfectly adequate to maintain homeostasis of light-chains. However, if the kidneys are not functioning correctly, through disease or damage, this route will not be available. To compound this problem excess FLCs form cysts within nephrons of the kidney and block the filtering of material, resulting in poor kidney function or failure, often resulting in apheresis [Bradwell et al., 2005; Hutchison et al., 2007], see Section 2.3.2.

2.1.3 Human Leukocyte Antigen

Human Leukocyte Antigens, or HLAs, will become important in a later chapter (Chapter 6), but are relevant to the immune system described here. In short, contrary to the

name antigen, HLAs are markers that allow the immune system to recognise 'self' from 'non-self', and respond accordingly [Owen and Lamb, 1988]. HLAs are of great importance in the rejection of graft implants, for example, kidney transplants (see section 2.2.2). As the donated tissue is allograft[1] it will express 'non-self' HLA and as such antibodies will be produced against the tissues. These can be measured to assist in categorisation of the patient's response to the transplant but are simply IgG antibodies as described above.

2.2 Gammopathic immune conditions

Whilst the models created for the immune complex kinetics are generic in nature, and are applicable to many diseases and non-standard immune conditions, the models are validated against two primary conditions in this thesis: Multiple Myeloma and immune reaction to kidney transplant (due to the proximity of collaborators and availability of data). Whilst it is possible to conduct clinical trials of normal and unhealthy patients, this would require extensive cost, time and complex medical trial ethics to be overcome. However, as a patient undergoing clinical treatment will be routinely tested as part of their treatment, these data can be used to elicit information regarding the underlying kinetics. In addition, patients with the two conditions considered require apheresis and synthesis reducing drugs; as such the treatment regime can be viewed as a method of perturbing the system to expose the fundamental system features.

2.2.1 Multiple Myeloma

Multiple Myeloma (MM or Myeloma) is a cancer of the plasma cells which accounts for around 1% of all cancer cases in the UK. Approximately 3,750 people are affected by the condition annually in the UK; whilst in the United States there are around 13,000 per year. There are approximately 150,000 people with the disease globally. After non-Hodgkin lymphoma, it is the most prevalent hematological malignancy (cancers of the blood, lymph and bone marrow). Myeloma patients present with a variety of symptoms ranging from bone pain, to anaemia and renal failure. There is a median

[1] A transplant from a non-identical member of the same species.

survival time of approximately 62 months, depending on the stage of the disease at diagnosis [Bradwell et al., 2005; Cancer-Research, UK; Pozzi, 1987].

In a healthy individual 'normal' plasma cells produce a range of antibodies against invading antigens. In a patient with Myeloma the plasma cells become tumourous and as such single antibody types, referred to as monoclonal antibodies, are generated that react only against a single antigen. This prevents an effective immune response to a range of invading antigen, leaving the patient immunocompromised and prone to infection. Myeloma is categorised into four main groups [Bradwell et al., 2005]:

- **Light-chain Multiple Myeloma**

  These patients do not produce whole monoclonal antibodies in excess, only the Light-chain fragments. 15% of Myeloma patients have this form of the disease.

- **Intact Immunoglobulin Multiple Myeloma**

  Patients in this category produce fully constructed monoclonal antibodies; in addition, the majority of patients also produce excess FLCs. The monoclonal antibodies can be of any of the isotype (IgA, IgG, etc). Around 83% of Myeloma patients over-produce intact immunoglobulin G (IgG).

- **Nonsecretory Myeloma**

  $1-5\%$ of Myeloma patients do not secrete monoclonal antibodies; the disease is diagnosed through bone marrow biopsies, and as such is difficult to detect.

- **Asymptotic (or Smoldering) Myeloma**

  Asymptotic Myeloma patients have low levels of monoclonal antibody production and tumourous plasma cell concentrations but the condition is not adversely affecting other organs. Less than 1% of patients present with this type of Myeloma.

The latter two conditions (Nonsecretory and Smoldering MM) do not produce measurable excess immunoglobulin, or associated fragments. They are therefore not considered in this study; these are rare conditions representing a small percentage of patients with Multiple Myeloma, therefore obtaining data and information is more difficult.

2.2.2 Antibody incompatible transplantation

The second medical condition that will be a focus in this thesis is antibody incompatibility during organ replacement; specifically kidney transplants. However, the techniques presented are equally applicable to other organs or medical device implants.

Unlike the disease condition Multiple Myeloma, the immune response to transplant is natural but unwanted. Once a foreign substance has been implanted into the body the host produces a range of antibodies against it, all with the purposes of attacking the antigen and removing it from the body, which can ultimately result in failure of the organ or medical device, and as such is a response that must be kept to a minimum, or stopped altogether. In standard transplants the clinicians will attempt to match the donated organ to the host immune system in terms of both the blood group type (A, O, AB and B; rhesus status positive or negative) and HLA types [Campbell and Halloran, 1996]. The probability that the patient and donor are compatible to reduce rejection are extremely restrictive and in many cases cannot be found in a time frame applicable for the patient's survival. In the UK 25% of all patients on the donor transplant list have HLAs that prevent transplant, resulting in about 250 transplants being cancelled annually due to HLA issues [Higgins, 2007]. To reduce the number of patient deaths whilst awaiting transplant Takahashi et al. [2004] began using kidneys from incompatible donors in transplants. The methods developed by Takahashi et al. [2004] have been replicated and modified by several groups [Higgins, 2007; Jordan et al., 2003; Montgomery et al., 2000].

The immune response to incompatible transplants is considerable, and much greater than that produced when equivalently matched donor organs are used. It is this strength of response that enabled analogies between the cancer driven synthesis of Multiple Myeloma to be drawn and the modelling techniques used laterally in the description of the underlying biological processes of transplantation.

2.3 Medical treatments

In medical conditions with excess antibodies the primary concern of the medical treatment is to reduce the high concentrations of antibodies within the patient to ranges

considered normal for a healthy individual. There are two mechanisms by which this can be achieved: to reduce the synthesis, and to artificially clear more from the system than is produced. The former, reduction of synthesis, is achieved primarily through drug combinations; to increase clearance, apheresis techniques are used.

2.3.1 Drug treatments

There are three possible types of treatment that affect the immune system [Galton and Brito-Babapulle, 1990; Halloran, 2004]:

- Immunotherapy

  Although the term 'immunotherapy' refers to any treatment that modulates the immune systems, in the conditions considered the main concern is immunosuppression, i.e. drugs that reduce the efficacy of the immune system from responding to antigens or to increase the reaction against cancerous plasma cells.

- Chemotherapy

  Chemotherapy drugs intervene in the cell life-cycle to reduce the number of cancerous plasma cells, either by encouraging apoptosis (cell death) or preventing mitosis (cell division). Combinations of drugs are used (e.g. Vincristine, Adriamycin, Dexamethasone, Thalidomide) often including the immunosupressing drugs discussed previously.

- Radiotherapy

  Radiotherapy uses radiation to kill tumourous cells and prior to chemotherapy was solely used to treat Myeloma patients. However, with the introduction of chemotherapy, its use has been limited to analgesic relief or locally identifiable tumours.

Chemotherapy and radiotherapy are used on patients with production linked to tumourous conditions (e.g. Multiple Myeloma) and after each of these treatments the patient will require bone marrow or stem cell transplants to replace the destroyed plasma cells [Alexanian et al., 1977]. Immunosuppresion is not applicable for cancer patients, although the same immunosuppressing drugs may be used in the combinatorial treatment. It is used primarily in transplant patients [Halloran, 2004].

The specific drug combination used in all cases is largely down to clinical experience and availability. The effectiveness of the treatment is monitored by measuring concentrations of antibodies in the patient's blood.

2.3.2 Extracorporeal removal (Apheresis)

Apheresis is the term for medical techniques that remove elements from the patient's blood stream. In most cases this involves taking the blood out of the patient into an external apparatus that is responsible for the separation of elements which are to be kept and those that are to be removed. A schematic diagram outlining the elements of an extracorporeal circuit can be seen in Figure 2.4. The circuit is used to extract the blood from the patient's venous blood stream and re-inject the cleaned blood back to arterial blood flow; the circuit consists of: a device for extracting proteins (in this case a dialysis filter); pumps for maintaining a consistent blood flow through the device; connective tubing and connectors; monitors and values to monitor patient safety. The connective tubing includes injection points along the length to allow for the introduction of drugs during treatment (e.g. heparin, to prevent blood clotting) or to remove small blood samples without disturbing the patient [Henrich, 2004; Nissenson and Fine, 2005].

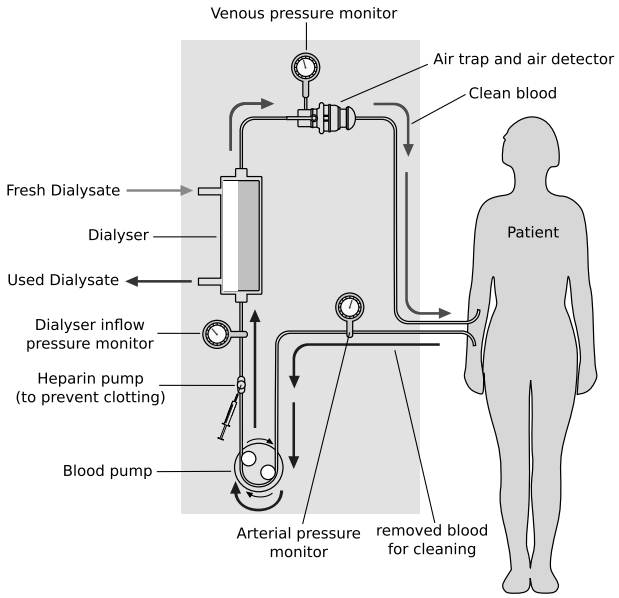


Figure 2.4: *Haemodialysis schematic [Mrabet., 2008].*

In general apheresis can be broadly categorised into dialysis and therapeutic apheresis [Daugirdas, 2006]. Which technique is used is based largely on the size of the molecule to be removed. For small proteins (< 10 KDa) haemodialysis is used, it is relatively cheap and easy to implement. In dialysis, or more precisely haemodialysis (as this study is concerned solely with clearance from blood) the particles are removed by passing them across a semipermeable membrane which, through convection (pressure gradients) and diffusion (concentration gradients), can be used to remove different elements from the blood. Haemodialysis is most commonly used to remove toxins from the blood in the advent of kidney failure. However, recently filters have been developed by manufacturers ([Toray Industries Inc., 2009] and [Gambro Dialysatoren GmbH, 2009]), which are able to pass larger molecules. These have been used to remove FLC antibody fragments [Hutchison et al., 2007]. This process is studied in more detail in Chapter 4. For larger molecules, therapeutic apheresis is used.

There are several methods categorised under the term therapeutic apheresis and two treatment types are used in this study: plasmapheresis and plasma exchange. Plasmapheresis uses a centrifuge device that the blood is passed through to separate out molecules of a particular size, whilst proteins that are not to be removed are returned to the blood and re-introduced in to the patient's blood stream. Plasma exchange is a simple process that involves removing a pre-defined volume of plasma over the treatment period, approximately three hours, and replacing it with an equal quantity of a saline solution. This has the effect of removing a portion of the offending substance; however, it is a nonspecific treatment removing all substances from plasma, so unlike plasmapheresis and haemodialysis, which extract particular elements from plasma, plasma exchange is applied with strict limits, and is usually limited to three or four treatments over a two week period [Clark et al., 2005; Zucchelli et al., 1988].

2.4 Measurement techniques

In order to begin the modelling process it is important to have an understanding and recognition of the methods clinicians have available to them to determine the levels of antibodies in the patient. It is known that there is a linear relationship between a plasma cell and the number of antibodies it produces [Sullivan and Salmon, 1972].

Therefore the most effective diagnostic measurement tool is a tissue sample to confirm and quantify the presence of plasma cells, from which the antibody production can be inferred. However, as the plasma cells exist solely within bone marrow this becomes an extremely intrusive and painful treatment and as such is rarely performed. No biopsy data are available for any of the patients treated during this work. Clinicians therefore measure the concentration of antibody in plasma.

There are several methods used to measure antibody concentrations (e.g. immuno-electrophersis, immunoassay, immunofluorescence, immunoblot) [Bradwell et al., 2005; George and Urch, 2000]. A study of these measurement techniques is beyond the scope of this work. Fortunately, clinicians provided measurements of the antibody titres in standard units of concentration, generally g/L, in addition to possible errors that may exist in the measurement. The exception to this is the measurement of Donor Specific Antibodies, in which only a mean fluorescence indicator is available, therefore requiring pre-processing prior to use of the data; this is covered in detail in Chapter 6.

Chapter 3

Compartmental modelling

In the modelling of physiological systems an approach that is often used is that of compartmental analysis [Brown, 1985; Godfrey, 1983; Jacquez, 1996]. It is a lumped system approach which focuses upon the quantity of material to be observed (e.g. antibodies), and how it moves around the system (e.g. the human body). Crucial to this is the understanding of a compartment: a compartment is simply a subsystem that contains homogeneously well mixed concentrations, or quantity, of the material of interest. For example, as previously mentioned (see Chapter 2), antibodies are synthesised by plasma cells and 'released' into the patient's blood stream; therefore, a candidate compartment would be blood. It should be noted that not all compartments have direct physical analogies, nor exist as a separate entity in the physical space, e.g. antibodies bound to a receptor or antigen share the same physical space but may represented by separate compartments. A system is modelled as a series of interconnected compartments. For all the models considered in this work the exchange between these compartments is assumed to occur instantaneously and all the individual elements within a compartment (e.g. antibodies or antibody fragments) are equally likely to exchange with other compartments.

3.1 Model construction

Given the above definition of a compartmental model the question naturally arises on how a complex system, such as the immune system kinetics can be modelled using

such an approach. A method of model construction is defined by Walter and Contreras [1999]:

1. Divide the problem into a collection of compartments and sub-compartments by reviewing the problem domain.

2. Determine the material flow between the compartments.

3. Perform preliminary analysis and validation of the model using mathematical techniques.

4. Identify appropriate functions for the transfer of material between compartments and the system inputs/outputs.

5. Establish initial conditions and compartment levels from established data.

6. Perform analysis on the model by manipulating the differential equations manually, or through computer simulation depending on model complexity.

Whilst it may at times be possible to skip steps (e.g. by deriving differential equations directly) compartmental analysis forces a more rigorous methodology upon model construction that can often lead to clearer understanding and mitigate against simple human error.

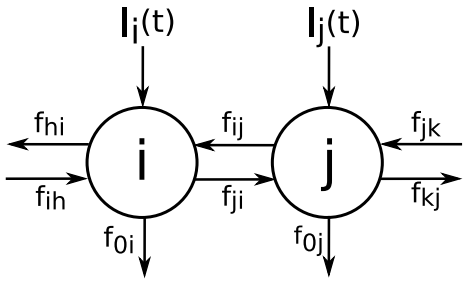


Figure 3.1: Generic compartmental model structure

Compartmental models are often expressed through simple schematic diagrams (see Figure 3.1), showing the compartments, inputs, outputs and compartment connectivity. Such a simple representation leads to a feature of compartmental models which is often overlooked, their ability allow multidisciplinary research teams to express ideas;;

ensuring mathematicians, engineers and medical staff are able to communicate at different levels about a common problem, facilitating model development to achieve a simple form understandable by each group.

The process of modelling begins with the investigation of the biology involved, from which the underlying process and governing equations can be extracted. Through the study of the biology, compartments should be identified as areas of the body which interact with the material in question; the interactions high-light the movement of material between compartments. Once the compartments and interactions are defined it is possible, through mass balance principles, to generate a system of differential equations mathematically describing the compartmental quantities over time [Godfrey, 1983; Jacquez, 1996]:

$$\frac{\mathrm{d}q_i(t)}{\mathrm{d}t} = -f_{0i}(t)q_i(t) + \sum_{j\neq i} f_{ij}(t)q_j(t) - \sum_{i\neq j} f_{ji}(t)q_i(t) + I_i(t) \tag{3.1}$$

where q_i is the quantity of material in compartment i, f_{ij} is the flow of material to compartment i from compartment j and the subscript 0 indicates exchange with the system environment; I_i is the rate at which the material enters compartment i. All the functions f in the above equation can be either linear or nonlinear, depending on the process modelled. Due to physical properties of biological systems in question, and the mass-balance principles evoked, a compartmental model inherits several key features of positivity that are important when validating the model; namely $q_i(t) \geq 0, f_{ij}(t) \geq 0$ and $I_i(t) \geq 0$ for all $t > 0$.

Whilst the representation of the modelled system shown in equation 3.1 is used through out this work, alternative forms are used. When referring to general functional form of a model, without the need for specific structure, the system may be expressed as:

$$\begin{aligned} \dot{q}(t,p) &= f(q(t,p), u(t), p, t), \\ y(t,p) &= h(q(t,p), p), \\ q(0) &= q_0(p), \end{aligned} \tag{3.2}$$

where $q(t)$ is the state variable vector ($[q_1(t)\ q_2(t)\ ...\ q_n(t)]^T$) and p a vector of model parameters; $u(t)$ defines the input vectors. However, it is often necessary to refer to the elements relating to the model structure without defining the full differential form (as in equation 3.1), in which case a state-space representation can be defined as:

$$\begin{aligned} \dot{q}(t,p) &= A(q(t,p),p,t)q(t) + B(p)u(t), \\ y(t) &= C(p)q(t,p), \\ q(0) &= q_0(p), \end{aligned} \tag{3.3}$$

where $A(.)$ is the system matrix relating the movement of material between compartments; $B(p)$ the input gain matrix and $C(p)$ the observation matrix describing which compartments are measured by the specific experiment. In the general state-space definition it is possible that the matrices (A, B and C) may depend on the system states, parameter vector and time; however, for the systems studied in this work a more restrictive structure is used: the matrices B and C are strictly linear with simple scalar elements representing input and output gain, but the system matrix (A) may be nonlinear and non-autonomous, depending on both the system states and time in addition to the parameter vector.

3.2 Model validation

The purpose of constructing a model to predict the real system response to new inputs (e.g. medical treatments) or environmental conditions (e.g. patient health) and elicit information regarding the current state of the system (e.g. the patient); however prior to evaluating these scenarios it is necessary to validate the model: that is, does the model adequately describe the real system for the intended purpose. For the case of treating patients in a clinical setting this involves comparing the model's output to samples taken, from a single patient, during a known procedure or treatment. In order to compare the output with the data, the model must first be initialised with values for the internal parameters (e.g. initial conditions, transfer co-efficients, etc). Whilst it may be possible to obtain some of these from the literature, or previous experiments, this may not always the case. In addition, there will be significant variation in the

parameter values for the same patient on different days, with different conditions, and between patients undergoing the same treatment.

It is therefore necessary to perform parameter estimation during model validation, this is often referred to as the 'inverse problem' [Bard, 1974; Jacquez, 1996], in which the parameters of the model are modified to allow the output of the model to closely match the real system output. In addition to allowing validation the parameter estimation process can offer insight into how different treatments affect the underlying process and gain knowledge on how intra and inter-patient variation may alter the impact of a specific treatment.

Prior to performing the parameter estimation, two important questions must be answered:

- Is it possible to estimate the model parameters given the observations available? This can be addressed analytically by a field of research referred to as Structural, or *A priori*, Identifiability [Bellman and Aström, 1970; Jacquez, 1996; Walter, 1982].

- If the parameters are, in theory, identifiable does the experiment accuracy and observations made allow the parameters to be reliably estimated in practice? The system outputs response to perturbations in key parameters can be used to investigate this issue, a field of study referred to as Sensitivity Analysis [Tomovic and Vukobratovic, 1972; Vajda et al., 1985].

3.2.1 Structural identifiability

The structural identifiability of a model can be categorised through the following definition.

Definition 3.2.1. Given a system (as defined in eqns (3.3)), an ideal measurement of the output ($y(t)$) and known input functions ($u(t)$) and a parameter vector p, the system can be either:

- Globally Identifiable (S.G.I), if for all values of p, the input-output behaviour (defined by eqn. (3.3)) of the system uniquely determines p.

- Locally Identifiable (S.L.I), if there is a countable set of values for p that give the same output behaviour.

- Unidentifiable, if there is an uncountably infinite set of values for p that give rise to the same observed output.

For a formal mathematical definition of identifiability see Chapman and Godfrey [1996] or Evans et al. [2002]. An extension to structurally identifiability is that of structural indistinguishability. That is whether two model structures are indistinguishable in terms of the observable input-output behaviour [Chapman and Godfrey, 1996; Evans et al., 2004; Godfrey and Chapman, 1989]. This is not an issue for the models presented in this work due to their uniqueness with no opposing theoretical structures.

Several analytical methods are available for performing identifiability analysis depending on whether the model is linear or nonlinear [Chappell et al., 1990; Jacquez, 1996; Walter, 1982]:

- Transfer Function Approach, often referred to the Laplace Transform Approach. (linear systems only)

- Similarity Transform Method (linear and nonlinear systems).

- Taylor Series, (linear and nonlinear)

- Differential Algebra, (linear and nonlinear).

Approaches based on the latter two methods (Taylor series and Differential Algebra) are used in this work due to their applicability to both linear and nonlinear models. In addition, the methods are either simple to implement in standard computer algebra systems, or have reliable implementations freely available.

Taylor Series approach

The Taylor series approach [Pohjanpalo, 1978] to identifiability relies on the Taylor series expansion of the observation/output of the system. The output ($y(t, p)$, from equation 3.3) can be expanded as a Taylor series about a known time point (e.g. $t = 0$):

$$y(t,p) = y(0,p) + y^{(1)}(0,p)t + ... + y^{(n)}(0,p)\frac{t^n}{n!} + ... \tag{3.4}$$

where

$$y^{(n)}(0,p) \quad = \tfrac{d^n}{dt^n}y(0,p). \tag{3.5}$$

As the coefficients of the Taylor series are, in principle measurable and unique, given two parameter vectors p and $\bar{p}$ then if $y(t,p) = y(t,\bar{p})$ it must be the case that

$$\begin{aligned} y^n(0,p) = \quad & y^n(0,\bar{p}) \quad n \geq 0 \\ y^{(0)} = \quad & y. \end{aligned} \tag{3.6}$$

Thus yielding a set of simultaneous equations which can be solved for each of the parameters p_i in the vector p. The number of solutions to these equations determine the categorisation of the system according to definition 3.2.1.

Difficulties arise with the Taylor series approach when determining how many coefficients of the expansion are necessary to determine the identifiability of the parameters. Upper bounds have been found for specific system [Chappell et al., 1999] (where n is the dimension of the state vector q):

- $2n - 1$, for linear systems.
- $2^{2n} - 1$, for bilinear systems
- $(q^{2n} - 1)/(q - 1)$, for nonlinear systems of homogeneous polynomial form (q is the degree of the polynomial).

However, in general an upper limit is not available. Due to this fact the Taylor series method can prove difficult in determining if a system is unidentifiable. Calculation of the derivative terms in equation (3.4) becomes increasingly difficult, it is therefore sensible to use computer algebra systems, e.g. MATHEMATICA [Wolfram, 1999] and MAPLE [Heal et al., 1998], in their calculation.

Differential Algebra methods

Differential Algebra techniques [Diop and Fliess, 1991; Margaria et al., 2001; Saccomani et al., 2001] attempt to redefine the system equations, through algebraic manipulation

into an input-output relationship, e.g. a function depending only on the known output and input functions. Through elimination theory, presented originally in work on differential algebra by Ritt [1950], a characteristic set is generated, which is a finite set of differential equations with the same solution set as the original system, from which the identifiability of the parameter vector (p) can be determined [Saccomani et al., 2002]. To generate the characteristic set, a set of differential polynomials is formed directly from the system equations:

$$\begin{aligned} \dot{q}(p) &- f(q,u,p), \\ y &- h(q,p). \end{aligned} \tag{3.7}$$

Then given a ranking, that specifies the elimination precedents of the indeterminates, it is possible from this differential ideal generated by the differential polynomials to derive an input-output map ($g(.)$) for the system can be created. This map will only contain terms of the measured variables and known inputs, associated derivatives and the parameter vector (p); thus eliminating the unknown states ($q_i(t)$). As with the Taylor series approach, given perfect data the derivatives of $y(t)$ and $u(t)$ are in theory calculable [Margaria et al., 2001]. It is therefore possible to set up a replica input-output function instantiated with an alternative parameterisation ($\bar{p}$), e.g.

$$g(y,...,y^n,u,...,u^n,p) = g(y,...,y^n,u,...,u^n,\bar{p}), \tag{3.8}$$

if these functions are linearly independent their co-efficients can be equated to determine the parameter identifiability in accordance with definition 3.2.1.

Implementation of the Differential Algebra method is a non-trivial computational mathematical problem; fortunately it is freely available in two software packages: Bellu et al. [2007] have created a software module, called Daisy, that runs in the REDUCE [Hearn, 1995] algebraic tool; an alternative implementation, DiffAlg [Hubert, 2005], is available as a module for MAPLE [Heal et al., 1998]. It must be noted that the elimination procedure is a computationally expensive procedure and as such both of these software tools can encounter difficulties when generating the characteristic set from the system equations, and with complex systems may fail to return a result.

3.2.2 Numerical sensitivity

Once a model has been confirmed to be S.G.I (or restricted to a unique subset in the case of S.L.I parameters) it must be determined whether the system output is sensitive to perturbations of the parameter set (at a nominal value) in order to influence the output. If the parameter does not impact on the output significantly parameter estimates will not be found reliably during the parameter estimation procedure (see section 3.2.3). It is therefore important to determine the system sensitivity to each parameter to be estimated.

If an analytical solution to the system of differential equations is available then it may be possible to obtain an expression to determine the amount the solution changes with respect to the parameter directly; unfortunately, this is rarely possible in complex systems. An alternative approach is to perform the sensitivity analysis manually by modifying a singe parameter from its initial value, performing a simulation and comparing changes in the simulated output and rejecting those parameters that do not have an observable influence. Depending on the model complexity and the number of parameters to be estimated, this can be an error-prone and time-consuming process. A automated numerical approach is suggested by Vajda et al. [1985].

In order to determine the sensitivity of the output to the input parameters, a matrix is created of changes in the normalised output values to changes in the parameters at each observation time point. For a multi-output system at n time points,

$$S = [S_1 \ ... \ S_n]^T \tag{3.9}$$

with the submatrices (S_i) describing the effects at time point i defined as

$$S_i = \frac{\partial \, ln(y_{ji})}{\partial \, ln(p_k)} \tag{3.10}$$

where y_{ji} is the observed compartment j at time i and p_k the k^{th} parameter of parameter vector p. Natural logs are used to non-dimensionalise the values and allow comparison of parameters of different units. The partial derivatives of the S matrix can be generated from finite differencing methods [Chapra and Canale, 2002] or through the application of the chain-rule to the above system to define the 'forward sensitivity equations' [Tomovic and Vukobratovic, 1972] which can be solved interactively during

the simulation of the system.

Principal component analysis is then performed on the square of the sensitivity matrix, which approximates the sum of squares response of the output to changes in the parameters [Vajda et al., 1985]. If singular value decomposition is used the matrix can be deconstructed into orthogonal components

$$V = S^T S = U \Sigma U^T \tag{3.11}$$

where U is a matrix of orthonormal eigenvectors and Σ contains the 'singular values' of the matrix. The square of the singular values are equal to the eigenvalues of V. The magnitude of the eigenvalues relate to the influence a parameter has on the principal axis, or ultimately, changes in the output response $y(t)$. Vajda et al. [1985] suggest parameters with eigenvector components in U with magnitudes less > 0.2 should be considered for rejection from the parameter estimation procedure, as this relates to less than 4% change in the sum of squares of the output. This can be implemented in standard computer tools, however it is handled automatically in the parameter estimation software used for the work presented in this thesis.

3.2.3 Parameter estimation

Parameter estimation of compartmental models is commonly performed by a least-squares process, using a distance metric (χ^2 in the example below) between the simulated output and the observed, or measured, response, with respect to the estimated parameter values. In a multi output system this can be represented as

$$p_0 = \arg\min_p \chi^2(p), \tag{3.12}$$

$$\chi^2(p) \equiv \sum_j \sum_i \left(\frac{\hat{y}_{ji} - y_j(t_i, p)}{\sigma_j} \right)^2, \tag{3.13}$$

where p is the vector of parameter to be fitted, p_0 the values at the minimum, y_j is the simulated output of compartment j at time t_i, $\hat{y}_{ji}$ the measured output of appropriate compartment and time samples point, and σ_j the standard error in the measurement. When dealing with *in vivo* measurements the observations available will be a noisy,

sparse, unevenly sampled time series of measurements from patients undergoing treatment.

3.2.4 Parameter estimation tools

It is possible to develop the above validation procedures in customised software; however, there are a plethora of pre-built tools designed specifically for this purpose. Two such tools that are predominately aimed at compartmental modelling are: Berkeley Madonna [Oster, 2009] and Facsimile [M.C.P.A, 2009]. Berkeley Madonna is a useful prototyping tool, allowing quick model construction and feasibility studies to be performed. Facsimile is a robust modelling environment with excellent parameter estimation facilities. Facsimile performs statistical analysis during the parameter fitting process and will reject parameters in accordance with their sensitivity (as defined in section 3.2.2), ensuring numerical validity in the choice of parameters and initial values. Unless otherwise stated, any parameter estimates given in this work were obtained using Facsimile. The variance of parameter estimates are given as the standard deviation of the log of the parameter (SDLN), this can viewed as an approximation to the coefficient of variation (CoV) for the parameter estimated [Bland and Altman, 1996].

3.3 Summary

Compartmental modelling provides a framework under which physiological and kinetic models can be described and analysed. In addition, due to the depth of research in this field, techniques to assist in model validation are well known, documented and often pre-built into available modelling tools. The approach identified above is used on all clinical models developed in later chapters. For each model three key stages are addressed: firstly, identification of potential compartments, inputs, outputs and transfer co-efficients from reviewing the available literature or discussion with clinical experts; secondly, the model and experiment is validated using identifiability analysis to ensure the parameters required can theoretically be determined by the measurements; finally, parameter estimation with sensitivity analysis is employed to obtain appropriate parameter values. Once this is complete the model can be used to investigate alternative

treatment scenarios, or investigate unmeasured patient responses.

In the next chapter compartmental models, and mathematical analysis, are used to investigate the *in vivo* kinetics of an immunoglobulin sub-complex, namely Free Light Chains, in patients with Multiple Myeloma. This investigation allows information to be obtained regarding the underlying disease, the effectiveness of treatment and the possibility of improving the patient's condition through alternative therapies or medication.

Chapter 4

Free-Light Chain clearance through haemodialysis

Haemodialysis is the removal of substances from the blood through an extracorporeal semi-permeable membrane. The first dialyser, or 'artificial kidney', was developed in 1913 [Abel et al., 1914], this was a proto-type system tested on animals with reduced renal function. It was not until 1943 that haemodialysis was used in a clinical setting on patients with kidney failure, initially with limited success [Kolff, 2002]. However, its use is now ubiquitous with regard to patients with renal failure and kidney damage.

Haemodialysis is conventionally used to remove waste products (e.g. urea, creatinine, potassium) from a patient's blood and ensure appropriate fluid levels are maintained. In order to determine efficacy of treatment and predict the required dose, mathematical models have been developed. These models focus upon urea, which is a waste product that is created when the body metabolises proteins. It is passed by kidneys and cleared by the dialysis and is an indicator of renal and dialyser performance [Gotch, 2001; Sargent and Gotch, 1996]. Although urea is not toxic it is used as a marker for dialysis efficiency as it is simple to measure, high in concentration in patients with reduced kidney function and has been linked, statistically, with patients' outcome in terms of dialysis dose [Tattersall, 2005].

The first mathematical models to describe urea haemodialysis were single pool models, in which a single compartment represented the total urea in the body, with first order removal representing dialysis clearance and constant (level) input for urea

production [Gotch, 2001]. Calculations and estimates from the single model are routinely used during clinical treatment; however, it has been found that with increasingly effective clearance mechanisms the single pool kinetics are insufficient and can result in serious over-estimation of dialyser efficiency and patient fluid levels [Daugirdas et al., 1999; Vanholder et al., 1996]. The errors present are due to the ability of urea to pass through the cell membrane, the kinetics of which are slower than that of dialyser clearance, resulting in disequilibrium of concentration across the cell wall. This inbalance is quickly restored when the dialysis treatment stops (a process known as rebound); this leads to an over-estimation of clearance and under estimation of the total toxins remaining in the patient. To account for this, models have been developed with two compartments (see Figure 4.1): the first representing the urea within plasma; the second, urea concentrations within cells [Sargent and Gotch, 1996]. The kinetic exchange between these two compartments is governed by the exchange of urea across the cell wall.

Although the models defined for urea are useful indicators of the mechanism that should be considered when modelling dialysis, they cannot be used directly for the analysis of Free light chains (FLC); urea is a small molecule, with a molecular weight (mw) of 60 Da, whereas FLC is considerably larger (mw 24-48 kDa) [Fielding, 2002, p. 184], and as such it will not be removed as readily by standard dialysis filters. β_2-microglobulin (B2M) is a middle-molecule of similar dimensions to FLC (mw 11.8 kDa) that has been modelled during dialysis [Gotch et al., 1989; Leypoldt, 2005; Leypoldt et al., 1997; Ward et al., 2006]. However, as it was found that high levels of B2M build-up during dialysis are highly correlated to patient morbidity during treatment [Winchester et al., 2003] mathematical models have been developed to simulate middle-molecule removal in haemodialysis.

As with urea, single-pool models were considered to describe B2M but it has been found [Leypoldt et al., 1997], through measurement of B2M in the blood and dialysate fluids across the dialyser, that a single pool model was inadequate in predicting dialyser clearance and once again a two-compartment model produced more accurate results. An example of a two compartment B2M dialysis model can be seen in Figure 4.1. The two pools are given the names of perfuse (when a region of the body has direct access to a blood supply) or non-perfuse (when the blood supply is not accessible from a region).

Unfortunately there are contradictions as to what these pools represent, anatomically. Leypoldt et al. [2003] refer to them as intracellular (perfuse) and extracellular (non-perfuse) while others [Gotch and Keen, 2005; Ward et al., 2006] describe the volumes of distribution as plasma and interstitial fluid. This distinction is important as it determines the transport mechanism defined by the intercompartmental transfer: if the boundary is intra/extracellular it is the cell membrane, whilst plasma/interstitial fluid would denote transfer across the plasma membrane. This confusion leads to differences in the input function (i.e. generation rate of B2M) and whether the B2M is created directly into the perfuse, non-perfuse or into both compartments. Further, this B2M model ignores renal clearance which if present, even when severally limited in function, could greatly affect the reported efficiency of the dialysis clearance.

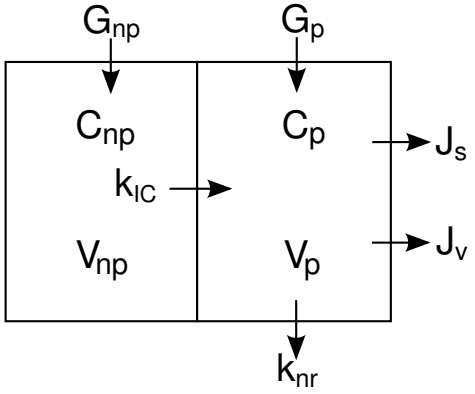


Figure 4.1: *Schematic representation of Gotch β_2-microglobulin model [Gotch et al., 1989]. Reproduced from Leypoldt et al. [2003] with permission of the copyright owner. P - Perfuse pool, NP - non-perfuse pool, C - B2M concentration, V - apparent volume of the pool, J_v - volume removal from dialysis, J_s - solute removal from dialysis, K_{IC} - intercompartmental transfer, K_{NR} - non-renal clearance and G the generation rate of B2M.*

Although B2M and FLC are of comparable sizes they are believed to have different *in vivo* kinetics [Waldman and Strober, 1969] and as such require new models to be constructed; however, the structure of the B2M models can be considered as a starting point from which to build upon.

FLC offers clinicians different treatment options in addition to dialysis, namely the use of chemotherapy treatments. Therefore in the first section of this chapter the effects of chemotherapy on FLC concentration without the influence of haemodialysis

are investigated; after which an FLC model will be described and its application to comparison of therapeutic regimes to determine the treatment's effectiveness in the presence of Free Light-Chain secreting tumours. This model is then used *de novo*, to determine the impact of using different apheresis and chemotherapy regimes on a single patient, and ultimately to assess the adequacy and benefits of the treatment.

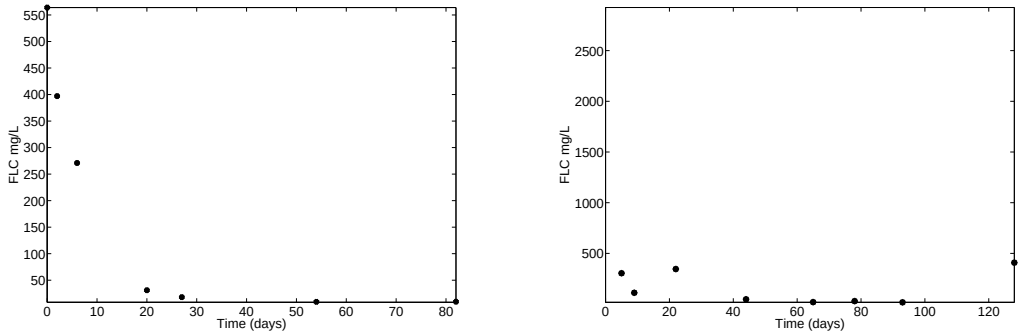
4.1 Patient response to chemotherapy

As previously stated in Chapter 2 a range of treatments are available for patients with Multiple Myeloma (MM). Whilst it is possible to treat some patients with apheresis, this is an experimental therapy and as such is only used in the clinical trial described in detail in Section 4.2. The standard procedure for MM patients, irrespective of the type (e.g. FLC, Intact secreting) is to use combined chemotherapy. In Mead et al. [2006] the first few cycles of chemotherapy, often referred to as induction chemotherapy, were monitored. Whilst all patients had tumours producing intact immunoglobulin, only FLC were observed due to their shorter half-life; as such they are a better indicator of the effects of chemotherapy than measuring whole immunoglobulin. Data from patients receiving three different chemotherapy drug combinations were monitored:

- Vincristine, Doxorubicin plus Dexamethasone (VAD): eight patients.
- Melphalan and Prednisone (MP): five patients.
- Thalidomide plus Dexamethasone (CTD): five patients.

Patients were monitored over an extended period to observe their response to the treatments. Measurements of FLC in plasma were taken periodically at between 0, 4, 8 and 21 months; occasionally, if the patient was re-admitted to hospital between these key periods further measurements were often available. From this minimal set of measurements, clinicians wished to determine the relationship between a particular chemotherapy treatment and tumour kill, and ultimately, patient health. An example of the data available can be seen in Figure 4.2.

The patient's clinical response is analysed by the treating physician and is a medical assessment of the patient's conditions. Whilst this is an unquantified measurement it



(a) Dataset 1 (b) Dataset 2

Figure 4.2: *Typical data available for the chemotherapy treatment comparison.*

is an indicator of the effect of the treatment, and in general it would be expected that analytical results should correlate well with this assessment. Clinical response is estimated through various non-invasive measurements, for example, observation of heart-rate, oxygenation, mobility, patient responsiveness, which are incorporated by clinicians in a general diagnosis. This is discussed further in Section 4.1.2.

4.1.1 Single compartment model

To assist in this analysis a mathematical model is beneficial to determine tumour kill and baseline processes, such as natural clearance and production of FLC. Due to the period and availability of observations, a simple one compartment model is used. This can be described by a single ordinary differential equation

$$\begin{aligned} \dot{c}(t) &= -k_0\, c(t) + f \\ c(0) &= c_0 \end{aligned} \tag{4.1}$$

where $c(t)$ represents the concentration of FLC within the patient, k_0 the natural clearance of FLC and f the synthesis of healthy FLC, which is assumed to be constant [Bradwell et al., 2005, ch. 3].

A clinical measure that is often used in assessing the treatment effectiveness is the time taken for the plasma concentration to reach half the initial concentration. This is an indicator of both effectiveness and the speed of response of the chemotherapy. However, with the minimal measurements normally available it is difficult to estimate

the point at which the biological half-life has occurred. Fortunately, with this simple linear model, an analytical solution can be obtained from which the apparent half-life ($\tau_{\frac{1}{2}}$) of FLC in the patient can be estimated:

$$\tau_{\frac{1}{2}} = \frac{1}{k_0} \ln\left(\frac{2(f - k_0\ c_0)}{2f - k_0\ c_0}\right) \tag{4.2}$$

The vector of parameters to be estimated is $p = (c(0), k_0, f)^T$. Using the Taylor series method (as described in Section 3.2.1), it can be shown that all parameters are uniquely determined from the observation of concentration of FLC in the patient ($c(t)$). Due to the simplicity of the model the identifiability analysis can be performed by hand. Given the model defined above (eq. 4.1), the first three Taylor series coefficients of the observation are given by:

$$\begin{aligned} y(t,p) &= c(t) \\ y(0,p) &= c(0) \\ y'(0,p) &= -k_0\ c(0) + f \\ y''(0,p) &= -k_0\ (-k_0\ c(0) + f) \end{aligned} \tag{4.3}$$

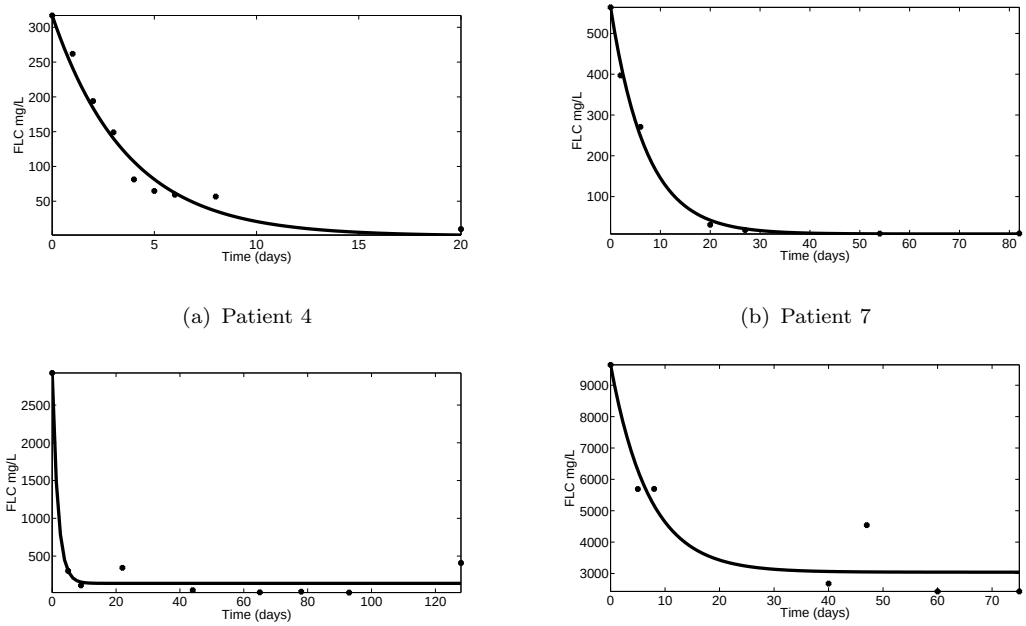
An alternative vector ($\phi = (\phi_1, \phi_2, \phi_3)^T$) of observable parameter combinations can be generated directly from these equations:

$$\begin{aligned} \phi_1 &= c(0) \\ \phi_2 &= -k_0\ \phi_1 + f \\ \phi_3 &= -k_0\ \phi_2 \end{aligned} \tag{4.4}$$

The initial conditions$c(0)$ is clearly identifiable from ϕ_1, and the clearance k_0 is identifiable from ϕ_3. Finally, it is evident from the ϕ_2, that if k_0 and $c(0)$ are globally identifiable f is also. In subsequent identifiability analyses only the results will be presented; however, the analysis and code listings, when symbolic packages have been used, can be found in Appendix A.

4.1.2 Results

The model (eqn. 4.1) was fitted to the data gathered for all eighteen patients. Examples of the simulation results can be seen in Figure 4.3; for brevity only four patients are shown (4, 7, 10 and 11); the other patient data and fits displayed similar behaviour.



(a) Patient 4

(b) Patient 7

(c) Patient 10

(d) Patient 11

Figure 4.3: *Estimated results for four induction chemotherapy patients.*

FACSIMILE was used to estimate the natural clearance k_0, baseline synthesis reached (f) and the initial concentration of FLC ($c(0)$) when the patient was presented for treatment. It should be noted that although the initial concentration is measured, to assume that it is equal to the measured value would neglect the possible error present in the measurement. Unsurprisingly, in all cases the estimated parameter was very close to the measured value. The results of the parameter estimation can be seen in Table 4.1.

The model described and results obtained allow clinicians to easily compare the different chemotherapy drug combinations. The results of the study are succinctly stated by Mead et al. [2006]:

> *"Patients showed slower FLC reductions with MP treatment than with regimens containing Dexamethasone. The extent of SFLC [Serum Free Light-*

Table 4.1: *Estimates for patient parameters from single pool model, with associated standard deviation of the natural log (SDLN). To aid in interpretation, the SDLN (shown in brackets) can be viewed as an approximation to the coefficient of variation for the parameter [Bland and Altman, 1996]. NWD - not well determined, indicates the parameter could not be estimated from the given model against the supplied data. N/A - not applicable as the patient did not reach the desired 50% reduction in FLC concentration within 21 months.*

Patient	$\tau_{\frac{1}{2}}$ (days)	f (mg L^{-1} day^{-1})	k_{01} (day^{-1})	$c(0)$ (mg L^{-1})	Drug Type
1	0.16	1.68E+02(0.09)	8.21E+00(NWD)	6.20E+01(0.12)	VAD
2	25.18	6.46E+00(0.33)	4.76E-02(0.41)	4.77E+02(0.26)	MP
3	0.33	7.14E+01(0.26)	4.27E+00(NWD)	4.84E+01(0.44)	VAD
4	2.54	4.09E-03(NWD)	2.72E-01(0.06)	3.17E+02(0.36)	CTD
5	1.05	1.26E+03(0.50)	7.50E-01(0.18)	1.97E+04(0.05)	CTD
6	4.06	3.56E+00(0.75)	1.97E-01(0.73)	1.93E+02(0.17)	VAD
7	5.05	1.17E+00(0.46)	1.40E-01(0.11)	5.64E+02(0.08)	VAD
8	15.10	1.01E+00(0.66)	7.54E-02(0.44)	5.08E+01(0.04)	VAD
9	3.97	2.01E+01(NWD)	1.11E+00(0.11)	3.68E+01(0.43)	VAD
10	1.31	7.86E+01(0.32)	5.66E-01(0.41)	2.93E+03(0.25)	VAD
11	9.16	4.32E+02(0.45)	1.42E-01(0.37)	9.65E+03(0.13)	CTD
12	N/A	8.16E+00(0.68)	2.24E-01(0.68)	4.53E+01(0.28)	CTD
13	N/A	1.45E+00(0.35)	1.70E-02(0.03)	1.57E+02(0.22)	MP
14	N/A	1.59E+03(NWD)	1.67E+00(0.08)	1.71E+03(0.12)	MP
15	N/A	1.93E+04(0.15)	8.08E+00(NWD)	3.50E+03(0.09)	CTD
16	N/A	4.74E-01(0.31)	3.47E-02(0.17)	1.69E+01(0.15)	VAD
17	N/A	9.61E+02(0.30)	2.95E-01(0.37)	5.50E+03(0.10)	MP
18	N/A	4.12E+00(0.05)	1.63E-01(0.47)	1.97E+01(0.01)	MP

Chains.] response was generally predictive of the final clinical assessment at 3 months and this may help to separate non responders from good responders as early as within 1 month of starting therapy. This may represent an important observation in the treatment of Myeloma and is now being assessed in a larger cohort study."

This can be clearly seen in Table 4.1, with only a single Melphalan and Prednisone patient achieving a 50% reduction in serum FLC in the twenty one month observation period; the time taken for this patient is at least an order of magnitude greater than the majority of patients reaching a half-life reduction. From these results it is unlikely that this would be down to either high initial concentrations in plasma, or low clearance rates, as greater extremes are seen in several patients undergoing VAD and CTD treatments. However, whilst Dexamethasone is clearly beneficial for these patients, it is unclear from the results which combination would be preferable, with Thalidomide or Vincristine and Doxorubicin showing similar outcomes.

4.1.3 Discussion of chemotherapy results

The above method shows how even a simple model can be extremely beneficial in categorising patient response, allowing analysis and comparison of drug treatments. To the author's knowledge this is the first time the chemotherapy treatments for FLC have been analysed with respect to patients' response; prior to this the treatment used was driven solely by previous clinical experience. The above process, whilst simple to understand is by no means ideal and has several flaws which should be addressed if it is to be considered for further treatment studies. The model relies on several key assumptions:

- **Instant, and consistent, chemotherapy effects**

  It is assumed that the chemotherapy has an instant effect, reducing the tumourous FLC synthesis to a baseline level that is then maintained over the treatment period. Whilst over the timescales of months this may be the case, it is not uncommon for patients to receive repeated chemotherapy due to a 'relapse' in tumour synthesis. An example of this can be seen in Figure 4.3(d). The fifth measurement

taken (at $\approx$ 47 days) shows a marked increase over the previous measurement, which is maintained in subsequent measurements. This is due to the patient having an additional set of chemotherapy treatments during this period, indicating that an exponential decay to a baseline can be an oversimplification.

- **Constant clearance**

  The clearance of FLC in the above model is assumed to be constant. Whilst it is believed that natural clearance is linear, apheresis is used in order to augment this catabolism. This is evident from some of the results obtained; with reference to Table 4.1, patients 1, 3 and 15 all have high clearance rates, an order of magnitude greater than the mean value[1] (mean of $k_0 = 0.22$ day^{-1}). This is due to these patients having apheresis, specifically plasma exchange, during early stages of the treatment regime, resulting in a clearance that is dependent on time, i.e. it is at natural clearance when the patient was not receiving plasma exchange but elevated to an increased value when plasma exchange was performed.

- **Single pool kinetics**

  For the model described it is assumed that the kinetics are described adequately by a single pool, which is equally impacted by both clearance and synthesis in equal measure. Whilst this may be appropriate over the time periods considered in the above treatment, it is known that FLC travel freely between the vascular and intravascular spaces within the body via the plasma membrane [Waldman and Strober, 1969]. As such a two-compartment model, as seen for B2M models, would be more applicable when analysising the kinetics of treatment.

The validity of the assumptions noted is vital to the accuracy of the analysis conducted. A patient response that does not adhere to the assumptions could easily allow misclassification of the patient's conditions with regard to long term recovery. Therefore, by more realistic modelling of the kinetics and treatment an improved classification may be obtained. In addition, features of the treatment may be revealed that were not apparent in the simple model described by equation (4.1). As such, the effectiveness of chemotherapy and its impact upon the underlying antibody synthesis will be discussed

[1]The mean value excluding parameters sets not well determined by the data.

further in Chapters 5 and 7. The remainder of this chapter will focus on adaptations of this model to deal with non-constant clearance, due to clinical intervention in the form of apheresis.

4.2 Two compartment FLC model

As seen in the previous section, the standard treatment for patients with Multiple Myeloma is to use chemotherapy drug combinations to 'kill' the tumour, relying on the patient's natural clearance to reduce the dangerously high concentrations of FLC in plasma. However, as described in Chapter 2, the majority of clearance of FLC is through the renal path, which for most FLC Multiple Myeloma patients has been severely compromised, and is usually reduced to zero, due the destruction of nephrons from FLC cysts [Hutchison et al., 2007]. In order to increase clearance, plasma exchange has been proposed, but after several studies this continues to be a controversial treatment [Leung et al., 2008]. Recently, an alternative treatment using highly porous, or 'protein leaking', dialysis filters has been tested [Bradwell et al., 2005]. These are filters with pores of sufficient dimensions to allow FLC proteins to be more readily cleared through the membrane. As patients with high-levels of FLC are already receiving haemodialysis to clear urea the use of dialysis filters to also remove FLC has the added benefit of working in conjunction with existing treatments. In a study by Evans et al. [2006] a two-compartment model was created and using parameter values gained from *in vitro* measurements, simulations were conducted to initially evaluate the individual effects of plasma exchange and haemodialysis for FLC patients. A schematic for this model can been seen in Figure 4.4.

As in the B2M model, Figure 4.1, the FLC model contains two compartments; rather than the abstract perfuse and non-perfuse separation of pools, the compartments represent more anatomically recognisable regions. The first compartment represents FLC in the plasma (q_1); the second, for FLC that exist outside of the plasma space, an area described as extravascular fluid. The rate constant k_{21} describes the flow of FLC across the capillary walls from plasma to extravascular fluid (EVF); k_{12} is the flow in the reverse direction. From equilibrium conditions k_{12} can be derived from k_{21} and compartmental volume ratios ($k_{12} = (v_1/v_2)k_{21}$) [Jacquez, 1996]. The rate

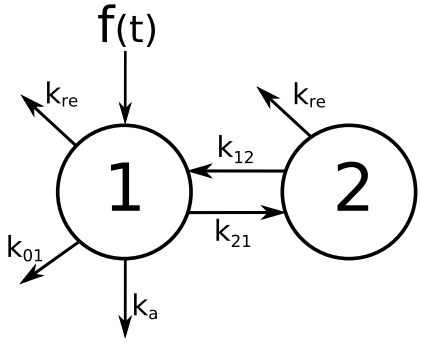


Figure 4.4: *The FLC model [Evans et al., 2006]. Compartment 1 - FLC in plasma; compartment 2 - FLC in extravascular fluid; k_a - aphereis clearance; k_{re} - reticuloendothelial clearance; k_{01} natural renal clearance; k_{12} and k_{21} exchange across the plasma membrane; $f(t)$ synthesis of FLC.*

constant k_{re} corresponds to the rate of the removal of FLC from plasma and EVF via the reticuloendothelial system (see Chapter 2). It is calculated from the half-life values of FLC when the patient has no natural renal clearance. k_{01} estimates the rate of FLC removal via the urinary tract. It is provided here for completeness as in all patients considered, the renal function is negligible and therefore under these circumstances $k_{01} = 0$. k_a is the rate of removal of FLC via apheresis, which is controlled by the time-varying function

$$k_a(t) = \begin{cases} 0 & t \notin [\tau_{si}, \tau_{ei}] \\ k_a & t \in [\tau_{si}, \tau_{ei}] \end{cases} . \tag{4.5}$$

τ_{si} and τ_{ei} are the start and end times of the i^{th} apheresis treatment. A description of the model parameters and related units can be found in Table 4.2.

By using mass-balance principles a system of equations can be derived from the compartmental schematic in Figure 4.4:

$$\begin{aligned} \dot{q_1}(t) &= -(k_{21} + k_{re} + k_{01} + k_a)q_1(t) + k_{12}q_2(t) + f(t) \\ \dot{q_2}(t) &= k_{21}q_1(t) - (k_{re} + k_{12})q_2(t) \\ q1(0) &= q_{10} \\ q2(0) &= q_{20} \end{aligned} \tag{4.6}$$

Table 4.2: *FLC model parameter description.*

Parameter	Description (Units)
$q_1(t)$	Quantity of FLC in plasma (mg)
$q_2(t)$	Quantity of in EVF (mg)
$c_1(t)$	Concentration of FLC in plasma (mg/L)
$c_2(t)$	Concentration of FLC in EVF (mg/L)
$f(t)$	FLC synthesis (mg/min).
k_{21}	Transfer rate constant from plasma to EVF (min^{-1})
k_{12}	Transfer rate constant from EVF to plasma (min^{-1})
k_{re}	Reticuloendothelial clearance (min^{-1})
k_{01}	Normal renal clearance of FLC-Kappa (min^{-1})
v_1	Apparent plasma volume (L)
v_2	Apparent EVF volume (L)
k_a	Apheresis clearance rate constant (min^{-1})
k_d	Haemodialysis clearance rate constant (min^{-1})
k_p	Plasma exchange clearance rate constant (min^{-1})

where $q_1(t)$ and $q_2(t)$ are the quantities of FLC in plasma and EVF respectively and $f(t)$ is the generation of new FLC proteins; all other rate constants are as defined previously (see Table 4.2).

During treatment clinicians are able to take periodic blood samples from the patient, from which the concentration (mg/L) of FLC can be measured using immunoassays. If it is assumed that the assay is able to measure 100% of the FLC in the sample, the following output structure can then be assumed:

$$y_1(t) = \frac{q_1(t)}{v_1} \tag{4.7}$$

In order to solve equations (4.6) an initial value problem can be constructed if the initial conditions of the compartments are known. As it is impossible to measure the concentration of the FLC in EVF fluid assumptions must be made to overcome this. As with all biological processes the generation of FLC via the tumourous plasma cells is bounded. Due to the nature of cancerous cells, growth is prohibited by the availability

of resource; in the case of Multiple Myeloma this is believed to be predominantly down to limited space for the cancerous cells to grow within the bone-marrow [Bataille, 1996]. Patients with Multiple Myeloma are difficult to diagnose, therefore the disease is usually in a state of equilibrium, where the synthesis has reached a constant maximum value; therefore, without the initiation of chemotherapy or apheresis the concentration across the plasma membrane should have had sufficient time to equilibrate. As such the initial concentrations in plasma and the EVF space can be considered equal. Thus the initial conditions for equations (4.6) are therefore

$$\begin{aligned} q_1(0) &= c_1(0)v_1 \\ q_2(0) &= c_2(0)v_2 \end{aligned} \tag{4.8}$$

The steady-state assumption prior to treatment was validated against available patient data. Prior to treatment plasma samples were taken of several patients and the FLC concentration calculated. Over a 3-4 day period it was seen that the concentration levels in plasma remained constant, to within $\pm 10\%$ which is within the measurement error.

Identifiability analysis

It can be shown (see Appendix A.1.1) that given two unknown parameter vectors, p_1 when the patient is off dialysis and p_2 during dialysis,

$$\begin{aligned} p_1 &= (k_{re}, k_{21}, q_1(0), q_2(0), f, v_1, v_2)^T, \\ p_2 &= (k_a, k_{re}, k_{21}, q_1(0), q_2(0), f, v_1, v_2)^T, \end{aligned} \tag{4.9}$$

the above model (eqn. 4.6) is SGI only if the parameter k_{re} and EVF volume are known. As in the model description the rate of reticuloendothelial clearance can be approximated from the half-life and the EVF volume can be estimated from the plasma volume of the patient, $v_2 \approx 3\ v_1$ [Malesker and Morrow, 2007], resulting in a globally identifiable parameter vector $p_f = (k_a, k_{21}, q_{10}, q_{20}, f, v_1)^T$.

If plasma samples are taken during the dialytic period only the parameters v_1 and $q_2(0)$ become unidentifiable; however, the plasma volume v_1 can be estimated from the

patient's weight (v_1 =8% of body weight in kilograms [Malesker and Morrow, 2007]) and if it is assumed the patient is in steady-state prior to treatment ($t < 0$) the quantity of FLC in the EVF compartment and the synthesis rate can be calculated from

$$\begin{aligned} q_2(0) &= \frac{k_{21}q_1(0)}{(k_{re}+k_{12})}, \\ f(0) &= \frac{(k_{21}(k_a+k_{re})+k_{re}(k_{12}+k_a+k_{re}))q_1(0)}{(k_{re}+k_{12})}. \end{aligned} \tag{4.10}$$

4.2.1 Apheresis analysis

Once the model has been validated and parameter estimates obtained for the model, it is possible to investigate clinical questions regarding optimal treatment, for example:

- Which treatment modality (plasma exchange or haemodialysis) is 'best' for FLC patients?
- Which of the available dialysis filters is best at clearing FLC?
- Is there a difference in the removal of FLC-Kappa or FLC-Lambda variants of Free Light Chains?
- What is the impact of apheresis in conjunction with chemotherapy?
- Does the recovery of natural renal clearance during treatment significantly improve recovery?

In order to answer the above questions the FLC model must first be initialised with appropriate values for the model parameters. Fortunately, parameter estimates for the FLC model were obtained through a study on the comparison of different dialysis filters; this is described in the following section. The parameters obtained are then used in a series of simulated experiments to answer the question listed above.

4.2.2 Comparison of filter manufacturers and FLC type

Dialysis filters are traditionally used for the removal of small molecular toxins (< 1000 Da) from a patient, e.g. urea. However, only three high-flux membranes with

sufficient pore size to remove proteins the size of FLC (> 25 kDa) are currently available: BKF (Toray Industries Inc, Tokyo, Japan), HCO (Gambro Dialysatoren GmbH, Hechingen, Germany) and Hi-Pres 18, (B. Braun Medical Ltd, Sheffield, UK). Although each membrane has similar physical characteristics (e.g. pore size, average number of pores, membrane thickness and length) due to the complexity of manufacture and design it cannot be assumed that the filters will all behave identically in a treatment scenario. It should be noted that manufacturers do perform *in vivo* tests to determine filter clearance capabilities, but these are not yet available for FLC, only for more common proteins e.g. urea, albumin [Toray, 2006].

To determine the filter clearance, four patients were tested with the three available membranes; three patients (DM, MF and VF) had the Kappa FLC Myeloma, the other had the Lambda variant. During the dialysis session plasma samples were taken to allow measurement of FLC concentration; in addition, one patient, due to complications, was monitored for a period post-dialysis, giving two off-dialysis measurements. Example data sets can be seen in Figure 4.5.

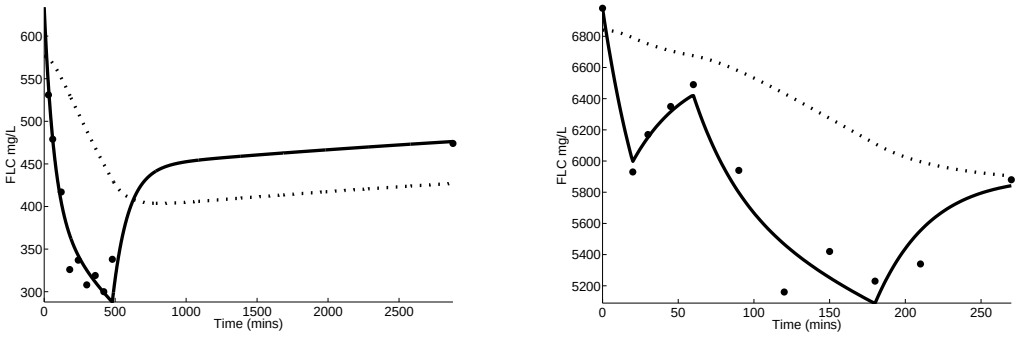
Traditionally only pre- and post-dialysis measurements are taken and the percentage change in plasma concentrations compared as an indicator of total clearance and filter efficiency. The results of comparing the percentage changing plasma concentration can be seen in Table 4.3. In this experiment FLC assays were used to measure the concentration in plasma samples; the Coefficient of Variation (CV) has been estimated to be 10% [Herzum et al., 2005]. With the exception of the Hi-Pres filters and FLC-Lambda patient, in which the clearance was poor (a drop of only 7%), the remaining combinations of filter and MM appear approximately equivalent. Given the magnitude of the error estimates it is difficult to differentiate between the remaining filters and MM types.

The FLC model was used to evaluate the clearance, inter-compartmental transfer and synthesis of FLC in the same patients. The model was fitted to the time series data provided to estimate the unknown parameters. Initial guesses for the parameters were obtained from previous work conducted by Evans et al. [2006]. The longitudinal data varied in number and frequency of samples, with between 6 and 10 measurements per patient, generally taken every 30 minutes, with some missing data points. For some patients, off-dialysis data measurements were available, but the simulation was

Table 4.3: *Percentage change in patient FLC levels, pre- and post-dialysis. The estimated propagated error is shown assuming a CV of 10%.*

Patient	MM Type FLC	Filter Type	% Reduction in FLC (% Error)
DM	Kappa	HCO	32.17(12.5)
VF	Kappa	BKF	28.12(12.62)
VF	Kappa	Hi-PeS	23.5(12.8)
MF	Lambda	BKF	36.82(12.39)
MF	Lambda	Hi-PeS	7.3(13.66)

conducted over a full dialysis and interdialytic period. As most patients are on a daily dialysis treatment, this involved a 24 hour simulation. However, in one patient the filter was changed under the belief that is was clogged; as a clean filter was used this results in an extra estimate for the Hi-Pres filter for one of the Kappa Myeloma patients (VF). Two examples of the data set and corresponding simulation can be seen in Figure 4.5.



(a) Patient DM - single six hour dialysis session with two post dialysis measurements.

(b) Patient VF - single three hours dialysis session with a change of filter after thirty minutes incurring a thirty minute period where the patient was measured but not on dialysis; then two post dialysis measurements.

Figure 4.5: *Patient data (dots) and simulation results for plasma (solid-line) and extravascular fluid (dotted-line).*

The results of the parameter fitting can be seen in Table 4.4. A clear distinction can be seen between the clearance for the filters, with the Braun Hi-PeS's performance being much poorer than the Gambro HCO and Toray BKF membranes. In this limited study, there appears to be no significant difference between the HCO and BKF filters

used. This prediction fits in with clinical evaluations conducted regarding patient health and recovery.

Table 4.4: *Parameter estimates for FLC Model with different high-flux filters. The value in brackets is the estimated standard deviation of the natural log of the parameters estimates.*

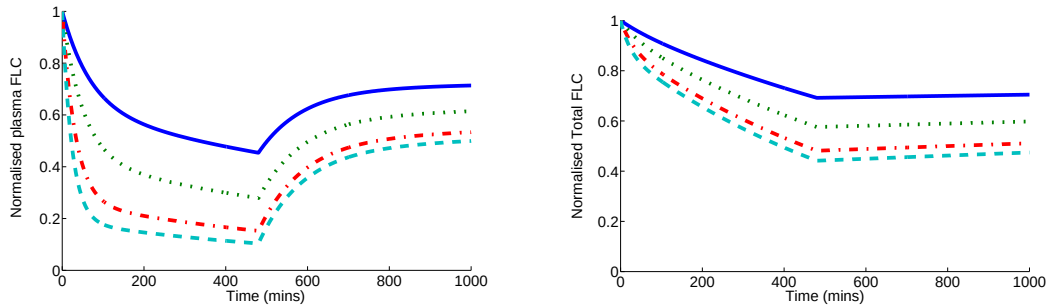
Patient	MM Type	Filter Type	k_{21} min^{-1}	k_a min^{-1}	f mg min^{-1}	$c_1(0)$ g L^{-1}
DM	Kappa	HCO	0.038(0.30)	0.015(0.10)	2.0(0.16)	0.6(0.05)
VF	Kappa	BKF	0.031(0.10)	0.011(0.06)	28.5(0.20)	9.5(0.03)
VF(i)*	Kappa	Hi-PeS	0.022(0.08)	0.009(0.07)	2.3(0.13)	7.5(0.05)
VF(ii)*	Kappa	Hi-PeS	0.022(0.18)	0.005(0.12)	2.3(0.13)	7.5(0.05)
MF	Lambda	BKF	0.021(0.11)	0.015(0.15)	23.3(0.34)	1.7(0.02)
MF	Lambda	Hi-PeS	0.053(0.21)	0.004(0.26)	4.1(0.3)	2.6(0.04)

* During the treatment, two filters were used, therefore clearance (k_a) is unique, but production (f), initial concentration ($c_1(0)$) and inter-compartment transfer (k_{21}) are fitted simultaneously. See Figure 4.5(b) for simulation results.

In addition to allowing the filters to be ranked in order of clearance rate this analysis offers insight into the intercompartmental transfer dynamics during treatment. The estimates for the intercompartmental rate constant (k_{21}) can be seen for each patient in Table 4.4, ranging from 0.022 to 0.053 min^{-1}, showing little intra or inter-patient variability; however, with reference to patients VF and MF using the BKF filter there is some evidence of an increase between the FLC-Kappa and FLC-Lambda intercompartmental transfer rate, as would be expected from the increased size of Lambda.

The inter-compartmental transfer rate is a key factor in the total body clearance during a dialysis session. This effect can be seen in Figures 4.5(a) and 4.5(b); in each the EVF compartment, which is not directly accessible by the dialysis clearance, is reduced, but at a much slower rate due to restriction by the limited transfer rate (k_{21}). The effect of this can also be seen when the patient is taken off dialysis. For a period after dialysis the compartments equilibrate, resulting in a sharp increase in the

plasma concentration due to the large difference in concentration between plasma and EVF. The rate of change in each compartment slows as the concentration difference is minimised. This phenomenon is seen in standard urea dialysis [Spiegel et al., 1995] and is known as 'rebound' but it was unclear, until this analysis, whether it would occur in middle-molecule clearance for the size of FLCs. The increase in FLC after the rebound period has elapsed, as seen in Figure 4.5(a), is due to the production of new FLC proteins.



(a) Normalised plasma quantity ($\frac{q_1(t)}{q1(0)}$). (b) Normalised Total Quantity ($\frac{q_1(t)+q_2(t)}{q_1(0)+q_2(0)}$).

Figure 4.6: *Comparison of increased filter clearance (k_a), all values taken from patient DM in Table 4.4. Solid-line DM fitted $k_a = 0.011$ $mins^{-1}$, dotted-line 2 k_a, chained-line 4 k_a and dashed-line 8 k_a.*

The effects of intercompartmental transfer in addition to explaining post dialysis rebound also have an impact upon the maximum clearance achieved by a dialysis membrane. Clinicians believe that merely increasing the clearance to as large a value as possible will ultimately remove all the FLC contained within the patient in the shortest time, to give the kidneys the maximum possibility of recovery. Due to the intercompartmental transfer rate there is a diminishing return on the observable FLC removed from plasma, which is magnified in terms of the total FLC removed.

In order to demonstrate this to clinical staff, a series of simulations have been conducted. The FLC model is initialised with the parameter estimates obtained for patient DM and simulated over a single dialysis session with increasing values of dialysis clearance k_d, the results of which can be seen in Figure 4.6. Over the four simulations the value of clearance is doubled. The normalised plasma quantities are shown in Figure 4.6(a) and the effects of doubling the clearance are reduced by intercompartmental

transfer; when the aphersis clearance is $8 \times k_a$ the 'plateauing' is within the first hour of treatment. Figure 4.6(b) shows the total FLC quantity ($q_1(t)+q_2(t)$), normalised to the initial values, where the impact of rapidly clearing FLC from the plasma compartment has a relatively small effect on the total FLC quantity. It should be noted that as the volume of each compartment is considered constant the above analysis applies equally well to FLC concentration.

4.3 Treatment prediction and comparison

During the previous analysis of filter types it was necessary to perform parameter estimation in order to determine the clearance rates of the filters. The information gained during this process can be used to investigate alternative treatments and compare the benefits of using haemodialysis. The current clinical 'best practice' to increase the removal of FLC is to use plasma exchange [Bradwell et al., 2005]; as explained in Chapter 2, this is the simple removal of a fixed amount of plasma volume over the treatment period which is replaced with a saline based solution. In order to extend the use of high-flux haemodialysis it must be shown that the new treatment modality is a significant improvement on the previous best practice. Mathematical modelling is a fundamental part of this process, allowing clinicians to analyse the treatment modalities prior to validating the methods against a larger cohort study and change current safe and known procedures.

In order to investigate haemodialysis and FLC in a meaningful context for clinical experts it is essential to replicate the whole of the treatment regimes; for Multiple Myeloma the modelling must therefore account for both chemotherapy and the alternative apheresis schedule.

4.3.1 Apheresis schedules

As previously stated, for patients with FLC there are two possible methods to assist in the clearance of FLC: plasma exchange and haemodialysis. However, due to the nature of the treatment mechanism these methods can be used with different schedules and treatment patterns. Plasma exchange removes all materials present in the blood that

have been exchanged, including proteins and fluids that are beneficial to the patient; as such over use of plasma exchange can have a detrimental impact on the patient's health. It therefore can only be used for a limited period after which the patient must be taken off plasma exchange to recover. The standard treatment is six plasma exchange sessions over a two week period, which is then stopped for an unspecified period until the patient's plasma has recovered, generally for a period of one to two months. In the study below a more extreme plasma exchange regime is tested. With severe medical conditions patients can be treated daily with plasma exchange for up to ten days [Hutchison et al., 2007].

Haemodialysis is less aggressive relative to plasma exchange. By careful balancing of the contents of the dialysate (a process that is handled by the dialysis machine) clinicians can remove only the toxic proteins, leaving beneficial elements in the blood to re-circulate. It is therefore possible to have a greater range of treatment schedules. The following schedules were suggested by clinical collaborators: four hours - three times per week, four hours daily, eight hours on alternate days, eight hours daily, twelve hours daily and eighteen hours daily. The first four treatment schedules reflect current practices in a standard treatment for a renal compromised patient [Hutchison et al., 2007; Kumar et al., 2004]; the patient would be moved between schedules depending on response. The eight hour treatment schedules would be rare, but are occasionally used in practice. The remaining treatments are considered experimental, with such periods on dialysis still being under investigation [Kumar et al., 2004]; however, with the advent of home haemodialysis facilities this may become more feasible.

4.3.2 Modelling chemotherapy effects

If it is assumed that the patient has reached a maximal concentration of FLC prior to treatment, and that unless treated this would be maintained indefinitely, i.e. the system is in steady-state, the synthesis rate of FLC prior to treatment (time $t = 0$) can be derived from the system equations ($f(0)$ in equation 4.10). This can be also considered to be the constant production required to maintain an *in vivo* plasma concentration at a steady-state level (e.g. 10 g/L) without treatment (chemotherapy or haemodialysis).

To simulate a patient's response to chemotherapy the FLC synthesis rate, previ-

ously assumed to be constant, must provide a dynamic response to treatment. As the direct link between synthesis and chemotherapy is not known several simple assumptions will be made to provide a clinically significant model of chemotherapy effect of FLC generation.

It is assumed that chemotherapy reduces the FLC synthesis rate by killing a set proportion of plasma cells per day (e.g. 50%) until non-cancerous levels of production are reached. As each individual plasma cell produces FLC at a constant rate [Sullivan and Salmon, 1972] the impact on FLC generation will be equivalent to the effective tumour kiil. This is approximated by a function that exponentially decreases to a known baseline. Therefore the FLC synthesis term ($f(t)$ in equations 4.6) can be expressed as:

$$f(t) = (p_{max} - p_{min})e^{-\lambda t} + p_{min}, \tag{4.11}$$

where p_{min} is the asymptotic baseline production, representing natural polyclonal FLC synthesis; p_{max} is a constant that encapsulates the maximum tumourous synthesis; in real terms this is calculated as the malignant production rate at the start of treatment to maintain a pre-treatment concentration. The decay constant (λ) can be changed to achieve the desired percentage tumour kill per day. This simple model can be extended to accommodate two complications that occur clinically during chemotherapy treatment: delay to starting chemotherapy and interruptions during the chemotherapy schedule. These may occur for either clinical or procedural reasons, e.g. the patient is unable to continue treatment or resources are unavailable.

The effectiveness of the chemotherapy drugs will vary between patients, with some seeing no alteration in FLC levels after chemotherapy; therefore a range of tumour kills were simulated, after consultation with clinical experts the following were chosen to reflect a range of responses: 0%, 2%, 5%, 10% and 100%. The first (0%) representing a worst case scenario of a completely ineffectual chemotherapy treatment, and conversely 100% the ideal situation were the tumour production of FLC is reduced to normal levels instantly. This approximation is necessary as the actual impact of the chemotherapy on plasma cells is currently unknown; this will be investigated further in Chapter 7.

Delayed chemotherapy

In some patients the chemotherapy is delayed until after apheresis; in such cases the production (eqn. 4.11) can be modified to a simple piecewise continuous function

$$f(t) = \begin{cases} p_{max} & 0 < t \leq t_s \\ (p_{max} - p_{min})e^{-\lambda t} + p_{min} & t_s < t \leq t_e \end{cases} . \tag{4.12}$$

the interval $t_s \leq t \leq t_e$ denotes the effective period for chemotherapy.

Interruptions in chemotherapy

To replicate interruptions in the chemotherapy treatment it is assumed that whilst the patient is on chemotherapy the FLC synthesis will decay at a constant rate (e.g. reduces by 5 % per day), and conversely, when the patient is not receiving chemotherapy the synthesis will exhibit growth of the same magnitude as the previous decay (e.g. increases by 5 % per day). Therefore, chemotherapy interruptions can be generically modelled by

$$f(t) = (f(\tau_i) - p_{min})e^{\lambda_i(t-\tau_i)} + p_{min}, \quad t \in (\tau_i, \tau_{i+1}] \tag{4.13}$$

with $f(0) = p_{max}$ (e.g. the production required to maintain 10 g/L) and λ_i dictates if the synthesis grows ($\lambda_i > 0$) or decays ($\lambda_i < 0$) in response to the application or removal of chemotherapy treatment over that period.

4.3.3 Patient FLC exposure

To elicit information visually for all simulation results would be extremely difficult; therefore, a quantitative approach has been developed to assist clinicians in the comparison of treatments. During treatment two key clinical factors are paramount: the time taken for the plasma FLC concentration to reach clinically acceptable levels and the total exposure the body has to the monoclonal FLC. The first of these can be simply handled by estimating the time the simulation takes to reach the required concentration, and for exposure concepts from pharmacokinetics can be used [Ritshel and Kearns, 2004, chap. 20].

Drug exposure, calculated as the integral of the plasma concentration time profile (often referred to as the Area-Under-the-Curve or AUC method), is a commonly used estimate of the total amount of drug in the body in pharmacokinetics [Ritshel and Kearns, 2004, chap. 20]. As clinicians are primarily interested in the FLC concentration delivered to the kidneys via plasma an equivalent FLC exposure is appropriate. The exposure of FLC between the start ($t = 0$) and end of patient monitoring ($t = T$) is calculated for each scenario

$$E_{ij} = \int_0^T y_{ij}(t)\, dt. \tag{4.14}$$

The exposure E_{ij} is a scalar value for the AUC of the simulation with tumour kill i (e.g. 10%) and treatment regime j (e.g. 12 hours daily); y_{ij} is the simulated output from the model, in all the comparisons made below this the plasma concentration of FLC. The percentage improvement gained from using the treatment j can thus be quantified by

$$\frac{E_{i0} - E_{ij}}{E_{i0}}, \tag{4.15}$$

where E_{i0} is the area under the curve with tumour kill estimate i and no apheresis, E_{ij} is the AUC calculated from the simulated result with the same tumour kill and apheresis treatment j. In all cases the integration in equation (4.14) is calculated numerically.

4.3.4 Comparing plasma exchange and haemodialysis

To compare the combined apheresis and chemotherapy treatments a series of simulations were conducted using the FLC model, instantiated either with parameter estimates from the patient DM in the filter comparison or values generated to match clinical settings. A full list of parameter values can be seen in Table 4.5. For each of the treatment regimes outlined (Section 4.3.1) a simulation was performed that estimated the patient FLC concentration with one of the effective tumour kill scenarios (Section 4.3.2).

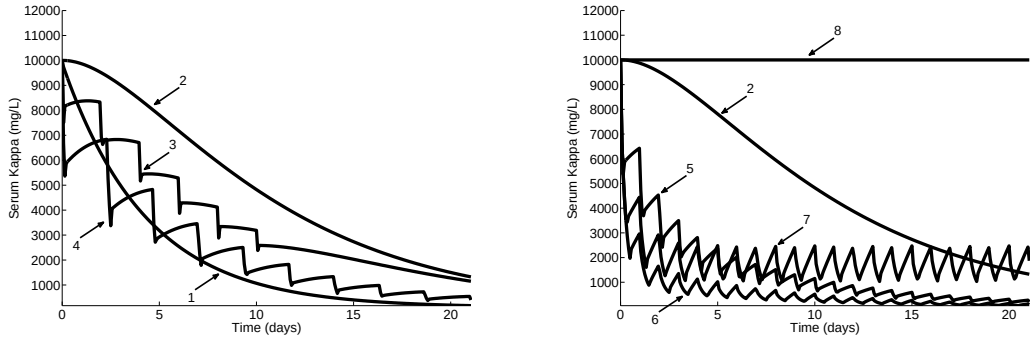
Table 4.5: *Parameter values for treatment comparison simulation: a Multiple Myeloma patient with a FLC-Kappa tumour producing steady state initial FLC concentration of 10 g/L.*

Parameter	Description	Value	Units
$c_1(0)$	Initial concentration in plasma due to tumourous synthesis.	10	g/L
p_{max}	Tumourous synthesis of FLC-Kappa.	22.95	mg/min
p_{min}	Normal FLC-Kappa production.	0.22	mg/min
k_{21}	Transfer rate from plasma to EVF (DM, Table 4.4)	0.031	min^{-1}
k_{re}	Reticuloendothelial clearance (estimated from literature, see text).	1.6×10^{-4}	min^{-1}
k_{01}	Normal renal clearance of FLC-Kappa (estimated from literature, see text).	0.012	min^{-1}
v_1	Plasma volume (assumed 70 kg male)	2.58	L
v_2	EVF volume (assumed 70 kg male)	12	L
k_d	Haemodialysis clearance (DM, Table 4.4)	0.011	min^{-1}
k_p	Plasma exchange clearance of $2\frac{1}{3}$ litres per hour.	0.015	min^{-1}

To allow different apheresis techniques and regimes to be compared the model parameters are all maintained as defined in Table 4.5. FLC synthesis ($f(t)$) is controlled by equation (4.11), with the decay rate (λ) calculated from the percentage reduction in plasma cells. The apheresis clearance is controlled by the time-varying function

$$k_{ai}(t) = \begin{cases} 0 & t \notin [\tau_{si}, \tau_{ei}] \\ k_p \text{ or } k_d & t \in [\tau_{si}, \tau_{ei}] \end{cases} . \tag{4.16}$$

τ_{si} and τ_{ei} are the start and end times of the apheresis treatment (i); with k_a set to either k_p if the treatment is plasma exchange or k_d for haemodialysis (see Table 4.5).



(a) Simulations 1-4

(b) Simulations 5-8

Figure 4.7: *Simulations of serum free light chain (FLC) removal by plasma exchange versus haemodialysis. Simulations: 1) 100% tumour kill on day one with only reticuloendothelial system removal. 2) 10% tumour kill per day with reticuloendothelial system removal alone. 3) 10% tumour kill per day with plasma exchange (3.5 litres exchange in 1.5 hours × 6 over 10 days). 4) 10% tumour kill per day with haemodialysis for 4 hours, three times a week. 5) 10% tumour kill per day with haemodialysis for 4 hours per day. 6) 10% tumour kill per day with haemodialysis for 12 hours per day. 7) No tumour kill with 8 hours haemodialysis on alternate days. 8) No tumour kill and no haemodialysis*

The results of the simulations for a patient with 10% tumour kill can be seen in Figure 4.7; the results are split between Figures 4.7(a) and 4.7(b) for clarity only. Referring to Figure 4.7(b), line 8 represents a patient that is not responding to chemotherapy and is not undergoing any form of apheresis. In contrast line 1 shows a patient responding extremely well to chemotherapy with 100% tumour kill instantly (time $t = 0$). For reference both graphs contain line 2, no apheresis and with 10% tumour kill per day. The

remaining plots in Figure 4.7(b) show the effects of continuous dialysis, with lines 5 and 6 showing the difference between 4 and 12 hour dialysis sessions. As would be expected the 12 hour sessions reduced the FLC to normal levels (500 mg/L) approximately 10 days earlier than the 4 hour sessions. Line 7, however, shows an interesting facet of the treatment revealing that, with no chemotherapy, if the patient can be maintained on an 8 hour daily session the FLC levels can be maintained at levels that may prevent the kidneys from further damage.[2]

Figure 4.7(a) compares plasma exchange and haemodialysis to the related chemotherapy kill. Line 3 depicts the plasma exchange treatment, 6 sessions in two weeks, with the remainder of the treatment period relying solely on natural clearance. Using plasma exchange shows a marked increase in clearance compared to the no apheresis values (Line 2). However, if the equivalent haemodialysis schedule is used, 4 hours 3 times a week, the clearance seen is 'close' to the clearance seen if the chemotherapy had been 100% at the first injection.

The results of all the simulations are summarised in Table 4.6, from which the suggested benefit of using apheresis can be seen. In each instance an improvement over the non-apheresis treatment is evident. Of greater interest is the comparison between plasma exchange and haemodialysis; for chemotherapy responses the shortest haemodialysis schedule (4 hours $\times$ 3/week) are a major improvement over the most active plasma exchange treatment possible (10 sessions in 10 days). Plasma exchange schedules also took much longer to reach the 500 mg/L concentration levels; for an ineffective chemotherapy result the peak plasma exchange concentration did not fall below the initial concentration for either of the plasma exchange treatments. As would be expected, as the length of haemodialysis increases the percentage improvement increased and the length of time to reach 500 mg/L reduced. It is important to note the relatively small improvement gained in moving to an 18 hour daily schedule from an 8 hours session; this reflects the impact of the intercompartmental transfer as noted in Section 4.2.2. However, as with all *de novo* experiments, care should taken when interpreting the results without validation to real patient response data.

[2]Bradwell et al. [2004] define this as 1 g/L.

Table 4.6: *Model calculations of the efficiency of therapeutic removal of free light chains (FLC). 100% to 0% are the chemotherapy tumour kill rates per day. Numbers are the additional % of FLCs removed by intervention beyond that from tumour cytoreduction from chemotherapy. Numbers in brackets are the time in days for FLC concentrations to reduce from 10 g/L to 0.5 g/L. * Serum FLC concentrations at day 150 when 0.5 g/L was not achieved. PE: plasma exchange. HD: haemodialysis. NA: not applicable. 1-8 Simulations shown in Figure 4.7.*

Method of FLC removal	Percentage of FLCs removed by apheresis (and time, in days, to reduce to 0.5 g/L)				
Tumour kill per day	100%	10%	5%	2 %	0%
None	NA(14)[1]	NA(30)[2]	NA(52)	NA(121)	NA(*10 g/L)[8]
PE x 6 in 10 days	37(10)	28(29)[3]	24(52)	21(121)	19(*10 g/L)
PE x 10 in 10 days	51(8)	40(29)	34(52)	30(121)	27(*10 g/L)
HD 4 hrs x 3/week	62(7)	53(19)[4]	42(31)	39(73)	37(*6.2 g/L)
HD 4 hrs daily	77(4)	72(13)[5]	70(23)	69(55)	68(*3.7 g/L)
HD 8 hrs alternate days	81(4)	73(13)	71(19)	69(47)	67(*4.4 g/L)[7]
HD 8 hrs daily	88(3)	84(7)	83(14)	82(29)	81(*2.2 g/L)
HD 12 hrs daily	92(2)	89(5)[6]	88(8)	87(16)	87(*1.8 g/L)
HD 18 hrs daily	94(2)	93(3)	92(4)	92(8)	91(*1.2 g/L)

4.3.5 Intensive haemodialysis schedules

In previous examples several dialysis schedules were simulated to investigate the effects of frequency and duration of haemodialysis when compared to plasma exchange. In clinical settings the amount of haemodialysis a patient will be able to receive will change depending on the local facilities. To investigate the impact of this on the benefits of high-flux haemodialysis for FLC patients several schedules were generated that reflect a realistic schedule for hospital resources and procedures:

1. No apheresis
2. Standard Schedule - four hours every Monday, Wednesday and Friday.
3. Extended Schedule 1 - six hours every Monday, Wednesday, Friday and Sunday.

4. Extended Schedule 2 - six hours day one and then 8 hours, days two to five, followed by eight hours, alternate days.

In order to compare these schedules it was assumed that the patient was receiving chemotherapy during the entire period and achieved a 5 % tumour kill per day, to the polyclonal baseline for FLC-Kappa production, see Table 4.5. Each of the schedules was simulated over a thirty day period. The patients' exposure to FLCs was calculated, using equations (4.14) and (4.15), for these periods; exposure was also calculated for the regime with and without haemodialysis, see Table 4.7.

Table 4.7: *Reductions in serum FLC exposure with intensive haemodialysis schedules. Peak FLC is the maximum plasma concentration after four days.*

Schedule (days)	FLC exposure (g days)	FLC reduction (%)	Peak FLC (g/L)
No HD	422	0	9.2
Standard	242	42	3.6
Extended 1	182	56	3.0
Extended 2	141	66	1.3

Whilst the results are unsurprising with regard to the improvement seen when haemodialysis is used, regardless of the treatment intensity, it shows the relatively small improvements gained from using the extended schedules. If the total amount of time the patient spends on haemodialysis is considered over the 30 day period, for the standard schedule (52 hours), extended schedule 1 (102 hours) and extended schedule 2 (134 hours), it can be seen that there is not a simple relationship between the FLC exposure and dialysis frequency and duration. However, the benefit of schedule 2 with the initial intensive chemotherapy can be seen if the peak plasma concentrations are observed during the first four days, with extended schedule 2 reducing to almost 10 % of the original concentration. As previously mentioned, this may have clinical significance by allowing the kidneys to recover some function, which as will be seen in Section 4.3.7 can have implications on long term patient recovery.

4.3.6 Effects of modified chemotherapy treatment

Whilst the previous study included the concept of chemotherapy effectiveness, clinicians are equally concerned with the timing of chemotherapy. During treatment, alterations to the chemotherapy schedule are often required to incorporate either patient health or resource availability. Therefore, two scenarios were investigated:

1. Delay in initiating chemotherapy.

2. Interruptions to chemotherapy.

Delayed chemotherapy

The importance of rapid initiation of chemotherapy was evaluated by comparing the total FLC exposure when chemotherapy was started at days 0, 5, 10, 15 and 20. The potential role of haemodialysis to protect the kidneys from ongoing high FLC concentrations was explored by repeating these simulations with haemodialysis. Once again, for comparison, a baseline effectiveness of chemotherapy was assumed to be a 5% reduction in synthesis rates of monoclonal FLCs per day and the extended schedule 2 (see Section 4.3.5).

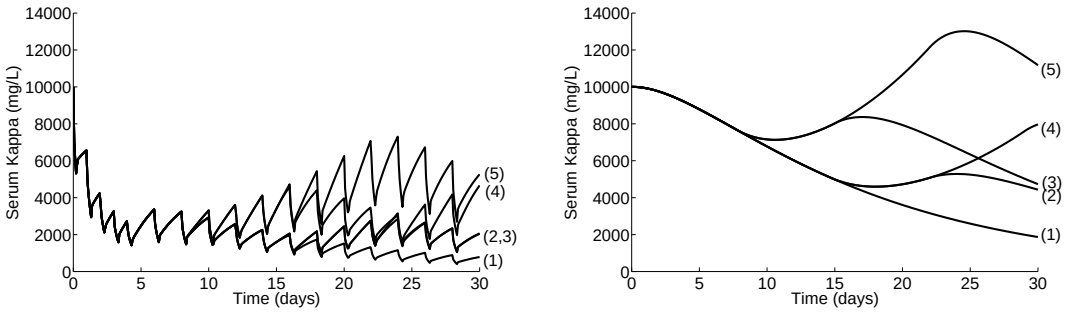
Table 4.8: *Changes in FLC exposure when initiation of chemotherapy is delayed for a variable period, with and without haemodialysis.*

Delay (days)	FLC exposure (g days)		FLC reduction %	
	without HD	with HD	without HD	with HD
0	417	106	0	74
5	519	138	-24	67
10	609	169	-46	59
15	685	196	-64	53
20	739	218	-77	48

The results of delaying chemotherapy can be seen in Table 4.8. When compared to a 5% tumour kill per day effectiveness, delaying had a dramatic impact on the kidney's exposure to FLC, increasing from 24% (5 days delay) to 77% (20 days) without

haemodialysis. However, if it is assumed that haemodialysis is administered immediately at the start of observation, the FLC exposure is significantly reduced. The importance of haemodialysis increases as the delays in initiating chemotherapy became greater.

Interrupted chemotherapy



(a) Haemodialysis

(b) No haemodialysis

Figure 4.8: *Impact of interrupting chemotherapy on serum concentrations of FLCs with and without dialysis.*

The influence of stopping and re-starting the treatment were assessed through simulating interruptions at days eight and fifteen, the simulated patient response was then left for either seven or fourteen days before beginning treatment again. For the purpose of the simulation it was assumed that the chemotherapy killed 5% of the tumourous cells per day, and conversely, when the chemotherapy was stopped the cancerous cells would increase by 5% per day.

Interruptions to chemotherapy increased the FLC concentration in plasma (Figure 4.8) and the total FLC exposure significantly (Table 4.9). As would be expected, the earlier the interruption occurred the greater the increase in FLC exposure. Similarly, the longer the interruption continued the greater the exposure. As can be seen clearly in Figure 4.8, whilst the interruptions had dramatic effects on the FLC concentrations in the simulated patient, haemodialysis was beneficial in controlling the variations in the monoclonal synthesis; with haemodialysis maintaining a lower level than the initial concentration when compared to the 'no dialysis' simulation. This is most evident in scenario five with the increase in FLC synthesis forcing the concentration up to 14 g/L if

Table 4.9: *Changes in FLC exposure when chemotherapy was interrupted, with and without heamodialysis.*

Chemotherapy Interruptions (days)				FLC exposure (g days)		FLC reduction (%)	
Line	Start	Stop	Re-start	without HD	with HD	without HD	with HD
1	0	n/a	n/a	423	142	0	66
2	0	15	22	486	173	-16	58
3	0	8	15	586	215	-41	48
4	0	15	29	515	192	-23	54
5	0	8	22	752	299	-80	28

haemodialysis was not used, whilst with haemodialysis the peak value is well below the initial concentration ($\approx$ 8 g/L). An interesting feature is shown in treatment scenario four, although the patient is off chemotherapy for the longest period (fourteen days) a lower exposure is seen; this is due to the impact of a longer time on chemotherapy prior to the interruption. However, it should be noted that, if the simulations were left to run, scenario four would exhibit much worse behaviour due to the natural rising of the synthesis rate.

4.3.7 Renal recovery

As stated in Chapter 2, it has been noted [Bradwell et al., 2004; Leung et al., 2008] that when FLC levels are reduced significantly some patients may recover partial renal function. In addition, if patients are diagnosed with FLC secreting Myeloma before the nephrons are completely damaged there may be some natural renal clearance maintained. It is of interest to determine the minimum recovery of renal function recovery required to benefit the patient with respect to FLC clearance. For this the extended haemodialysis schedule 2 (see Section 4.3.5) was used with 0, 5, 10, 20, 40% of normal renal function (all simulations were undertaken assuming a 5% reduction in FLC production per day in response to chemotherapy).

In previous examples the renal function was assumed to be zero, however for this study an estimate is required for the renal clearance of FLC (k_{01} in equation (4.6)). The

patients observed during this study were in end-stage renal failure, and as such had no renal clearance of FLC, it was not possible to obtain values for k_{01} during parameter estimation. However, renal clearance can be estimated from the glomerular filtration rate (GFR), which is an indicator of renal function obtained from measuring other products in urine, commonly from urea [Coresh et al., 2005]. For FLC-Kappa renal clearance is approximately 40% of GFR, whilst for FLC-Lambda it is approximately 20% [Bradwell et al., 2005, pp. 17-19]. The base GFR is taken from Coresh et al. [2005], assuming a 60 year old male patient.

Table 4.10: *Effects of renal recovery on FLC exposure in an FLC-Kappa patient.*

Renal function (%)	FLC exposure (g days)		FLC reduction (%)	
	without HD	with HD	without HD	with HD
0	420	133	0	68
5	253	108	40	74
10	181	91	57	78
25	97	62	77	85
40	67	47	84	88
100	30	24	93	94

The results of increasing renal function can be seen in Tables 4.10 and 4.12. It clearly shows that renal function must recover significantly before haemodialysis can be discontinued; for both FLC-Kappa and FLC-Lambda renal function only has to be greater than 10% of the healthy value (see Table 4.5 and 4.11) to achieve results similar to that through using extracorporeal filters. This is due to the fact that natural clearance occurs constantly, whilst haemodialysis clearance is limited to the times when the patient is on dialysis. The results also convey the message that, if it is possible, haemodialysis is extremely beneficial for clearing FLC quickly, even if the kidneys have recovered their clearance function.

Table 4.12 shows the effect of the recovery of kidney function on FLC-Lambda. As with the FLC-Kappa patient a significant increase is seen when renal function is present and dialysis is not used. The effect is smaller than that for FLC-Kappa patients

Table 4.11: *Parameter values for treatment comparison simulation: a Multiple Myeloma patient with an FLC-Lambda tumour producing an FLC concentration of 10 g/L. All other parameters are as for the FLC-Kappa patient Table 4.5.*

Parameter	Description	Value	Units
p_{max}	Tumourous synthesis.	22.74	mg/min
p_{min}	Normal (polyclonal) production.	0.12	mg/min
k_{21}	Transfer rate from plasma to EVF (MF, Table 4.4)	0.021	min^{-1}
k_d	Haemodialysis clearance (MF, Table 4.4)	0.015	min^{-1}
k_{01}	Normal renal clearance of FLC-Lambda (see text).	0.006	min^{-1}
k_{re}	Reticuloendothelial clearance.	9.6×10^{-4}	min^{-1}

due to the slower underlying renal clearance of FLC-Lambda proteins. The differences seen in the exposure between the FLC-Kappa and Lambda in Tables 4.10 and 4.12 can be related to the differences in polyclonal synthesis, recticuloendothial clearance, intercompartmental transfers rates and differences in haemodialysis clearance.

4.4 Analytical asymptotic prediction

In the previous work all the results were obtained from numerically evaluating the system of differential equations (eqns. 4.6) and solving them using numerical integration. For engineers and mathematicians such techniques are relatively simple; this is not the case for clinicians. It would be beneficial if a simpler method could be derived that encapsulates the treatment information in an analytical expression, which could be implemented in a more accessible or familiar analysis tool, e.g. a spreadsheet; thus providing a mechanism for non-technical staff to interpret the results during patient treatment.

Table 4.12: *Effects of renal recovery on FLC exposure in FLC-Lambda patient.*

Renal function (%)	FLC exposure (g days)		FLC reduction (%)	
	without HD	with HD	without HD	with HD
0	469	107	0	77
5	311	94	34	79
10	231	84	51	82
25	130	64	72	86
40	90	52	81	88
100	41	30	91	94

4.4.1 Switched systems

In a standard system as defined by equations (4.6) the asymptotic values can be described using steady-state assumptions on the model. However, during the apheresis treatment the state-matrix is not continuous and changes structure depending on whether the patient is on or off apheresis. To accommodate this discontinuity it is useful to consider the concept of a switched system.

Definition 4.4.1. A switched system is made up of a finite subset of systems with a rule that defines when each subsystem is active.

$$\dot{q}(t,p) = f_i(q(t,p),p)\,q(t) + g_i(q(t,p),p)\,u(t) \tag{4.17}$$

$$y(t,p) = h_i(q(t,p),p) \tag{4.18}$$

$$q(0,p) = q_0(p) \tag{4.19}$$

where the switch-state i is controlled by

$$i = \begin{cases} 1 & \text{if } t \in [t_0, t_1) \\ \vdots & \\ n & \text{if } t \in [t_{n-1}, t_n) \end{cases} \tag{4.20}$$

with switching between states occurring instantaneously.

The apheresis models presented thus far fall into a specialised set of switched systems that are periodic linear time-invariant (LTI) state-switch systems.

Definition 4.4.2. A periodic LTI state-switched system is of the form

$$\dot{q}(t,p) = A_i(p)q(t) + B(p)u(t) \tag{4.21}$$

$$y(t,p) = C(p)q(t) \tag{4.22}$$

$$q(0) = q_0 \tag{4.23}$$

with a switching sequence of period T

$$i = \begin{cases} 1 & \text{if } t \in [t_0 + lT, t_1 + lT) \\ \vdots & \qquad\qquad l = 0, 1, \ldots \\ n & \text{if } t \in [t_{n-1} + lT, t_n + lT) \end{cases} \tag{4.24}$$

Research into switched systems has received attention from the control community; for a review see Liberzon and Morse [1999]. However, the focus is on the stability of such systems under an arbitrary switching sequence ([Cevat, 2004; Phat and Pairote, 2006]) or the control of the switched systems through the manipulation of the switching sequence and input function. Whilst excellent for the handling of electrical and mechanical systems, these approaches are not readily applicable to biomedical systems. As previously stated, the control over the system input is rudimentary at best and the application of dialysis as treatment is not applicable to automated switching due to the preparation and planning required to implement it. The analysis here is therefore reduced to estimating *in vivo* results from known conditions and a predefined switching sequence.

4.4.2 Estimates for pre- and post-apheresis measurements

Assuming the *in vivo* kinetics are modelled by a linear period switched system, an analytical solution for the asymptotic maximum (pre-dialysis) and minimum (post-dialysis) solutions are available.

Proposition 4.4.1. *Given definition 4.4.2, the asymptotic minimum and maximum values of a linear single-switch*[3] *state-space system with constant input are given by*

[3] A system with only two possible system matrices

$$\begin{aligned}\lim_{t\to\infty} \max c(t) &= (I_d - e^{A_1\tau_1}e^{A_2\tau_2})^{-1}(e^{A_2\tau_2}\alpha + \beta) \\ \lim_{t\to\infty} \min c(t) &= (I_d - e^{A_2\tau_2}e^{A_1\tau_1})^{-1}(\alpha + \beta e^{A_1\tau_1})\end{aligned} \tag{4.25}$$

where e^{Ai} is the matrix exponential, $A_1 \in \mathbb{R}^{d\times d}$ and $A_2 \in \mathbb{R}^{d\times d}$ are the two possible system matrices, τ_1 is the duration system one is operational, τ_2 system matrix two, I_d is the identity matrix of dimension d; with α and β defined as

$$\alpha := \int_0^{\tau_1} e^{A_1(\tau_1 - s)}\, P_\infty\, ds \tag{4.26}$$

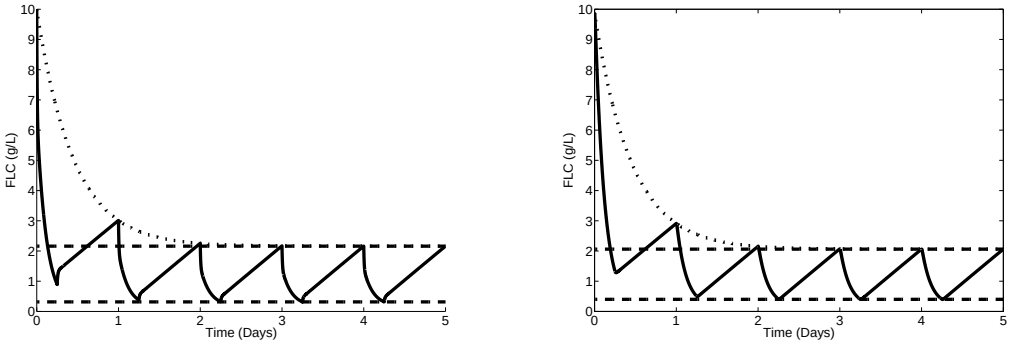
$$\beta := \int_0^{\tau_2} e^{A_2(\tau_2 - s)}\, P_\infty\, ds \tag{4.27}$$

where P_∞ is a vector of the asymptotic baseline of the input functions; for FLC patients this is $P_\infty = [f_{min}, 0]^T$.

The proof of this proposition can be found in Appendix B; it uses a recursive expansion of the analytical solution of consecutive phases of treatment and rest periods to derive the above expressions. It is also possible from the asymptotic estimation, see Appendix B, to construct an envelope function that approximates the maximum values of patient concentration during the period when the initial conditions continue to influence the system

$$q_{env}(t) \quad = (e^{A_2(T-D)}e^{A_1 D})^{\frac{t}{T}}(c(0) - \max c(t)) + \max c(t); \tag{4.28}$$

where T is the period between dialysis sessions and D is the duration; however, this function should only be interpreted at the time points $t = nT$, $n \in \mathbb{Z}^+$. Using these equations it is possible to show the maximum concentration a treatment schedule will provide, and the time at which the patient's plasma concentration will reach a certain level, e.g. 500 mg/L, without the necessity of implementing the system as a set of time-variant ODEs, as has been done previously.



(a) Plasma (b) EVF

Figure 4.9: *Asymptotic pre- and post-dialysis quantities and exponential envelope function. Chained-line is the minimum and maximum values, dotted-line, the envelope function and the solid-line is the simulated patient response from the full model.*

4.4.3 Patient example

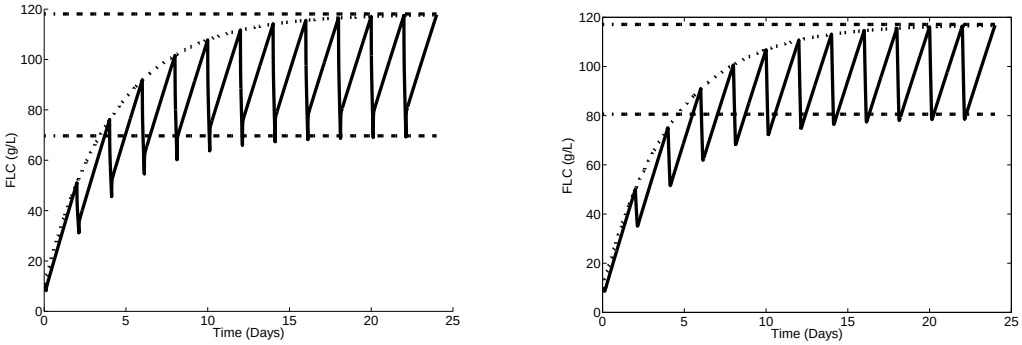
As an example, consider a patient receiving periodic haemodialysis treatment; using the linear two compartment FLC model, results in a switched system with the two system matrices defined as

$$A_1 = \quad A_2 + \begin{pmatrix} -k_d & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \quad \begin{pmatrix} -(k_{1e} + k_{21}) & k_{12} \\ k_{21} & -(k_{1e} + k_{12}) \end{pmatrix}. \quad (4.29)$$

If the parameters for the patient and treatment are known, the asymptotic values of FLC concentrations can be estimated from equation (4.25) and the peak value attained pre-dialysis can be estimated by equation 4.28. An illustration of this can be seen in Figure 4.9, where the parameters for patient DM, as defined in Table 4.5, are used. Figure 4.9(a) shows the results of simulating the patient treatment with a standard ODE solver (solid-line); overlaid onto this graph are the asymptotic maximum and minimum values (chained-line). It can clearly be seen that within six days of beginning treatment the patient's plasma concentration has reached the estimated asymptotic minimum and maximum values. In addition, an estimate can be made for the levels of FLC in the unobservable EVF compartment, see Figure 4.9(b).

An alternative example is shown in Figure 4.10, where rather than assuming the production remains at the value required to maintain steady-state, it is assumed to

increase, at time $t = 0$, by a factor of 10. All other parameters are as previously defined; this step in production results in asymptotic min-max values that are greater than the initial concentrations for both plasma and EVF.



(a) Plasma (b) EVF

Figure 4.10: *Asymptotic pre and post plasma exchange quantities and exponential envelope function. The chained-line shows the minimum and maximum values, the dotted-line, the envelope function and the solid-line is the simulated patient response from the full model.*

4.5 Extended filter model

In the previous section of this chapter models have been developed that address the problem of patient treatment from a whole-body perspective. The assumptions made during the model construction allowed a linear model to be used to compare the relevant treatments; however, whilst the *in vivo* kinetics are unknown, those of a standard counter-flow filter can be investigated *in vitro* [Galach et al., 2003; Waniewski, 1994]. In the following sections closer scrutiny is made of the filter kinetics, investigating the dynamics of a counter flow filter, and its impact on whole body clearance. This leads on to an extension of the above model to include the extracorporeal dialysis circuit (see Chapter 2).

4.5.1 Single filter dynamics

In the models presented thus far the dialysis removal rate (k_d) was considered constant over the dialysis session. However, from filter theory this is known not to be true

[Fournier, 2007, ch. 8]; the movement of particles across a semi-permeable membrane is driven by concentration (diffusive transport) and pressure gradients (convective transport). The amount of substance (q) crossing the membrane at time t can by described by the Kedem-Katchalsky equations [Kedem and Katchalsky, 1958]:

$$\dot{q}(t) \quad = D(c_1(t) - c_2(t)) + J_s\, c_m(t) \tag{4.30}$$

where $c_1(t)$ and $c_2(t)$ are the substance concentration either side of the membrane, D the diffusive permeability to the material in question, J_s the ultrafiltration rate and $c_m(t)$ the mean of intramembrane concentration (most often taken as the arithmetic mean i.e. $\frac{c_1-c_2}{2}$ [Williams, 1996]). For the patients considered in this study the ultrafiltration rate was zero, which reduces equation (4.30) to its diffusive components only:

$$\dot{q}(t) \quad = D(c_1(t) - c_2(t)). \tag{4.31}$$

It should be noted that these equations ignore other phenomena known to occur at the membrane-fluid interface (Donnan effect [Petitclerc, 1998] or membrane absorption [Clark et al., 1994]) and these will be discussed in Chapter 8. Whilst the dialysis equipment was set to ensure the ultrafiltration was zero, due to internal pressures between the filters ultrafiltration may still occur during treatment.

Equations (4.30) and (4.31) make it apparent that the concentration of the substance in the dialyzer is a key component, and as such should be considered in the model description. Therefore, in the following section a model is presented that will develop the membrane dynamics described above into a generic structure, enabling multiple filters configurations to be modelled, allowing the overall impact on whole body clearance to be assessed.

4.5.2 Multiple filter models

As stated in the biological introduction modern dialysis filters are a collection of semi-permeable fibres contained within a polyurethane jacket, that have inputs and outputs

for blood and dialysate fluid to flow through. The dialyser is connected into the patient/dialysis machine circuit through flexible tubes. This simple connection mechanism allows multiple filters to be connected together in a single circuit. Figure 4.11 shows how the plasma and dialysate flows are commonly connected for two filters.

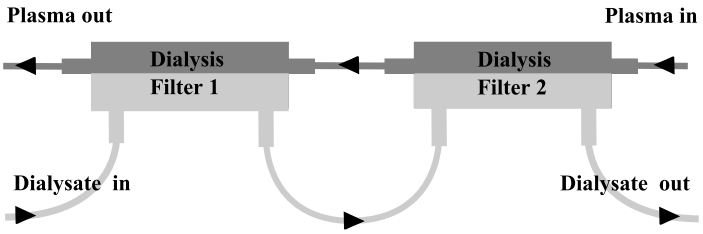


Figure 4.11: *Schematic dialysis filter configuration*

Some studies on multiple filter configurations have been conducted previously. Powers et al. [2000] used two large filters to determine their effectiveness at clearing urea, but the study was only concerned with pre- and post-dialysis urea concentrations, rather than the dynamics during treatment. An excellent *in vitro* comparison of serial and parallel configurations was conducted by Eloot et al. [2005], with varying flow rates. Blood flow and hematocrit were closely evaluated in a clinical setting by Kato et al. [2007] on β_2-microglobulin when a series dual-dialyser was used. However, these latter two studies used monitors on the inputs and outputs of filters to determine flow rates and pressure differences between the filters; however, no attempt was made to model the system or attempt prediction on patient *in vivo* kinetics. In addition, the methods used are not feasible for standard clinical treatment, i.e. it is not possible to have continuous monitoring for all patients during treatment outside of a specialist trial. Therefore, in the following sections a model is described that can use observations of plasma and dialysate concentrations of FLC only to determine parameter values for different filter configurations.

4.5.3 The model

If the FLC is assumed not to collect in the lines connecting the patient and filters (see Figure 2.4, Chapter 2) and there is no time lag between the FLC molecules leaving one compartment and entering another a generalised form of the compartmental model can be seen, Figure 4.12; for a description of the nomenclature used for the model please

refer to Table 4.5.

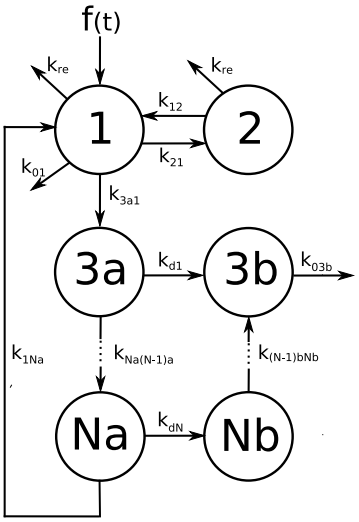


Figure 4.12: Extended filter model for N filters with plasma and dialysate in counter-flow configuration.

As in the FLC model [Evans et al., 2006] compartments 1 and 2 represent the flow of FLC *in vivo*, between plasma and EVF. The remaining compartments all relate to the extracorporeal dialysis circuit. The structure surrounding compartments 1 and 2 will remain unchanged during and after dialysis treatment, whilst the other compartments are only 'in-situ' during the dialysis period.

Compartments $3a$ to Na represent the FLC flowing in plasma in the filters; $3a$ donates the first filter, Na the $N-2$ filter. There may be numerous filters between these but in practice the maximum currently used is 3 filters. It should be noted that this is a closed-loop system and the output of the last filter is returned to the patient. This means that any FLC not removed by the filters is recirculated into the patient. It is assumed that the FLC remains well mixed whilst progressing through the tubes connecting the patients and filters and the volume of plasma contained in the connecting tubes is negligible.

The compartments $3b$ to Nb represent the FLC flowing (in the opposite direction for a counter-flow circuit) through the dialysate side of the filter. As with the plasma flow, the tubes are connected in series with respect to the dialysate fluid. Fresh dialysate is constantly fed into the first filter, the output of which is then fed as an input to the second filter. So any FLC removed from filter 1 (compartment Nb) will be passed

onto filter 2 ($(N-1)b$) and so on. Unlike the plasma side of the filter the output from the final compartment ($3b$) is not fed back, it is removed from the system. This has the effect of maintaining a constant volume in the dialysate compartments without the need for any form of input.

The rate constants between filters are derived from known fluid flows and volumes, e.g. $k_{3a1} = Q_p/v_1$, where Q_p is the FLC flow in plasma. The parameter k_{di} represents the diffusive permeability of the filter membrane (between FLC plasma and dialysate fluid).

The system of equations generated from the model (Figure 4.12) is

$$\begin{aligned}
\dot{q_1}(t) &= -(k_{21} + k_{re} + k_{01} + k_{3a1})q_1(t) + k_{12}q_2(t) + k_{1Na}q_{Na}(t) + f(t) \\
\dot{q_2}(t) &= k_{21}q_1(t) - (k_{re} + k_{12})q_2(t) \\
\dot{q_{3a}}(t) &= -k_{4a3a}q_{3a}(t) - k_{d1}(\frac{q_{3a}(t)}{v_{3a}} - \frac{q_{3b}(t)}{v_{3b}}) + k_{3a1}q_1(t) \\
\dot{q_{3b}}(t) &= -k_{03b}q_{3b}(t) + k_{d1}(\frac{q_{3a}(t)}{v_{3a}} - \frac{q_{3b}(t)}{v_{3b}}) + k_{3b4b}q_{4b}(t) \\
&\ldots \\
\dot{q_{Na}}(t) &= -k_{1Na}q_{Na}(t) - k_{dN}(\frac{q_{Na}(t)}{v_{Na}} - \frac{q_{Nb}(t)}{v_{Nb}}) + k_{Na(N-1)a}q_{(N-1)a}(t) \\
\dot{q_{Nb}}(t) &= k_{dN}(\frac{q_{Na}(t)}{v_{Na}} - \frac{q_{Nb}(t)}{v_{Nb}}) - k_{(N-1)bNb}q_{Nb}(t)
\end{aligned} \tag{4.32}$$

Prior to treatment the filters are 'primed' with the patients plasma by drawing it manually through the entire blood-side circuit; then clean dialysate is fed through the dialysate side of the filter at which point treatment is considered to have started. Therefore the quantity of FLC in the dialysate can be considered to be zero at the start of treatment, and the quantity of FLC in the blood side of the filter is equal to the concentration in compartment 1, resulting in the following initial conditions:

$$\begin{aligned}
q_1(0) &= c_{10}\, v_1, \quad & q_2(0) &= c_{10}\, v_2 \\
q_{3a}(0) &= c_{10}\, v_{3a}, \quad & q_{3b}(0) &= 0 \\
&\ldots & & \\
q_{Na}(0) &= c_{10}\, v_{Na}, \quad & q_{Nb}(0) &= 0
\end{aligned}$$

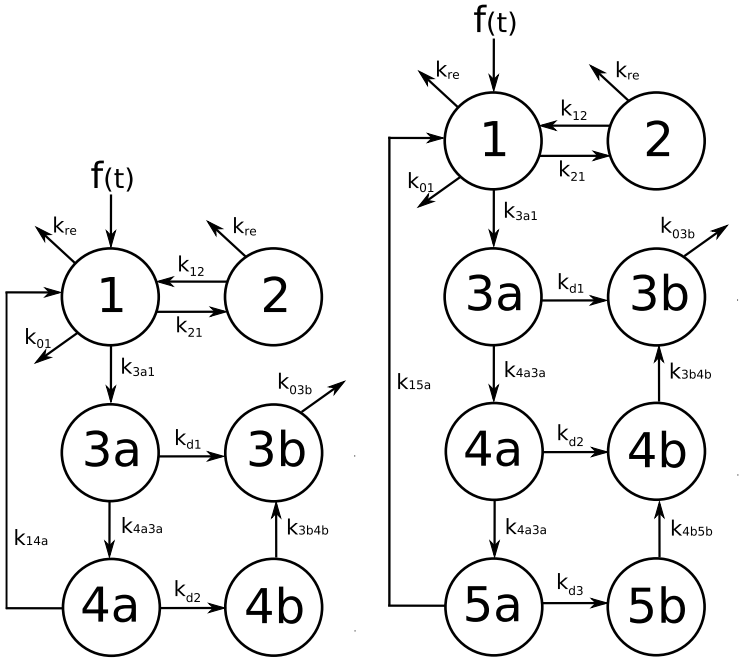
where c_{10} represents the concentration of the FLC in plasma at time $t = 0$ and v_i the volume of distribution of compartment i.

Previously only the FLC observed in plasma was considered; however, clinicians are able to take additional measurements during dialysis by extracting plasma samples from the patient directly or from plastic tubes between the dialysis filters, plasma or dialysate. However, in the above model the compartments representing the dialysis filter are lumped parameter models and thus model the average concentration within the filter, but measurements are taken at the ends of the filter, not the middle. This discrepancy can be accounted for by modifications to the observation functions. If it is assumed that the concentration change along the membrane is linear [Williams, 1996] and the compartments ($c_{3a}, ..., c_{Na}$ and $c_{3b}, ..., c_{Nb}$) represent the mean concentration at the mid-point of the filter (e.g. $c_{3a} = (c_1 + c_{4a})/2$) the observations ($y_i$) at the extremes of the filter can be described by

$$\begin{aligned}
y_1(t) &= c_1(t) && t \in [\tau_s, \tau_t) \\
y_{3a}(t) &= 2c_{3a}(t) - y_1(t) && t \in [\tau_s, \tau_e] \\
&... \\
y_{Na}(t) &= 2c_{Na}(t) - y_{N-1a}(t) && t \in [\tau_s, \tau_e] \\
y_{3b}(t) &= 2c_{3b}(t) - y_{4b}(t) && t \in [\tau_s, \tau_e] \\
&... \\
y_{Nb}(t) &= 2c_{Nb}(t) && t \in [\tau_s, \tau_e]
\end{aligned} \tag{4.33}$$

where τ_e equates to the end of the dialysis session, τ_s the start and τ_t the end of the patient's treatment. Whilst it is physically possible to measure at all connections points between the filters, only a limited subset of data were available: the concentration of FLC in plasma in the patient (y_1), between filters on the plasma side of the filter ($y_{3a}, ..., y_{Na}$) and the concentration leaving dialysate (y_{3b}).

Figure 4.13 shows how the generic extended model (Fig. 4.12) can be used to replicate different filter configurations. Figure 4.13(a) describes the kinetic model applicable when two filters have been connected in series and Figure 4.13(b), where a three-filter configuration is pictured. The extended filter model is not limited to a series configuration, with simple modification it is possible to model parallel configuration, in which either the plasma or dialysate flow is split across multiple filters. However, it was clinical infeasible to construct parallel configurations; therefore, the following validation is



(a) Two filters (b) Three filters

Figure 4.13: *A schematic of different filter configurations using the extended filter model.*

limited to serial configuration. The modelling of alternative filter configurations (e.g. filters in parallel) is discussed in Chapter 8.

4.5.4 Validation and parameter estimation

Prior to parameter estimation the model was tested for structural identifiability. As data were only available for the two filter model (Fig. 4.13(a)) the identifiability was only performed on this model. The differential algebra approach was used (using the DAISY package, see Chapter 3), gave rise to the results and input files are shown in Appendix A. It should be noted that the identifiability of the model was attempted using the Taylor series approach but due to the model structure the co-efficients become computationally intractable after the third derivative. With reference to equation 5.2 and Figure 4.13, for a two filter model the parameter vector for the model, assuming all constants are unknown, is given by

$$p_2 = \quad [c_{10}, v_1, v_{3a}, v_{3b}, v_{4a}, v_{4b}, k_{12}, k_{21}, k_{re}, f, k_{d1}, k_{d2}, k_{3a1}, k_{3b4b}, k_{14a}, k_{03b}, k_{4a3a}]^T.$$

It was found that by observing comparments $q_1(t)$, $q_{3a}(t)$, $q_{3b}(t)$ and $q_{4a}(t)$ the

parameter vector p_2 is globally identifiable. It can also be seen that several parameters relating to the *in vivo* kinetics are structurally identifiable that were unidentifiable in the FLC Model, notably, k_{12} and k_{re}. Previously these were estimated through physical assumptions which can be verified numerically with the extended model.

Although all the parameters were found to be identifiable, validating this model against real data gives extremely poor results in terms of both the fit to the data and the estimates obtained for the parameters, with several parameters being rejected from parameter estimation for being poorly determined numerically. On further investigation it has been found that in a dual filter urea haemodialysis an ultrafiltration circuit can develop between the two filters, even though ultrafiltration is not set in the dialysis equipment [Daugirdas, 2006]. This results in an ultrafiltration flow that pushes molecules between plasma and dialysate in the first filter and in the reverse direction (dialysate to plasma) in the second filter. These ultrafiltration rates must be equivalent, resulting in no net fluid loss during the treatment. Therefore, the full transport equations (eqn. (4.30)) are used over the diffusion equations (eqn. (4.30)). For a dual filter series configuration this results in the following system of equations:

$$
\begin{aligned}
\dot{q_1}(t) &= -(k_{re} + k_{21} + k_{3a1} + k_{01})q_1(t) + k_{12}q_2(t) + k_{14a}q_{4a}(t) + f \\
\dot{q_2}(t) &= -(k_{re} + k_{12})q_2(t) + k_{21}q_1(t) \\
\dot{q_{3a}}(t) &= -k_{4a3a}q_{3a}(t) + k_{3a1}q_1(t) - k_{d1}\left(\frac{q_{3a}(t)}{v_{3a}} - \frac{q_{3b}(t)}{v_{3b}}\right) \\
&\quad - \frac{uf_1}{2}\left(\frac{q_{3a}(t)}{v_{3a}} + \frac{q_{3b}(t)}{v_{3b}}\right) \\
\dot{q_{3b}}(t) &= -k_{03b}q_{3b}(t) + k_{4b3b}q_{4b}(t) + k_{d1}\left(\frac{q_{3a}(t)}{v_{3a}} - \frac{q_{3b}(t)}{v_{3b}}\right) \\
&\quad + \frac{uf_1}{2}\left(\frac{q_{3a}(t)}{v_{3a}} + \frac{q_{3b}(t)}{v_{3b}}\right) \\
\dot{q_{4a}}(t) &= -k_{14a}q_{4a}(t) + k_{4a3a}q_{3a}(t) - k_{d2}\left(\frac{q_{4a}(t)}{v_{4a}} - \frac{q_{4b}(t)}{v_{4b}}\right) \\
&\quad - \frac{uf_2}{2}\left(\frac{q_{4a}(t)}{v_{4a}} + \frac{q_{4b}(t)}{v_{4b}}\right) \\
\dot{q_{4b}}(t) &= -k_{3b34}q_{4b}(t) + k_{d2}\left(\frac{q_{4a}(t)}{v_{4a}} - \frac{q_{4b}(t)}{v_{4b}}\right) \\
&\quad + \frac{uf_2}{2}\left(\frac{q_{4a}(t)}{v_{4a}} + \frac{q_{4b}(t)}{v_{4b}}\right)
\end{aligned}
\tag{4.34}
$$

$$q_1(0) = c_{10}v_1 \qquad q_2(0) = c_{10}v_2$$
$$q_{3a}(0) = c_{10}v_{3a} \quad q_{3b}(0) = 0$$
$$q_{4a}(0) = c_{10}v_{4a} \quad q_{4b}(0) = 0$$

with all parameters as previously defined and where uf_1 and uf_2 are the ultrafiltration rates of the first and second filter respectively; similarly, k_{d1} and k_{d2} are the diffusive rate constants for each filter. It should be noted that in the above model the direction of ultrafiltration is assumed to be in the same direction for both filters; however, positivity is not enforced on either uf_1 or uf_2 during the parameter estimation process, thus allowing the flow direction to be determined during the parameter estimation process.

The transfer rate between filters is determined by the volumetric flow rate between them. It is therefore possible to reduce the number of parameters by redefining the transfer rates in terms of the flow rates and ultrafiltration losses. As the plasma (Q_p) and dialysate (Q_d) flow-rates are provided by clinical staff along with the FLC concentration measurements, the transfer rates between filters are modified by the loss or addition of the ultrafiltration flow, that is:

$$\begin{aligned} k_{4b3b} &= \frac{Q_d + uf_2}{v_{4b}}, \\ k_{03b} &= \frac{Q_d + (uf_1 + uf_2)}{v_{3b}}, \\ k_{3a1} &= \frac{Q_p}{v_1}, \\ k_{4a3a} &= \frac{Q_p - uf_1}{v_{3a}}, \end{aligned} \tag{4.35}$$

and

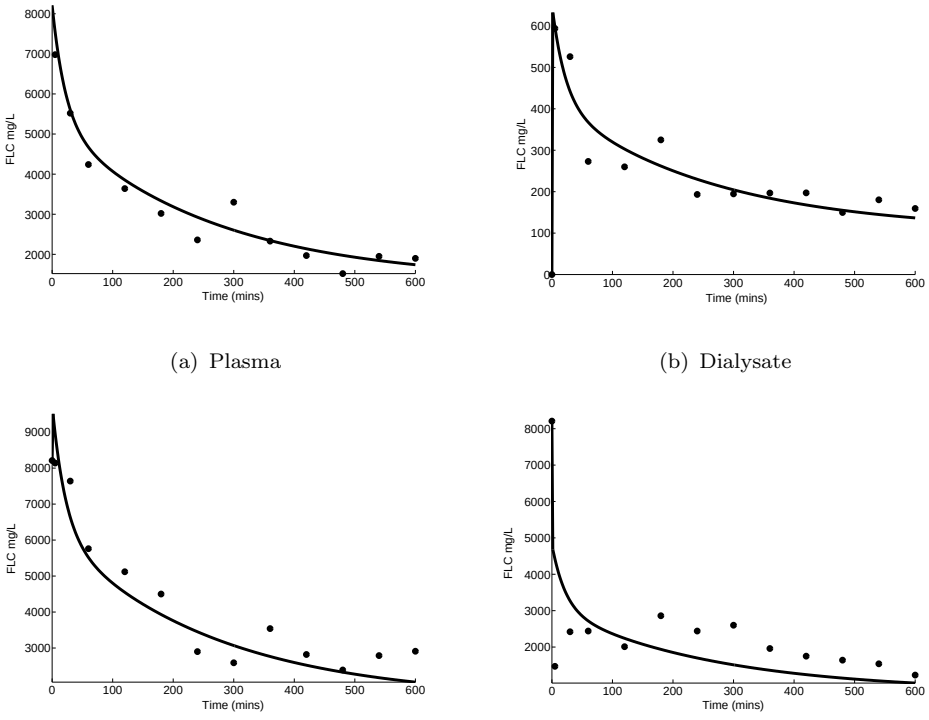
$$k_{14a} = \frac{Q_p - (uf_1 + uf_2)}{v_{4a}}.$$

In addition, the volume of fluid within the filters is constant and measurable, as it is provided by the manufacturers of the filters [Toray, 2006]. For the filter used in the validation below, the fluid volume was 70 millilitres for plasma and dialysate (i.e. $v_{4a} = v_{3a} = v_{4a} = v_{4b} = 0.07$). The filters used in this trial were the same make and

model, it is therefore assumed that the filters have the same diffusion characteristic ($k_{d1} = k_{d2} = k_d$), so a reduced parameter vector is available, i.e.

$$p_2 = \ [c_{10}, v_1, k_{12}, k_{21}, k_{re}, f\ k_d, uf_1, uf_2]^T.$$

The analysis was performed to ensure identifiability of the parameter vector p_2, and excluding the initial condition (c_{10}), the model was found to be globally identifiable. Unfortunately, due to computational limitations, if the initial condition is added to this set, identifiability results are not achievable. Therefore, the initial condition (c_{10}) is assumed known, and taken as the pre-treatment FLC concentration measurement. The DAISY input file and results for this analysis can be found in Appendix A.



(a) Plasma

(b) Dialysate

(c) Filter 1

(d) Filter 2

Figure 4.14: *Validation of simulated results against patient data for a dual-filter haemodialysis session; solid line - simulated FLC concentration; dots - measured values. Figure a) shows the FLC concentration in plasma; b) the FLC concentration leaving filter 1 in the dialysate; c) the FLC concentration between filters 1 and 2; d) FLC returned to the patient from filter 2.*

Data were available for a single patient who was treated with two dialysis filters

connected in series (as shown in Figure 4.11). Figure 4.14 shows a comparison of the simulation results and measured FLC values seen in plasma, dialysis filters and dialysate for the double filter configuration. The FLC plasma concentration can be seen in Figure 4.14. Figure 4.14(b) shows the comparison of measurements in the dialysate fluid on exit from the final dialyser; whilst the remaining two, Figures 4.14(c) and 4.14(d), describe the plasma-side concentration of each filter.

The model predicts the observed behaviour accurately, with the exception of initial measurements of filter two, for which the model fails to estimate the sharp decline in concentration within the first few minutes. It was not possible to have this sample re-calculated and it is not known if it is a true measurement or an outlier that should be ignored; it has therefore been left in the dataset. The plasma concentration is similar in nature to that already seen in the FLC model validation above. The simulations of the dialysate and between filter measurements, however, exhibit some surprising features. In filter one an increase is seen in the concentration from the initial starting value, whilst the reverse is seen in filter two. The dialysate concentration increases drastically in the first few minutes from its initial zero concentration to a maximum of approximately 600 mg/L. Further investigation is required to determine why these effects occur and if they can be exploited to assist in clearance.

The parameter estimates obtained by fitting to the data in Figure 4.14 are displayed in Table 4.13. The transfer rate between plasma and EVF (k_{21}) shows a similar value to that seen previously in the FLC model. The plasma volume is within the same range, approximately 2L. Using k_{21}, k_{12} and v_1 results in an estimated EVF volume of approximately 6L which suggests the previous approximation of transfer rate constants ratio being equal to the patient volume ratio is correct. However, the value obtained for reticuloendothelial clearance, based on the single compartment half-life of FLC may be underestimated, with the value shown here an order of magnitude greater. Of greater interest are the estimates given for the diffusive clearance and ultrafiltration rates, with the results suggesting that flow across the membrane due to ultrafiltration is much greater than that offered by diffusion; this agrees with the finding of Eloot et al. [2005], where *in vitro* tests were conducted to determine the effects of middle-molecule clearance through high-flux filter membranes. The two ultrafiltration rates are in opposition to each other, as can be determined by the sign, and of similar magnitudes,

Parameter	Description (Units)	Value (SDLN)
v_1	Plasma volume (L).	1.92 (0.18)
k_{21}	FLC transfer constant between plasma and EVF (min^{-1}).	0.011 (0.58)
k_{12}	FLC transfer constant between EVF and plasma (min^{-1}).	0.003 (0.23)
k_{re}	Clearance via reticuloenditilium system (min^{-1}).	2.49e-4 (0.31)
k_d	diffusive transfer across membrane (L/min).	0.010 (0.19)
f	FLC generation (mg/min).	22.66 (0.51)
uf_1	Ultrafiltration rate from first filter (L/min).	0.041 (0.64)
uf_2	Ultrafilteration rate from second filter (L/min).	-0.035 (0.61)

Table 4.13: *Parameter estimates (and SDLNs) for extended filter model, two filter configuration.*

which would prevent the dialysis monitors from detecting a net ultrafiltration rate. However, with only a single dataset available to validate these findings, care must be taken not to place too much emphasis on them at present.

4.6 Summary

In this chapter methods and techniques have been developed that allow the analysis of the *in vivo* kinetics of Free Light Chains. Three models have been presented:

1. A single compartment system to model the effects of chemotherapy on FLC MM patients over an extended evaluation periods of months.

2. A two compartment model that describes the impact of increasing the natural clearance rate via haemodialysis over a shorter time scale, i.e. hours.

3. An extendable multi compartment model to simuate the use of a multi-filter in a single dialysis session.

In addition, simple functions that describe quantitatively the impact of cytotoxics on the underlying FLC synthesis have been documented. Whilst very simple, the single

compartment model is beneficial in analysing the effectiveness of different combinations of drug treatments, that would be difficult to assess using only plasma concentration measurements. The second models were then used *de novo* to compare treatment for FLC patients, from which the benefits of using haemodialysis are clearly evident, offering indications of the ability of different manufacturers' proto-type high-flux filters to remove FLC for MM patients during haemodialysis. It was also shown that regular and repeated dialysis regimes are required to maintain the desired levels of FLC allowing patient recovery that would not be possible to achieve using more traditional apheresis methods, namely plasma exchange. However, whilst haemodialysis can assist in maintaining lower levels of FLC concentration, if it is combined with successful chemotherapy significant improvements in clearance levels are seen. In addition, the models have also been used to describe the impact of altering the application of chemotherapy, with the simple rule of initiating chemotherapy as soon as possible on presentation and avoiding interruptions in the treatment. However, in situations when interruptions are known or planned beforehand, the application of haemodialysis can assist in maintaining significantly lower levels of FLC. The effect of kidney recovery on FLC clearance has also been demonstrated, with even modest levels of recovery of the natural kidney function yielding significantly improved whole-body clearance.

An extension to the standard compartmental model developed by Evans et al. [2006] has been described. The extended model incorporates measurements of inter-filter and dialysate side concentrations. The extended model allowed assumptions made previously regarding parameter values to be validated and highlighted the potential recirculatory ultrafiltration issue that may account for anomalies seen in patient clearance when using filter combinations. Whilst it is evident that performing such extensive measurements on patients is neither financially justifiable nor possible in all clinical environments, conducting such detailed studies for initial haemodialysis treatments may offer insight into a particular patient's characteristics; enabling a better understand of treatment effectiveness. However, consideration must be given to intra- and inter-patient variability and the assumption that FLC synthesis and body volume is constant throughout the treatment. As all patients in this study are undergoing chemotherapy and, with falling renal output, ultrafiltration these assumption may be invalid.

Due to the underlying dynamics, FLCs lend themselves to the linear modelling techniques presented above, as they are inactive *in vitro*, in that they do not bind readily to other material, and the clearance routes can assumed be to be linear. For patients with Intact Secreting Myeloma or transplant subjects with healthy immune response, this is not the case. Therefore in the following chapter the kinetics of whole antibodies are investigated to reveal more complex dynamics resulting in nonlinear system models that build upon the simple single and two compartment models presented in this chapter.

Chapter 5

Kinetics of Intact Immunoglobulin IgG

Immunoglobulin G are known to be recycled through the Brambell, or FcRn, receptor, with the effect of prolonging the half-life of IgG circulating in plasma [Waldman and Strober, 1969]. In patients with Intact Secreting Immunoglobulin Multiple Myeloma, monoclonal Immunoglobulin G is produced in excess when compared to normal antibody production. In this chapter a mathematical model, incorporating nonlinear binding kinetics and delayed-logistic growth, is described, that relates the interaction of tumourous and non-tumourous immunoglobulin on the FcRn receptor. The model is validated against data from patients with Hypercatabolic Hypoproteinemia[1], Multiple Myeloma and normal subject kinetics. The results show the effects of a dynamic concentration-dependent catabolism on IgG concentrations and highlight the differences between perceived and actual tumour decay under the influence of chemotherapy.

As with FLC Myeloma patients, chemotherapy is given to patients with Intact IgG Myeloma to encourage cell-death of the tumourous plasma cells and ultimately reduce the concentration of tumourous IgG to normal levels; however, due to the size of the whole IgG antibodies haemodialysis is not available as a treatment. Other forms of apheresis are used to clear whole antibody, these are covered in the following chapter (Chapter 6). To determine the effectiveness of the chemotherapy on Intact IgG Myeloma, the concentration of total IgG is measured in plasma. The first section of

[1] A disease resulting in a lack of FcRn [Waldmann and Terry, 1990]

this chapter uses a mathematical model of IgG binding to investigate the effects of the tumour and non-tumourous IgG competing for the same receptor under the influence of chemotherapy. In the second section the benefits of calculating the tumour kill through the measurement of alternative IgG markers are analysed, using the model previously developed, through which it is shown that, given applicable assumptions regarding the treatment, measuring the ratio of tumourous and non-tumourous IgG may present a better marker than measuring plasma concentrations alone.

5.1 Model describing the recycling IgG

Immunoglobulins are removed from plasma via a process of pinocytosis, which would yield an effective *in vivo* half-life similar to that of FLC, i.e. between three and five days. However, IgG is known to have a much greater half-life, approximately 21 days in normal subjects, though it is frequently shorter in MM patients [Bradwell et al., 2005, pages 90-105]. The reason for this extended retention is due to receptors known as the 'Brambell Receptors'; the existence of which was first theorised by Brambell et al. [1964]. The receptors exist in the plasma membrane, in the endosomes[2], and prevent a portion of IgG from being catabolised through binding to the receptor. Any free, or non-bound, IgG is destroyed whilst those 'selected' by the receptors are recycled [Junghans and Anderson, 1996; Mariani and Strober, 1990]. In addition, this receptor is also responsible for transportation of IgG between mother and infant during pregnancy and is therefore referred to as the FcRn (Fc Receptor-neonatal) [Ghetie and Ward, 2002].

5.2 The model

As seen in the previous chapters, FLC are produced in two variants: Kappa and Lambda. A single antibody is constructed of only one type of FLC; as such two variants of IgG exists: IgG-Kappa (IgGK) and IgG-Lambda (IgGL). Multiple Myeloma patients will produce only a single monoclonal IgG (either Kappa or Lambda) far in excess of all other antibodies. The FcRn receptor does not differentiate between the types of

[2]Small vesicles within the plasma membrane [Metzer, 1990, ch. 3].

IgG it binds to, therefore all IgG are effectively 'competing' for the free FcRn to prevent destruction. It is therefore necessary to review the mathematics surrounding the kinetics of receptor and ligand binding in a competitive environment. This will subsequently be used in the construction of a complete metabolic model.

5.2.1 Competitive binding kinetics

The most common form used to represent receptor kinetics is to consider the process as a bimolecular reaction [Foreman and Johansen, 2002]; if the two ligands do not interact the competitive process can be expressed by the following chemical equations:

$$\mathrm{L}_1 + \mathrm{R} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \mathrm{L}_1\mathrm{R}$$
$$\mathrm{L}_2 + \mathrm{R} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \mathrm{L}_2\mathrm{R},$$

where L_1 and L_2 represent the free ligands, R the free receptor concentration and L_iR the bound receptor-ligand complex. The rate at which the molecules bind is governed by k_a (association rate - the rate at which the ligand binds to a receptor site, in M^{-1} min^{-1}) and k_d (the dissociation rate constant - the rate at which it is released, in min^{-1}). A schematic representation of this reaction can be seen in Figure 5.1.

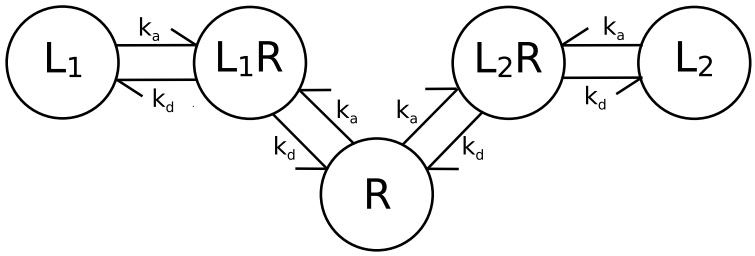


Figure 5.1: *Model of receptor binding.*

If receptor binding is considered to obey the law of mass-action, the rate of the reaction is directly proportional to the product of the concentrations. For two ligands binding to a single receptor, this leads to a five state model (Figure 5.1) that describes the reaction, where the hook-arrow represents kinetic binding, as opposed to material flow as seen previously. This model results in the following set of ordinary differential equations [Sklar et al., 1985], square brackets, e.g. [R], indicate concentration as opposed to quantity:

$$
\begin{aligned}
[\dot{L_1}] &= -k_a[L_1][R] + k_d[L_1R] \\
[\dot{L_1R}] &= k_a[L_1][R] - k_d[L_1R] \\
[\dot{L_2}] &= -k_a[L_2][R] + k_d[L_2R] \\
[\dot{L_2R}] &= k_a[L_2][R] - k_d[L_2R] \\
[\dot{R}] &= k_d[L_1R] - k_a[L_1][R] + k_d[L_2R] - k_a[L_2][R]
\end{aligned} \tag{5.1}
$$

However, as shall be seen later (section 5.3) this model can be significantly reduced using basic laws of conservation. It should be noted that Gurbaxani and Morrison [2006] give strong evidence for the existence of more complex interactions than a single-site receptor binding to IgG; however, this is not considered in this work due to the level of granularity of the model and the data available for validation.

5.2.2 Full metabolic model

In order to construct a model, the biology of IgG was reviewed, then mass-balance principles were applied. It was constructed without considering constraints on measurable variables or availability of known parameters; assumptions will be applied later to produce a more tractable model for simulation. A schematic of the full kinetic model is shown in Figure 5.2.

On first inspection the model is clearly symmetrical as the IgG Kappa (IgGK) and IgG Lambda (IgGL) proteins are treated identically by the metabolic processes. All compartments labelled K represent the quantity of IgG Kappa, whilst L denotes the quantity of IgG Lambda. The model was initially constructed by considering the flow of IgG in plasma. It is known, through radio-nuclide marking [Waldman and Strober, 1969], that IgG moves between plasma and extravascular (EVF) pools. Therefore, two compartments were created for each IgG type (K_1,L_1 representing plasma and K_2,L_2 EVF quantities) with two linear rate constants (k_{12} and k_{21} describing the flow across the plasma membrane. The flow between plasma/EVF is considered to be the same process as EVF and plasma; it is therefore possible to define this exchange by a single parameter, with respect to the volume ratio between the two compartments $k_{21} = \frac{v_1}{v_2}k_{12}$, where v_i is the apparent volume of distribution of compartment i. However,

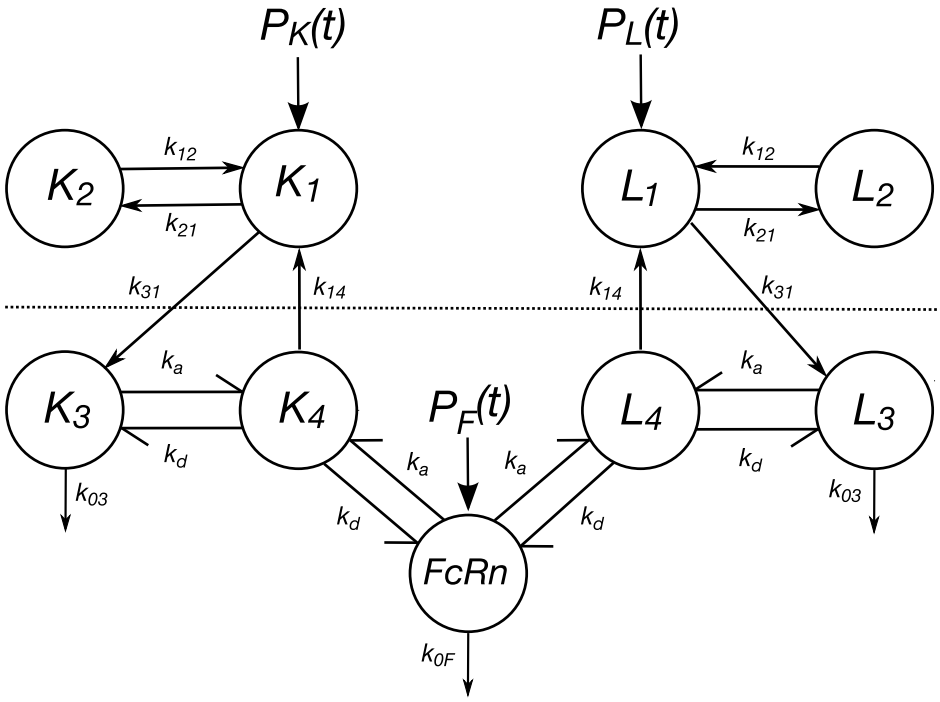


Figure 5.2: *IgG full model. IgGK (K) and IgGL (L) with the subscripts i, i.e K_i and L_i representing the following compartments: 1) Plasma, 2) EVF, 3) Free in the endosome, 4) bound in the endosome, and FcRn) bound to FcRn in the endosome.*

from radiolabelling experiments Waldman and Strober [1969] have demonstrated that for IgG the volume of distributions are approximately equal ($v_1 \approx v_2$). The synthesis of intact IgG is controlled by two time-dependent functions: $P_K(t)$ (for IgGK) and $P_L(t)$ (for IgGL).

Both IgGK and IgGL are continuously exchanged between plasma and the endosome through pinocytosis. This is represented by the linear rate constant k_{31}. The K_3 and L_3 compartments represents free IgG in the endosome, from which IgG is either cleared from the system, rate constant k_{03}, or bound to FcRn (bound IgG is represented by K_4 and L_4). The quantity of free receptors in the endosome is represented by the $FcRn$ compartment. FcRn receptors are also catabolised in the endosome; this is represented by the clearance of FcRn (k_{0F}) and synthesis of new receptors by the input function, $P_F(t)$. A concise list of the parameters and model nomenclature can be seen in Table 5.1.

Through the model the clearance mechanism of IgG can be easily described. The IgG is allow to move freely between plasma and EVF (k_{12} and k_{21}) but is passed into

Table 5.1: *IgG and FcRn competitive binding parameters description.*

Variable	Description
$P_K(t)$,$P_L(t)$ and $P_F(t)$	Production of IgGK, IgGL and FcRn (mg min^{-1})
K_i or L_i	Quantity of Kappa or Lambda in compartment i (mg)
$FcRn$	Quantity of FcRn in the Endosome (mg)
y_i	Observation of compartment i, c_i will be used as observation gain, in all cases this $\frac{1}{v_i}$
k_{ij}	Rate Constant of flow of material from compartment j to compartment i (min^{-1})
k_{0j}	Rate Constant removal of material from compartment j (min^{-1})
k_a	Association rate of IgG and FcRn receptor (mg min^{-1})
k_d	Dissociation rate of IgG and FcRn receptor (min^{-1})
v_i	Volume of compartment i (1=plasma, 2=extra-vascular and 3=endosome) (L)

and out of the endosomes via pinocytosis (k_{31} and k_{14}). If an individual IgG is bound to an FcRn receptor it is returned to plasma, whilst if it is not bound it is eliminated from the endosome (k_{03}).

The system of equations generated from the model (Figure 5.2) can be seen in equations (5.2). Only the equations for IgGK, with the equation for free FcRn receptors, are shown but due to the symmetry in the model, the IgGL variant can be derived by replacing K for L in each case.

$$\begin{aligned}
\dot{K_1}(t) &= -(k_{21}+k_{31})K_1(t) + k_{12}K_2(t) + k_{14}K_4(t) + P_K(t) \\
\dot{K_2}(t) &= k_{21}K_1(t) - k_{12}K_2(t) \\
\dot{K_3}(t) &= -k_b K_3(t) FcRn(t) + k_d K_4(t) - k_{03}K_3(t) + k_{31}K_1(t) \\
\dot{K_4}(t) &= k_b K_3(t) FcRn(t) - k_d K_4(t) - k_{14}K_4(t) \\
\dot{FcRn}(t) &= -k_b(K_3(t)+L_3(t))FcRn(t) + k_d(K_4(t)+L_4(t)) \\
&\quad - k_{0F}FcRn(t) + P_F(t)
\end{aligned} \tag{5.2}$$

The association constant (k_a) in equation 5.2 has been replaced by k_b, where $k_b = \frac{k_a}{v_3}$, to convert the concentrations as shown in equation 5.1 and quantity used above.

On clinical presentation, prior to treatment, it is assumed that the patient's tumour has produced sufficient IgG to saturate all compartments and receptor sites; with the system therefore in steady-state the initial conditions are (once again only IgGK examples):

$$\begin{aligned} K_1(0) &= C_{K10}\, v_1 & K_2(0) &= C_{K20}\, v_2 \\ K_3(0) &= C_{K10}\, v_3 & K_4(0) &= C_{K40}\, v_3 \\ FcRn(0) &= FcRn_0\, v_3 \end{aligned}$$

where C_{K10} represents the concentration of the IgGK in plasma prior to any chemotherapy treatment. Due to the steady-state assumption the concentration difference between plasma and EVF has equilibrated; in addition, as pinocytosis is a sampling process it can be assumed that IgG free in the endosome is of equal concentration to that in plasma.

Whilst in a restricted laboratory experiment it may be possible to take measurements relating to numerous aspects of the above model, the expectation of this modelling study is that it is used in a clinical environment and as such the only measurements available are those that can be taken *in vivo* and hence, are restricted to concentrations in plasma allowing the following measurements to be taken:

$$\begin{aligned} y_{K1}(t) &= c_1 K_1(t), \\ y_{L1}(t) &= c_1 L_1(t) \text{ or} \\ y_{\Sigma}(t) &= c_1 (K_1(t) + L_1(t)). \end{aligned} \tag{5.3}$$

where c_1 represents an observation gain. It is possible to observe IgGK (y_{K1}) and IgGL (y_{L1}) individually, and the total IgG present (y_{Σ}). As with FLC it is assumed that all IgGK and IgGL are captured by the measurement technique and as such equates c_1 to the volume of distribution in plasma, $c_1 = \frac{1}{v_1}$.

5.3 Model reduction

The above model is a comprehensive description of the process involved in IgG recycling, and it is extremely difficult to obtain information regarding key parameters, even initial estimates from literature. For example, the volume of the endosome would require knowledge of the average number and approximate volumes of the vesicles. Therefore, a simplified model is derived through the application of assumptions regarding the kinetic behaviour of the receptors.

If the total concentration of receptors is considered constant throughout the duration of the treatment, the total receptors available, R_T, is then given by:

$$\begin{aligned} R_T &= K_4(t) + L_4(t) + FcRn(t) \text{ or} && (5.4) \\ FcRn(t) &= R_T - (K_4(t) + L_4(t)) && (5.5) \end{aligned}$$

If equation (5.5) is applied to the full system model (Equation (5.2)) the quantity of IgG bound in the endosome can be rewritten as

$$\begin{aligned} \dot{K}_4(t) = &\quad k_b(R_T - (K_4(t) + L_4(t)))K_3(t) - (k_d + k_{14})K_4(t) \\ \dot{L}_4(t) = &\quad k_b(R_T - (K_4(t) + L_4(t)))L_3(t) - (k_d + k_{14})L_4(t) \end{aligned} \tag{5.6}$$

Further, if a quasi-steady state assumption is made regarding the bound IgG, namely that the bound IgG dynamics are much faster than the free IgG in the endosome, the IgG bound in the endosome can be considered in steady-state (i.e. $\dot{K}_4(t) = \dot{L}_4(t) = 0$). Equations (5.6) can thus be reduced to

$$K_4(t) = \frac{k_b(R_T - L_4(t))K_3(t)}{k_d + k_{14} + k_bK_3(t)} \text{ and } L_4(t) = \frac{k_b(R_T - K_4(t))L_3(t)}{k_d + k_{14} + k_bL_3(t)}, \tag{5.7}$$

which with re-arrangement gives

$$K_4(t) = \frac{k_bR_TK_3(t)}{k_d + k_{14} + k_b(K_3(t) + L_3(t))} \text{ and } L_4(t) = \frac{k_bR_TL_3(t)}{k_d + k_{14} + k_b(K_3(t) + L_3(t))}. \tag{5.8}$$

Focusing on K_4, as stated above, pinocytosis is assumed to be a constant extraction and recycling process, which ensures that the concentrations of IgG in the endosome and plasma are constant ($\frac{K_3(t)}{v_3} = \frac{K_1(t)}{v_1}$ and $\frac{L_3(t)}{v_3} = \frac{L_1(t)}{v_1}$). Then equation (5.8) can be re-written as

$$K_4(t) = \frac{k_b R_T \frac{v_3}{v_1} K_1(t)}{k_d + k_{14} + k_b \frac{v_3}{v_1}(K_1(t) + L_1(t))}. \tag{5.9}$$

If this is re-written in a more familiar Michaelis-Menten form, then

$$k_{14}\, K_4(t) = \frac{V_m K_1(t)}{K_m + (K_1(t) + L_1(t))} \tag{5.10}$$

where $V_m = k_{14}\ R_T$ and $K_m = \frac{v_1(k_{14}+k_d)}{k_b v_3}$. As only the bound IgG in equation (5.10) occurs into the plasma compartment of the system equations (5.2)

$$\begin{aligned} \dot{K_1}(t) &= -\left(k_{31} - \frac{V_m}{K_m + (K_1(t) + L_1(t))}\right) K_1(t) - k_{21} K_1(t) \\ &+ k_{12} K_2(t) + P_K(t) \end{aligned} \tag{5.11}$$

and applying the same process to the IgGL in plasma equation and yields

$$\begin{aligned} \dot{L_1}(t) &= -\left(k_{31} - \frac{V_m}{K_m + (K_1(t) + L_1(t))}\right) L_1(t) - k_{21} L_1(t) \\ &+ k_{12} L_2(t) + P_L(t) \end{aligned} \tag{5.12}$$

The first term in these equations $\left(k_{31} - \frac{V_m}{K_m+(K_1+L_1)}\right)$ is referred to as the Fractional Catabolic Rate [Waldman and Strober, 1969], and represents the fraction of the available plasma pool cleared per unit time.

5.3.1 Tumour synthesis of IgG

As seen with FLC patients, patients with Myeloma will have chemotherapy to destroy the tumour-producing plasma cells. It is therefore necessary to consider the dynamic effects of chemotherapy; for FLC patients this was modelled by a simple exponential decay to a baseline. When applied to data for a patient with Intact Immunoglobulin Multiple Myeloma the results were not appropriate; through the use of deconvolution techniques (see Chapter 7) it could be seen that the underlying synthesis was not responding to chemotherapy in the same manner seen by FLC MM; therefore, an alternative functional form for the tumour cell numbers and monoclonal IgG synthesis is required.

Logistic curves have been used to model tumour growth [Forys and Marciniak-Czochra, 2003] during chemotherapy. These models were used to describe the survival of tumour cells, rather than the substance produced by the tumour. These logistic models were then validated with measurements for the number of tumour cells, estimated from imaging the tumour mass and average cell numbers per unit volume of tumour. In MM a measure of the number of plasma cells killed during chemotherapy is only possible through an invasive biopsy and is a parameter that is not often available. However, the logistic curve equation parameters are not directly related to specific tumour properties, only to the initial condition, growth rate and maximum final value. As it is known from *in vitro* experiments [Sullivan and Salmon, 1972], a single cancerous cell produces IgG at a constant rate; therefore it is possible to use the same logistic equation approach to simulate IgG synthesis under chemotherapy treatment, up to a constant scaling factor. The synthesis is therefore modelled by the logistic equation and the rate of change of production is given by:

$$\dot{P_K}(t) = \quad rP_K(t)\left(1 - \frac{P_K(t)}{K}\right), \quad P_K(0) = P_{K0} \tag{5.13}$$

where r is the growth rate of cells (which may be negative for effective tumour kill) and K the carrying capacity. When applied to patients undergoing chemotherapy treatment two addition phenomena must be recognised:

- The chemotherapy may not affect the tumour sythnesis immediately, i.e. at time

$t = 0$. This may be due to the behaviour of the drug or the effects on the individual patient.

- The treatment may not be 100 % effective, with some baseline production of the tumourous IgG type still evident.

The first of these issues leads naturally to the use of a delayed logistic equation, a solution for which is given in Forys and Marciniak-Czochra [2003], with the addition of basal terms for minimum production, gives

$$P_K(t) = \frac{K}{1 + e^{r(t-\tau)}} + \beta, \tag{5.14}$$

where β is the basal production, τ the delay effect, or time shift. The baseline production may be solely due to normal (non-tumourous) cells if the chemotherapy is 100 % effective; however, it is more likely to be a constant production combining normal and tumourous cells. The non-tumourous production is assumed to be unaffected by the chemotherapy. The delay in equation (5.14) provides the possibility of a non-instantaneous response to the chemotherapy. If an instant response to chemotherapy is evident then $\tau = 0$.

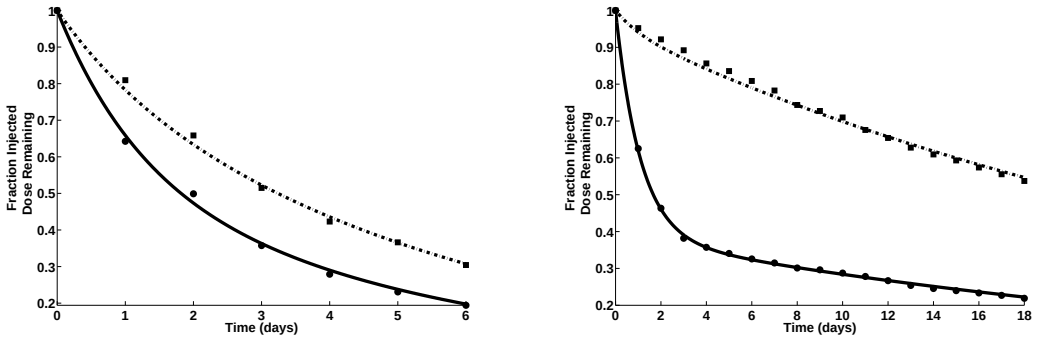
There are other functions that provide greater flexibility to describe tumour dynamics, e.g. Richard's Curve [Richards, 1959]; however, difficulties arise when fitting this function to data due to over-parameterisation [Karkach, 2006], as such the simpler logistic function is preferred; as it offers sufficient flexibility, capturing a range of production dynamics by altering just four parameters (K, β, τ and r).

5.3.2 Model parameters

Hypercatabolic Hypoproteinemia is a rare medical condition that results in a patient that does not express FcRn; Waldmann and Terry [1990] measured proteins of two patients with this condition. Markers were injected into these patients and the plasma and whole body concentrations of IgG were measured; whole body concentrations were determined from renal clearance of the radio marker. If total IgG is considered, combining equations (5.11) and (5.12) and removing the effects of the FcRn receptors, this results in a standard linear two-compartment model.

$$\begin{aligned} \dot{IgGT_1}(t) &= -(k_{01} + k_{21})IgGT_1(t) + k_{12}IgGT_2(t) \\ \dot{IgGT_2}(t) &= k_{21}IgGT_1(t) - k_{12}IgGT_2(t) \end{aligned} \tag{5.15}$$

where the subscripts 1 and 2 denote plasma and EVF pools respectively, $IgGT_i(t)$ is the total IgG ($IgGK_i(t) + IgGL_i(t)$) in compartment i at time t, k_{01} is the clearance from the system and the remaining parameters are as defined previously. As the study used radio-labelled IgG the production term for IgG is zero. Equations (5.15) were fitted to the data described in Waldmann and Terry [1990] to enable approximate values to be obtained for the parameters k_{01}, k_{21} and k_{12}; for patients with Hypoproteinemia the apparent volume of distribution is not considered equal Waldmann and Terry [1990], so k_{12} and k_{21} are estimated separately. The results of these fits can be seen in Figure 5.3 and the values obtained are displayed in Table 5.2.



(a) Hypoproteinemia Patient (b) 'Normal' individual

Figure 5.3: *Fits of equation 5.15 to [Waldmann and Terry, 1990];* □*=whole body IgG,* ○*= plasma IgG.*

The parameters obtained for k_{01} by applying the model to Hypoproteinemia patients provide a useful measure of the Fractional Catabolic Rate (FCR). For 'normal' subjects the value equates to the FCR, with recycling, at normal IgG concentration, approximately 12 g/L on average [Waldmann and Terry, 1990]. However, for Hypoproteinemia patients (with the FcRn missing) it is the maximum clearance possible by the FCR, i.e. k_{31}. The value distinctly shows the benefit of the FcRn receptor in maintaining high levels of IgG in plasma by ensuring a low clearance rate.

The results of fitting the simplified model (eqns. (5.11) and (5.12)) with logistic production (eq (5.13)), to patient data can be seen in Figure 5.4; it should be noted that

Table 5.2: *Model parameters for Hypoproteinemia patients and normal subjects, with SDLN. The model was checked to ensure all parameters were uniquely globally identifiable, see Appendix A.2.1.*

Parameter	Hypoproteinemia	Normal
(days^{-1})	Value (SDLN)	Value (SDLN)
k_{21}	0.20 (0.17)	0.51 (0.04)
k_{12}	0.54 (0.29)	0.41 (0.04)
k_{01}	0.28 (0.02)	0.07 (0.01)

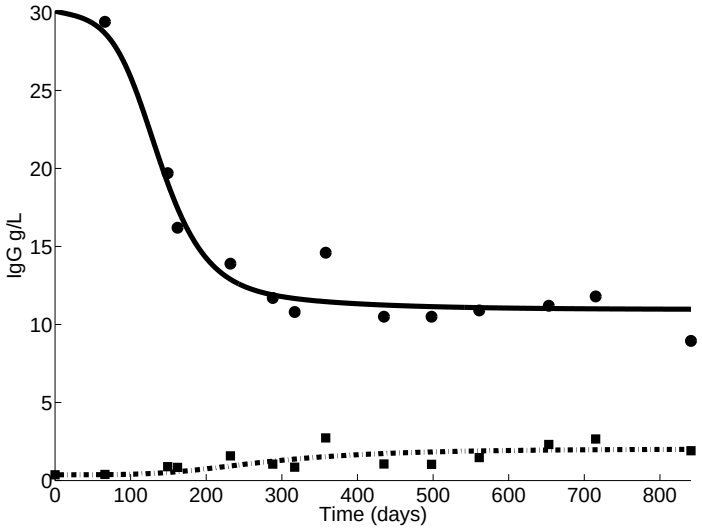


Figure 5.4: *Parameter estimation of IgG-FcRn and tumour synthesis parameters. Data: □-Lambda (non-tumourous) ○-Kappa (tumourous). Simulation: IgGK (solid line) and IgGL (chained-line).*

although the system appears to be numerically identifiable, i.e. the fitting procedure is sensitive to the parameters chosen, a formal identifiability analysis has not been completed. The reduced recycling model with the logistic input is a nonlinear time-delayed system and the identifiability analysis is a non-trivial problem. Both the Taylor series and differential algebra methods have been used but neither approach was able to produce informative results. This is the subject of ongoing research.

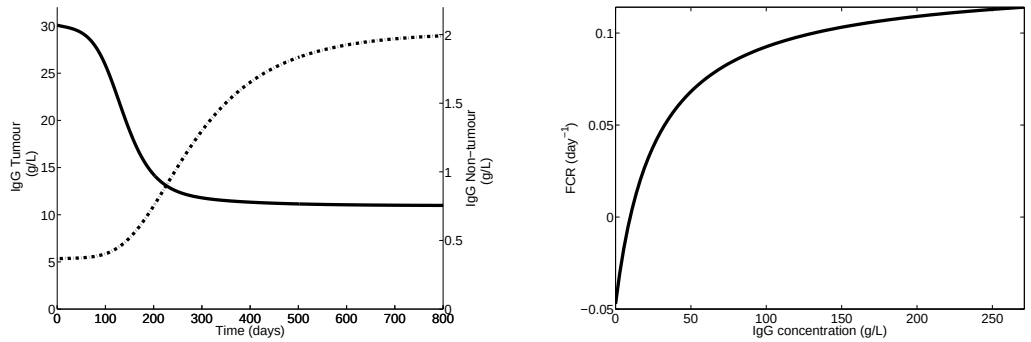
Table 5.3: *Model parameters for intact secreting MM patient, with SDLN for fitted parameters.*

Parameter	Description	Value (SDLN)
k_{21}	Plasma/EVF inter-compartment transfer (days^{-1}).	0.51 (N/A)
k_{12}	EVF/Plasma inter-compartment transfer (days^{-1}).	0.41 (N/A)
v_1	Plasma volume (L)	2.59 (N/A)
k_{31}	Maximum clearance rate of IgG through endosome (days^{-1}).	0.13 (0.31)
V_m	Maximum rate of FcRn recycling (g/day)	15.51 (0.29)
K_m	The quantity at which half the V_m is achieved (g).	79.29 (0.18)
r	Decay rate of the tumour (% per day).	0.03 (0.42)
τ	Chemotherapy delay effect (days)	98.11 (0.19)
β	Baseline production of tumourous IgG (g/day)	0.18 (0.12)

The patient for whom the data were collected had an IgGK secreting tumour and received treatment on the day of the first measurement. The simulation results show (Figure 5.4) excellent approximations to the data on visual inspection. The fall in IgGK is clearly evident and is dominated by the IgGK tumourous synthesis rate.

5.3.3 Predicted clinical response

Figure 5.5(a) compares the plasma concentrations over the treatment period for the patient given in Table 5.3, in which an interesting effect of the concentration dependent clearance can be seen. The tumourous IgG declines due to the chemotherapy killing the cancerous plasma cells and thus restricting production, as expected; however, the non-tumourous IgG, IgG-Lambda in this case, rises from around 400 mg/L to 2 g/L,



(a) Simulation results

(b) FCR

Figure 5.5: *Simualated results using parameters estimates in Table 5.3. Figure a) Simulation results of tumourous (IgGK - solid line) and non-tumourous (IgGL - chained line) concentrations in plasma over the time course of the chemotherapy treatment. Figure b) Esimated Fractional Catabolic Rate (eq. 5.11) for the patient parameters specified in Table 5.3)*

even though the non-cancerous synthesis remains constant throughout the treatment. The increase is due to the decrease in natural clearance (FCR) over the time course, effective in both tumourous and non-tumourous IgG concentrations. The FCR defined by the paramters in Table 5.3 is shown in Figure 5.5(b), the similarities between this and the FCR as measured by Waldmann and Terry [1990] (see Figure 2.3 in Chapter 2) can clearly been seen. Figure 5.5(b) shows two facets of the model that should be mentioned: firstly, that over the range of concentrations shown for the patient data considered (10-30g/L) the IgG clearance could be approximated by a simple linear function; secondly, that for low concentrations the FCR predicts a negative clearance, which by definition is not feasible.

In order to determine the effectiveness of chemotherapy, clinicians measure changes in concentration of the tumourous IgG in plasma. The change in concentration will be directly related to the concentration fall in plasma, the tumour kill estimated and ultimately how the patient is categorised clinically in terms of their response, e.g. a good response requiring no additional chemotherapy or a poor response which requires further treatment. The accuracy of this categorisation should therefore be of great significance to clinicians and patients alike. In Figure 5.6 a comparison[3] is shown of

[3]All simulations have been normalised to the maximum value, in all cases this is the value at time $t = 0$.

the change in total plasma tumourous IgG concentration, against that of IgG tumour production. For this patient, after the initial period, using the change in plasma IgG measurement underestimates the tumour kill, and at steady-state results in a 35% difference.

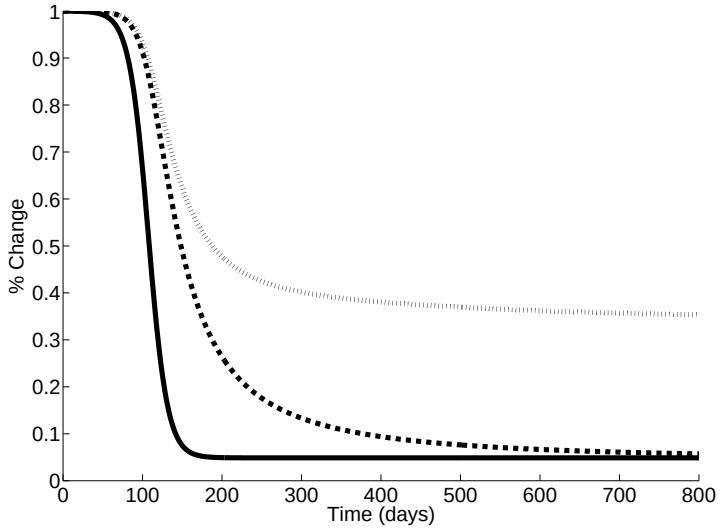


Figure 5.6: *Comparison of IgGK (dotted-line), IgG synthesis (solid-line) and the ratio IgGK/IgGL (dashed-line).*

An alternative measure of tumour kill has been suggested by clinicians, that is using the ratio of the tumourous and non-tumourous concentration as an alternative form of categorising patients. This has risen out of ancedotal evidence of patient recovery when the ratio has been taken. This ratio can also be seen in Figure 5.6 (dotted-line) and appears to offer a significant improvement in the estimation of tumour kill. The merit of using the ratio as a improved tumour marker over the measured period is studied in the following section.

5.4 Ratio of IgG Kappa/Lambda as an improved tumour marker

As has been described above, in MM patients, the excess tumourous monoclonal antibodies compete for free receptor sites in the endosome, resulting in a nonlinear relationship between IgG synthesis, concentration and catabolism. Due to this relationship the

total IgG cleared during chemotherapy will exhibit variation. The standard measure of chemotherapy effectiveness is to consider the drop in concentration of the tumourous IgG type in plasma. In this section a comparison is made using alternative plasma concentration measurements to describe the efficacy of the chemotherapy. The following tumour markers are compared:

- Total tumourous IgG type, inclusive of tumour and non-tumour production.
- Tumorous IgG alone.
- Ratio of total IgG type over non-tumourous IgG.
- Ratio of tumour IgG over non-tumour IgG.

5.4.1 Quasi steady-state model

In order to study the effectiveness of the ratio as a tumour marker, the competitive IgG FcRn quasi-steady state model is used. A schematic of the model is shown in Figure 5.2.

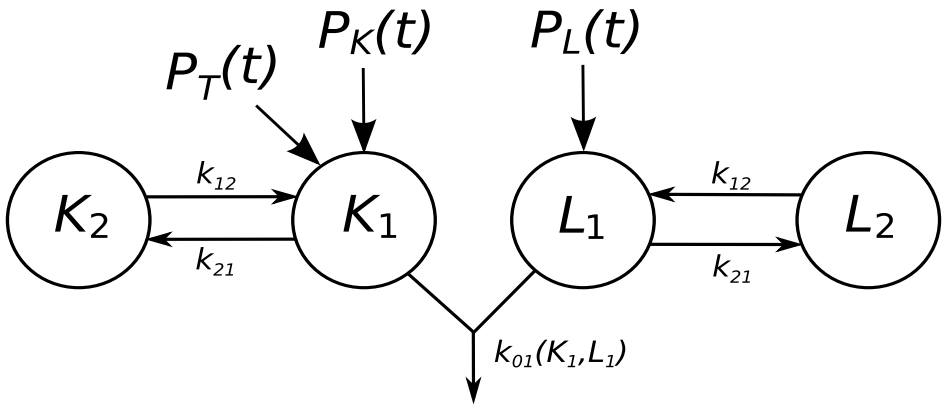


Figure 5.7: *QSSA compartmental model of IgG-K and IgG-L dynamics.*

The model is constructed from two coupled systems representing the IgGK and IgGL antibodies. All compartments labelled K represent the quantity of IgG Kappa, whilst L denotes the quantity of IgG Lambda in the equivalent pool. The model was initially constructed by considering the flow of IgG in plasma. It is known through

radio-nuclide marking, ([Waldman and Strober, 1969]), that IgG moves freely between plasma and EVF pools, across the plasma membrane. Therefore, two compartments were created (K_1,L_1 and K_2,L_2) with two linear rate constants (k_{12} and k_{21}) describing the flow between the plasma and EVF. However, for normal subjects, the quantity of IgG in plasma and EVF are approximately equal [Waldman and Strober, 1969], for this to be true, under steady state conditions k_{12}=k_{21}. The synthesis of intact IgG is controlled by three functions: $P_K(t)$ (IgGK), $P_L(t)$ (IgGL) and $P_T(t)$, representing the monoclonal IgG produced by the tumourous MM cells. In the above diagram, Figure 5.2, the tumour synthesis is shown only producing monoclonal IgGK. Tumours can produce either IgGK or IgGL.

Both IgGK and IgGL are continuously exchanged between plasma and the endosome through pinocytosis; IgG is either bound to FcRn receptors or remains free in the endosome. Bound IgG is protected and recycled into plasma, whilst the free IgG is catabolised. The system generated from the model (Figure 5.2) can be seen in equations (5.16a)-(5.16d):

$$\begin{aligned}
\dot{K_1}(t) &= -\left(k_{31} - \frac{V_m}{K_m + (K_1(t) + L_1(t))}\right) K_1(t) - k_{21}K_1(t) \\
&\quad + k_{12}K_2(t) + P_K(t) + P_T(t) \qquad (5.16a) \\
\dot{K_2}(t) &= k_{21}K_1(t) - k_{12}K_2(t) \qquad (5.16b) \\
\dot{L_1}(t) &= -\left(k_{31} - \frac{V_m}{K_m + (K_1(t) + L_1(t))}\right) L_1(t) - k_{21}L_1(t) \\
&\quad + k_{12}L_2(t) + P_L(t) \qquad (5.16c) \\
\dot{L_2}(t) &= k_{21}L_1(t) - k_{12}L_2(t) \qquad (5.16d)
\end{aligned}$$

with parameter definitions supplied in Table 5.3. On clinical presentation, prior to treatment, it is assumed that the patient's tumour has produced sufficient IgG to saturate all compartments and receptor sites. With the system therefore in steady-state the initial conditions are:

$$K_1(0) = C_{K1}(0)v_1 \quad K_2(0) = C_{K1}(0)v_2$$
$$L_1(0) = C_{L1}(0)v_1 \quad L_2(0) = C_{L1}(0)v_2$$

where $C_{K1}(0)$ and $C_{L1}(0)$ represent the concentration of the IgGK and IgGL in plasma prior to any chemotherapy treatment and v_1 and v_2 are the apparent volumes of dis-

tribution for plasma and EVF respectively. It is possible to observe four different measures relating to the IgG produced *in vivo*:

$$\begin{aligned} y_{K1}(t) &= C_{K1}(t) = c_1 K_1(t) \\ y_{T1}(t) &= C_{T1}(t) = c_1 T_1(t) \\ y_{L1}(t) &= C_{L1}(t) = c_1 L_1(t) \\ y_{KT1}(t) &= (C_{K1}(t) + C_{T1}(t)) = c_1(K_1(t) + T_1(t)) \end{aligned} \tag{5.17}$$

where c_1 represents an observation gain. It is assumed that all IgGK ($C_{K1}(t)$), IgGL ($C_{L1}(t)$) and monoclonal tumourous IgG ($C_{T1}(t)$) are captured by the measurement technique and as such this equates to the volume of distribution in plasma, $c_1 = \frac{1}{v_1}$ and the observation is simply the concentration in plasma. In equation (5.17) $y_{K1}(t)$ represents the concentration of IgGK in plasma; $y_{L1}(t)$ that of IgGL. The remaining two measurements relate to the measuring of tumour-specific concentrations. y_{T1} is the measurement of the monoclonal antibody produced by the tumour, whilst y_{KT1} is the total amount of IgG tumour type, in this case Kappa, but this may be Kappa or Lambda depending on the MM tumour type. If measurement of the monoclonal antibody is possible for a patient, two additional compartments would be added to the system model (eqns. (5.16a)-(5.16d)), representing the monoclonal antibody as follows

$$\begin{aligned} \dot{K_{T1}}(t) = &- \left(k_{31} - \frac{V_m}{K_m + (K_1(t) + L_1(t) + K_{T1}(t))} \right) K_{T1} - k_{21} K_{T1}(t) \\ &+ k_{12} K_{T2}(t) + P_T(t) \\ \dot{K_{T2}}(t) = &\ k_{21} K_{T1}(t) - k_{12} K_{T2}(t) \end{aligned} \tag{5.18}$$

where K_{Ti} represents the quantity of monoclonal antibody in compartment i and the remaining compartments are modified accordingly. Equations 5.16a and 5.16c would be modified as follows

$$\begin{aligned}\dot{K}_1(t) = &-\left(k_{31} - \frac{V_m}{K_m + (K_1(t) + L_1(t) + K_{T1}(t))}\right) K_1(t) - k_{21}K_1(t) \\ &+ k_{12}K_2(t) + P_K(t) \\ \dot{L}_1(t) = &-\left(k_{31} - \frac{V_m}{K_m + (K_1(t) + L_1(t) + K_{T1}(t))}\right) L_1(t) - L_{21}L_1(t) \\ &+ k_{12}L_2(t) + P_L(t)\end{aligned} \tag{5.19}$$

The above equations assume that the tumour is producing IgGK.

5.4.2 Chemotherapy markers

In order to compare the appropriateness of the different tumour markers it is necessary to define a mathematical framework by which markers can be compared. However, prior to this it is prudent to consider the medical definition of a tumour marker. A succinct description is offered by Krishna [Krishna, 2004, pp. 1163]

> "A tumour marker is a substance present in or produced by a tumour itself or produced by the host in response to a tumour, that can be used to distinguish a tumour from normal tissue, or can be used to detect tumours in large screening surveys."

To compare tumour markers, e.g. two measurable substances, it is necessary to define a metric between them to enable quantitive comparison. A 'better' marker is one that more closely predicts the underlying synthesis of the tumour during the patient's treatment. To view this concept from a different perspective the synthesis and markers can be viewed as discrete samples of a continuous function, yielding vectors in a real vector space $\mathbb{R}^N$, where N is the number of samples. It is then natural to consider the metric induced by the L^p norms as suitable candidates (e.g. Euclidean Norm, absolute value, etc) to produce a simple scalar comparison of two markers. For example, if two markers $m_1(t)$ and $m_2(t)$, are sampled at equivalent time points, $t_i, i = 1, ..., N$, and compared via the p-norm to a tumour synthesis $q(t)$; m_1 could be considered a better marker than m_2 if

$$(\sum_{i=1}^{N}(m_1[i] - q[i])^p)^{\frac{1}{p}} < (\sum_{i=1}^{N}(m_2[i] - q[i])^p)^{\frac{1}{p}} \tag{5.20}$$

where $[i]$ denotes discrete sampling of the relevant signal (i.e., $q[i] = q(t_i)$). However, if the use of the marker is considered in a clinical application, measurements will be periodically taken of the plasma concentrations (monthly or quarterly are not uncommon) with a sample interval that will be irregular and unspecified *a priori*. At each sample point the clinician will wish to make assessments of the chemotherapy's effectiveness. It is therefore imperative that for one tumour marker to be considered 'better' than another it must be closer to the monoclonal synthesis at any point in time during the patient's observation. This is not the case for the inequality in (5.20), and this relationship may hold true even if at a specific time point (t_i) marker m_1 is not a better representation of the tumour kill than m_2, i.e.

$$|m_1[i] - q[i]| > |m_2[i] - q[i]| \tag{5.21}$$

Therefore, a marker (m_1) can only be regarded better than another (m_2) if

$$|m_1(t) - q(t)| \;\leq\; |m_2(t) - q(t)| \qquad\qquad t \in (0, t_o] \tag{5.22}$$

where t_o is the maximum observation time. A perfect marker would be one that is exactly equal to tumour synthesis during this period.

5.4.3 Normalised ratio marker

Considering this information, a new marker ($R(t)$) has been suggested of the ratio of tumourous over non-tumourous IgG. For example, if the patient has an IgGK producing tumour the ratio is

$$R(t) := \frac{y_{KT1}(t)}{y_{L1}(t)}. \tag{5.23}$$

To relate this to the percentage tumour kill $R(t)$ is normalised with respect to the initial conditions to produce a precentage change in the ratio, e.g.

$$R(t) := \frac{y_{KT1}(t)}{y_{L1}(t)} \frac{y_{L1}(0)}{y_{KT1}(0)}. \tag{5.24}$$

Alternatively, if the clinician has the ability to measure the monoclonal IgGK ($y_{T1}(t)$) produced by the tumour in isolation from normal IgGK,

$$R(t) := \frac{y_{T1}(t)}{y_{L1}(t)}, \tag{5.25}$$

which is normalised to

$$R(t) := \frac{y_{T1}(t)}{y_{L1}(t)} \frac{y_{L1}(0)}{y_{T1}(0)}. \tag{5.26}$$

5.4.4 Analysis of ratio IgG marker

As described by equation (5.22) a better marker is one that is closest to the underlying tumour synthesis for all time. Unfortunately, the IgG-FcRn model proved intractable for this problem, in that an algebraic solution to inequality in equation (5.22) could not be found. Therefore an alternative problem is investigated: that is whether the ratio related markers are better indicators of tumour kill under asymptotic steady-state conditions. Given the definition of the normalised ratio above ($R(t)$), normalised tumour producing synthesis ($P_T(t)$) and the normalised IgG concentration in plasma, Kappa ($K(t)$) and Lambda ($L(t)$), the question reduces to: is the ratio a closer estimate of the overall tumour kill on completion of chemotherapy treatment? Numerical analysis, using typical patient parameters, is conducted (Section 5.4.4) to investigate the applicability of the markers during the entire treatment under typical clinical settings and patient conditions.

Steady-state analysis

If it is assumed that the patient secreting monoclonal IgGK has completed chemotherapy treatment and the concentrations of both the tumourous and non-tumourous IgG have reached steady-state, equations (5.16a)-(5.16d) all equal zero. Thus, for an IgGK secreting tumour, in steady-state we have

$$
\begin{aligned}
0 &= -FCR(\tilde{K_1}, \tilde{L_1})\tilde{K_1} - k_{21}\tilde{K_1} + k_{12}\tilde{K_2} + P_K + \alpha P_T(0) && (5.27a)\\
0 &= k_{21}\tilde{K_2} - k_{12}\tilde{K_1} && (5.27b)\\
0 &= -FCR(\tilde{K_1}, \tilde{L_1}))\tilde{L_1} - k_{21}\tilde{L_1} + k_{12}\tilde{L_2} + P_L && (5.27c)\\
0 &= k_{21}\tilde{L_1} - k_{12}\tilde{L_2} && (5.27d)
\end{aligned}
$$

where $\tilde{K_i}$ and $\tilde{L_i}$ are the steady-state quantities of Kappa and Lambda IgG respectively, α is the fraction to which the initial tumour production ($P_T(0)$) has fallen, i.e. 1 - α is the amount of synthesis reduced by chemotherapy, independent of the rate of kill. P_K are P_L is the constant production rates for normal IgGK and IgGL. In addition, $FCR(\cdot, \cdot)$ is the fractional catabolic rate of IgG from plasma defined by

$$FCR(K_1, L_1) = k_{31} - \frac{V_m}{K_m + (K_1 + L_1)} \tag{5.28}$$

with all other parameters as defined previously.

Proposition 5.4.1. *Assuming the above model is a valid representation of the patient response to treatment, irrespective of the dynamic tumour phase (growth or decay), under steady-state conditions prior to and after the treatment period, the ratio of tumourous over non-tumourous concentration will equal the fraction of tumour killed if either of the following conditions are met:*

(i) the initial tumour production is infinitely high, or

(ii) the normal production rate for the tumourous IgG type (Kappa or Lambda) is zero.

Proof. For ease of nomenclature an IgGK producing tumour will be assumed; however, the analysis is equivalent for either type of IgG.

Using the approximation $k_{12} = k_{21}$ the EVF compartments, Equations (5.27b) and (5.27d), yield $\tilde{K_1} = \tilde{K_2}$ and $\tilde{L_1} = \tilde{L_2}$ and Equations (5.27a) and (5.27c) simplify to give

$$
\begin{aligned}
-FCR(\tilde{K_1}, \tilde{L_1})\tilde{K_1} + P_K + \alpha P_T(0) &= 0, && (5.29a)\\
-FCR(\tilde{K_1}, \tilde{L_1}))\tilde{L_1} + P_L &= 0. && (5.29b)
\end{aligned}
$$

The ratio of the tumourous over non-tumourous IgG is then given by

$$\frac{\tilde{K_1}}{\tilde{L_1}} = \frac{P_K + \alpha P_T(0)}{P_L}. \tag{5.30}$$

To consider the percentage fall the ratio should be analysed with reference to the initial conditions

$$K_1(0) = \frac{P_K + P_T(0)}{FCR(K_1(0), L_1(0))} \tag{5.31}$$

and

$$L_1(0) = \frac{P_L}{FCR(K_1(0), L_1(0))}. \tag{5.32}$$

The ratio, normalised to the initial conditions, therefore becomes

$$\frac{\tilde{K_1}}{\tilde{L_1}} \frac{L_1(0)}{K_1(0)} = \frac{P_K + \alpha P_T(0)}{P_L} \frac{P_L}{P_K + P_T(0)} = \frac{P_K + \alpha P_T(0)}{P_K + P_T(0)}. \tag{5.33}$$

with the normalised ratio of the IgGK over IgGL purely dependent upon the terms relating to IgGK synthesis and the percentage of tumour killed (P_K, $P_T(0)$ and α). This suggests that if the patient is in steady-state post-chemotherapy treatment, which is often seen in patients, the normalised ratio is dependent solely upon the initial tumour production ($P_T(0)$), the underlying normal IgG production (P_K) and the fractional tumour kill (α). In summary, the smaller the normal production or larger the initial tumour, the closer the ratio will be to the tumour kill. Considering the size of the tumour

$$\lim_{P_T(0) \to \infty} \frac{P_K + \alpha P_T(0)}{P_K + P_T(0)} = \alpha \tag{5.34}$$

and the normal production of IgG, in this instance a patient with an IgGK tumour,

$$\lim_{P_K \to 0} \frac{P_K + \alpha P_T(0)}{P_K + P_T(0)} = \alpha. \tag{5.35}$$

□

This may at first seem an impractical statement as with real patients the above limiting conditions will cannot exist ($\lim_{P_T(0) \to \infty}$ and $\lim_{P_K \to 0}$); however, in MM patients tumourous production will be considerably increased, and the greater the synthesis the

worse the patient condition and more critical appraisal of the fraction of tumour killed. Also, it is believed that in a tumourous environment normal production is often suppressed in patients. Therefore studying the limits offers insight into the overall trend or behaviour.

Corollary 1. *Given the model of the system defined in Equations 5.16a-5.16d, if it is possible to measure only the tumourous IgG produced, i.e. excluding normal IgG of the same type (Kappa or Lambda), the ratio of tumourous over non-tumourous IgG concentration, in steady-state conditions, will give the exact fractional tumour kill.*

Proof. In this case, the extended model described in Equation (5.18) is used, as such Equation 5.33 in the previous proof becomes

$$\frac{\tilde{K_1}}{\tilde{L_1}}\frac{L_1(0)}{K_1(0)} = \frac{\alpha P_T(0)}{P_L}\frac{P_L}{P_T(0)} = \alpha. \tag{5.36}$$

□

As can be seen from Equation (5.36), the result under the assumptions of Corollary 1 is completely independent of the underlying production of normal IgG, of the same type as the tumour, in this case IgGK.

Proposition 5.4.2. *If the patient, as described by the QSS model (Equations 5.16a-5.16d), is in a steady-state condition, the ratio of the total IgG tumour over the non-tumourous variant will be an improved marker on measuring tumour IgG concentration alone if there is a change in the total IgG present from the pre-treatment values.*

Proof. Considering a patient that has an IgGK producing tumour which is decaying under the influence of chemotherapy, the total IgG concentration is expected to decrease, therefore

$$K_1(0) + L_1(0) > \tilde{K_1} + \tilde{L_1}. \tag{5.37}$$

Also, the definition of the Fractional Catabolic Rate (Equation (6.6)) implies that

$$k_{31} - \frac{V_m}{K_m + (\tilde{K_1} + \tilde{L_1})} < k_{31} - \frac{V_m}{K_m + (K_1(0) + L_1(0))}, \tag{5.38}$$

or

$$FCR(\tilde{K_1}, \tilde{L_1}) < FCR(K_1(0), L_1(0)), \tag{5.39}$$

and, as FCR in practice is always positive, ultimately results in

$$1 < \frac{FCR(K_1(0), L_1(0))}{FCR(\tilde{K_1}, \tilde{L_1})}. \tag{5.40}$$

The normalised steady-state concentration, using Equations (5.29a) and (5.31), can be written as

$$\frac{\tilde{K_1}}{K_1(0)} = \frac{P_K + \alpha P_T(0)}{FCR(\tilde{K_1}, \tilde{L_1})} \frac{FCR(K_1(0), L_1(0))}{P_K + P_T(0)}. \tag{5.41}$$

Therefore, comparing the ratio of tumour over non-tumourous concentration (Equation (5.33)) against the normalised plasma concentration (Equation (5.41)), the inequality 5.40 implies that

$$\frac{R(t)}{R(0)} = \frac{P_K + \alpha P_T(0)}{P_K + P_T(0)} < \frac{P_K + \alpha P_T(0)}{FCR(\tilde{K_1}, \tilde{L_1})} \frac{FCR(K_1(0), L_1(0))}{P_K + P_T(0)} = \frac{\tilde{K_1}}{K_1(0)}. \tag{5.42}$$

□

Therefore, for the ratio to be a better marker there must be a fall in total IgG plasma concentration, which in all patient data provided, is true. As with previous results, the reverse holds for tumours exhibiting growth.

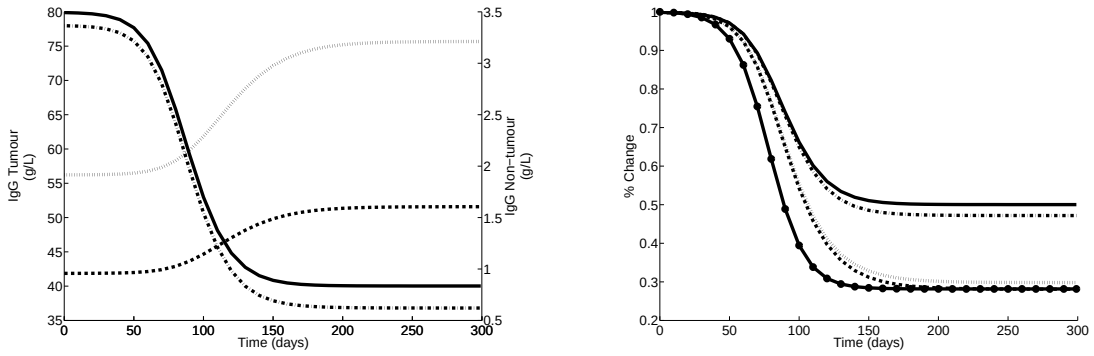
Numerical analysis

In order to exercise the model in a realistic environment, two patient scenarios have been considered:

- Patient A presents with a concentration of 80 g/L tumourous IgGK, with a 50% reduction seen post-chemotherapy.
- Patient B presents with a concentration of 30 g/L IgGK and a 100% reduction in synthesis is achieved.

Assuming that synthesis of normal IgGK and IgGL continues in line with normal concentrations, the synthesis parameters P_K and P_L, are set to achieve concentrations

of 12 g/L IgGK and 6 g/L IgGL. However, it should be noted that the concentrations of normal IgG will be suppressed under increased catabolism due to the raised total concentration of IgG present. If the normal level of production and the required initial level of tumourous IgG concentration are known, the remaining initial conditions can be found, see Appendix C.



(a) IgG in plasma

(b) Percentage Change in Markers

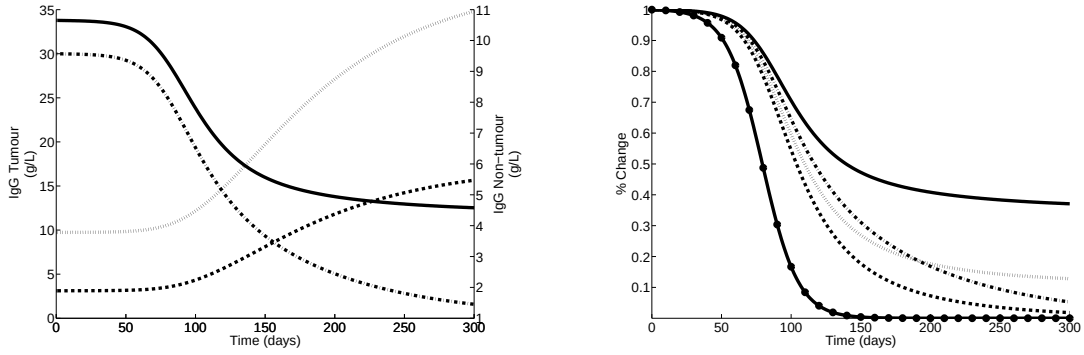
Figure 5.8: *Simulated results for patient A. Figure a) IgG in plasma (solid line total IgGK, chained-line IgGK tumour, dotted-line normal IgGK and dashed-line IgGL); figure b) percentage change in markers (bold-solid line tumour synthesis, solid-line total IgG K, chained-line IgG K tumour, dotted-line ratio total IgGK/IgGL, dashed-line ratio IgGK tumour/IgGL)*

For each of the above scenarios a simulated experiment has been conducted using the parameters derived in Table 5.3 (Figure 5.4), the results of which can be seen in Figures 5.8 and 5.9. Figure 5.8(a) shows the simulation results for Patient A for the concentrations of the IgG proteins present in plasma. In this figure the left hand axis is scaled for the tumour and total IgGK (tumourous + normal) whilst the right hand axis is formatted to display the normal IgGK and IgGL concentrations. The simulation clearly shows the fall in tumour IgG and the effect it has on the total tumourous IgG measured. In addition, the rise in normal IgG (both Kappa and Lambda) is evident with each rising above the initial concentration. However, they do not increase to the normal concentration ranges (12 g/L and 6 g/L) as noted previously.

Figure 5.8(b) compares the concentration change that will be seen if the four different marker measurements are considered. The solid line, marked with circles, represents the actual synthesis of the monoclonal IgGK produced by the tumourous cells; the ideal but immeasurable marker. Interestingly, if the actual tumour synthesis is compared to

the fall in plasma concentration a large difference is evident, with the plasma falling indicating a 50% reduction when the synthesis has actually drop by almost 70%.

During the initial phase (0-75 days) little difference is evident between the markers; however, as described by the analysis above, the ratio becomes a better marker than measuring the concentrations in plasma directly. As noted in Proposition 5.4.1, the ratio of total IgGK is a much improved marker due to the relatively high tumour production and low non-tumour production. The most accurate marker, as anticipated by Corollary 1, is the ratio of IgGK tumour and the non-tumour IgGL, with the final value at 300 days equalling the fractional kill (a 22 % difference seen between this measurement and that of total IgGK). The ratio of IgGK tumour over non-tumour is the most effective marker during the entire measurement process.



(a) IgG in plasma

(b) Percentage Change in Markers

Figure 5.9: *Simulated results for patient B. Figure a) IgG in plasma (solid line total IgGK, chained-line IgGK tumour, dotted-line normal IgGK and dashed-line IgGL); figure b) percentage change in markers (solid-line (∘) tumour synthesis, solid-line total IgG K, chained-line IgG K tumour, dotted-line ratio total IgGK/IgGL, dashed-line ratio IgGK tumour/IgGL).*

The simulation of the condition described by the 'Patient B' scenario can be seen in Figure 5.9, with the plots and axes scaled as for Patient A. With reference to Figure 5.9(a) the tumour concentration is seen to drop to zero, due to the 100% tumour kill; however, in this scenario the normal IgG levels return to those expected in a healthy patient. As shown in Figure 5.9(b) the ratio markers are once again better than the plasma concentration measurements, with one noticeable exception, that of measuring tumourous monoclonal IgG. As would be expected in the latter stages of treatment as the synthesis is zero the concentration will lag behind but will eventually

also reach zero. It is interesting to note, that even for this scenario it is still more effective to measure the ratio of monoclonal IgGK over IgGL, as the ratio approaches zero on a shorter time frame than the monoclonal tumour concentration. However, approximately 40% difference can be seen between measuring the ratio and total IgGK in plasma concentration.

5.5 Summary

In this chapter a model has been proposed describing the competitive binding of IgG-Kappa and IgG-Lambda to FcRn receptors. A reduced model was derived using basic assumptions from the underlying receptor kinetics. It was validated using data from normal subjects and patients with Hypoproteinemia, the results of which offered information on the maximum and normal clearance achieved through pinocytosis and binding to the FcRn. These parameters were then used to investigate the effects of FcRn recycling on a patient with IgG secreting Multiple Myeloma. It was shown that the model, when coupled with a delayed-logistic input function, could be used to successfully simulate the effects of chemotherapy in accordance with evidence provided by *in vitro* measurements.

Although certain biological complexities noted above have been excluded from the model presented, evidence has been provided (see Figure 5.6) that suggests the ratio of tumour over non-tumour plasma measurement may offer an improved tumour kill indicator. The ratio appears to give a significant improvement over the current best practice of measuring the total change in IgG tumour in plasma alone. This may offer an alternative method of categorising the chemotherapy response of patients with intact IgG secreting Myeloma giving improved accuracy and earlier detection of poor tumour kill or ineffective chemotherapy treatment.

In the next chapter investigation of the whole-antibody analysis is continued but rather than the tumourous synthesis seen in MM patients discussed in this chapter, the full immune responses to transplanted organs will be considered. Patients with Intact Secreting MM produce only monoclonal IgG all of which behave with similar dynamics *in vivo*; transplant patients, however, produce a range of antibodies and antibody types and as such additional models are derived from the IgG models presented here.

Chapter 6

Immune response to transplants

As described in Chapter 2, patients who receive transplants will incur an immune response to any transplanted device or implanted material. To avoid damage to the implanted substance clinicians use apheresis to clear antibodies from the blood stream; however, due to the size of whole antibodies haemodialysis is not an option. Patients also receive immunosuppressant drugs to prevent the immune system generating new antibodies [Higgins, 2007], rather than chemotherapy drugs given to Myeloma patients, as seen previously (Chapters 4 and 5). Unlike MM patients, the immune response generated by transplant patients is natural and as such a range of antibody isotypes are produced.

In this chapter, models are defined that attempt to estimate the synthesis and clearance of immunoglobulin (IgG, IgA and IgM) of patients undergoing kidney transplants. The models constructed previously worked solely upon either single antibody or antibody fragments; however, due to the similarity in the catabolism of the antibody types, with minor modification, these models are combined to analyse the total immune response. At first glance this may not appear comparable to the previous work; however, although transplant patients do not have cancerous tumours, they do have highly-elevated immunoglobulin concentration in plasma. In addition, although the treatments given to transplant patients are different from those of cancer patients, they are analogous in their effect on the body from a systems perspective. Multiple Myeloma patients receive chemotherapy which suppresses the production of immune complexes, whereas transplant patients are given immune suppressing drugs that re-

tard the production of antibodies to antigens; both have the effective of reducing the underlying synthesis. As was seen in Chapter 4, MM patients are treated with high-flux haemodialysis to clear excess FLC but due to the size of the molecules in question (immunoglobulin A, G and M) this is not possible for transplant patients. In comparison transplant patients are treated with plasmapheresis, hemoperfusion or plasma exchange. These analogies allow the previous models developed to be modified slightly and the analysis techniques employed to address the dynamics.

The aim of the work presented in this chapter is to enable clinicians to better categorise a patients response to the implanted organ, allowing them to modify procedures earlier on in the patient's treatment than was previously possible. In the first part of this chapter a simple technique to estimate the synthesis rates and efficiency of apheresis for different antibody types is presented and from this, understanding of the antibody synthesis can be gained regarding the effects of immune suppression drugs and apheresis, pre- and post-transplant. In the final section of this chapter an alternative method for automatic categorisation of patients is investigated. A new measurement technique is available to measure antibodies directly targeted against the kidney [Higgins, 2007]. However, due to anomalies in the measurements the process-driven modelling techniques used previously are not applicable. Therefore, time series cluster analysis techniques are used to separate the patients into groups in an attempt to assist clinicians in diagnosing a particular patients response in a nonparametric manner.

6.1 Describing the immune reaction to transplanted organs

After transplantation a patient's immune system will generate antibodies against the donated organ. It will be seen as an invading body and therefore marked for removal. Even in the best possible match of donor organ and patient this is the case. It is therefore of great interest to clinicians to have an accurate estimate for the generation of antibodies. As seen for Multiple Myeloma (Chapter 5) the current best practice is to measure the concentration of antibodies in plasma. To determine the response

the concentrations are compared to previous history to determine the current state of the patient. However, this expert analysis is further complicated by the different numbers of antibodies produced. For each patient it is common practice to measure three antibodies: IgG, IgA and IgM (please refer to Chapter 2 for a description of these). Each will be prevalent at different stages of the immune response. In addition, through new measurement techniques [Higgins, 2007] antibodies specifically generated to target the donor organ can be measured, referred to as donor specific antibodies (DSA). Due to inconsistency in the measurement techniques the DSA measurements cannot be used in process driven modelling; however, an alternative method to use these measurements is described in Section 6.2.

For each patient the following data were available (an example dataset can be seen in Figure 6.1):

- Measurements of IgM, IgG and IgA concentrations; unlike the previous haemodialysis study (Chapter 4) only two measurements were taken, one at the start and one at the end of apheresis. In addition, these data were not always complete (see Section 6.1.1) with either pre- or post-apheresis measurements missing.
- Indication of the apheresis treatment given at each session (plasmapheresis (PP), plasma exchange (PE) or plasma absorption (PA)). For details on the different treatments see Chapter 2.
- The total body weight prior to the apheresis treatment.
- Date of each of the apheresis sessions and the transplant operation.

The IgG models previously developed in Chapter 5 can be applied to the transplanted patient data allowing physicians to estimate the antibody synthesis and its reaction to immunosuppressive drugs and to investigate the effectiveness of the apheresis treatments in removing whole-antibody. From the previous work presented, the most practical approach would be to assume a known parametric form for the synthesis (e.g. Logistic growth, as seen in Chapter 5) and initial estimates for the system and fit the data using a reliable method. The results from this approach were extremely poor yielding numerically undetermined parameters and poor fits to data. The reasons for

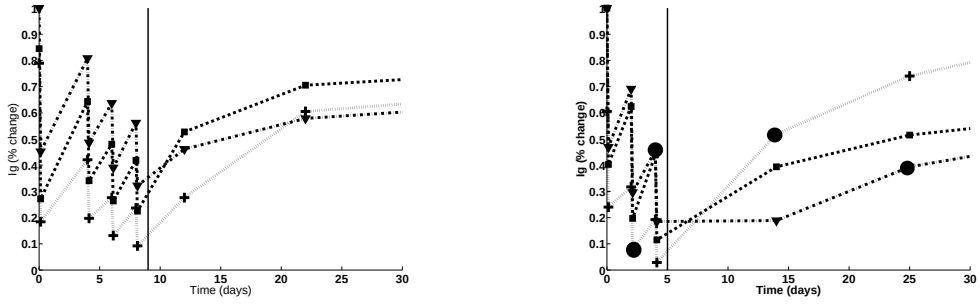
this are believed to be the sparsity of data and the variability of patient response to immunotherapy drugs and apheresis treatment. For example, a single patient could receive twelve apheresis treatments prior to their operation, each requiring a separate clearance parameter, with only a maximum of twenty four data points.

In order to provide estimates for the required parameters, assumptions are made regarding the system dynamics over a known time interval, allowing the problem to be discretised. It is assumed that the synthesis is piecewise constant over the treatment period. Further, all other parameters, with the exception of apheresis clearance, antibody production and EVF concentrations, are assumed to be known for each patient.

6.1.1 Patient data

Figure 6.1 shows two examples of the data provided[1]; in each case there are three types of measurement taken:

- Pre-apheresis, taken immediately before apheresis (pre).
- Post-apheresis, taken when apheresis has finished (post).
- No apheresis, measured when the patient is under clinical observation but not on apheresis (no ap).



(a) Patient with all observations present

(b) Missing data points - large dark circle are missing measurements

Figure 6.1: *Normalised example data sets. IgG ▽, IgA □ and IgM +. Vertical line represents the time of transplant.*

[1]All data have been normalised to the maximum value, to enable comparison.

A complete dataset is shown in Figure 6.1(a), with no missing values; it should be noted that even in this case the data are not ideal, in that they are sparse and irregularly sampled. The alternative case (a poor dataset) can be seen in figure 6.1(b); the figure has been annotated with large dots to highlight the missing data points. When using *in vivo* biomedical data, missing data points are a common problem. Two basic solutions to the problem exist [Altman and Bland, 2007]:

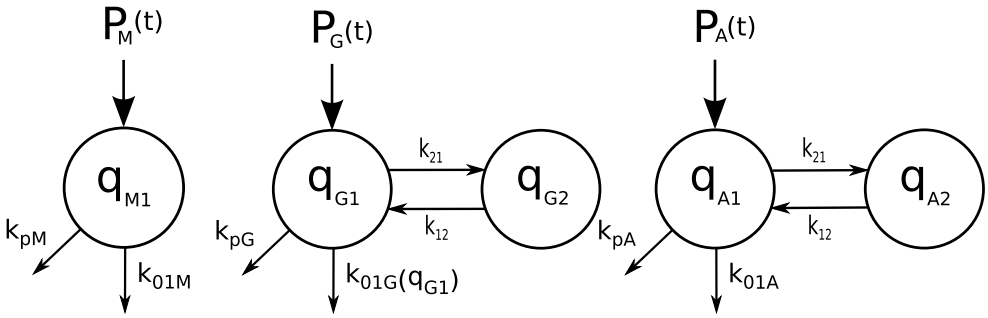
- to ignore all data sets that contain missing data.
- impute the missing values using statistical techniques.

Ignoring datasets that contain missing observations can be applied during the model development and validation phase; however, to be used in a clinical setting the model should be applicable to all patients and suggest a typical response even with the limited information available. When the observations provided are frequently sampled continuous time series data, the missing measurements may be simply imputed by interpolation techniques; for the discontinuous system of apheresis this is not possible. In the following work this problem is addressed by assuming that the parameter estimates that are not calculable due to the missing data points (e.g. apheresis clearance and Ig synthesis) remain the same as the previous estimated value, or are constant until the next observation. This approach is only applicable as the missing observations are relatively uncommon, with contiguous missing points never occurring.

6.1.2 Method

To investigate the kinetics three models were constructed, one for each of the Ig types of interest: IgA, IgG and IgM. A compartmental schematic of each can be seen in Figure 6.2. In the following subsections each model will be described in more detail. As the models produced are all linear, closed form solutions are available for the unknown clearance, synthesis and EVF concentrations given the available plasma concentration measurements.

In Figure 6.2, label $q_{*1}(t)$ indicates the quantity of the antibody present in plasma; whilst, $q_{*2}(t)$ is the quantity in EVF. $P_*(t)$ is the synthesis rate of the immunoglobulin, due to the immune suppression and subsequent immune response to the transplanted



(a) IgM (b) IgG (c) IgA

Figure 6.2: *The three Ig models. The variables q_{Mi}, q_{Gi} and q_{Ai} denote the quantity of IgM, IgG and IgA in compartment i respectively; compartment 1 is plasma, 2 is EVF. k_{pM}, k_{pG} and k_{pA} are the effective apheresis clearance. k_{01*} the natural clearance of the antibody.*

kidney; this is known to be non-constant over the period of observation. All other rate constants are discussed in the relevant model section.

IgG model

As has been previously discussed (Chapter 5) IgG is known to exist in both plasma and EVF, it has been shown via radio marking to be equally distributed between both compartments in terms of quantities [Waldman and Strober, 1969]. The model, as seen in figure 6.2(b), has plasma (q_{G1}) and EVF (q_{G2}) compartments with linear exchange between the two (k_{12} and k_{21}). Clearance (k_{01} and k_{pG}) and synthesis ($P_G(t)$) are via the plasma pool only. Unlike the FcRn model, described in Chapter 5, only the total IgG is measured therefore a simplified version of the FcRn model can be used. This model is equivalent to that presented by Kim et al. [2007] and Ferl et al. [2005]. The model for IgG dynamics can be described by the following equations:

$$\dot{q_{G1}}(t) = -(k_{01G}(q_{G1}(t)) + k_{pG} + k_{21})q_{G1}(t) + k_{12}q_{G2}(t) + P_G(t) \quad (6.1)$$

$$\dot{q_{G2}}(t) = -k_{12}q_{G2}(t) + k_{21}q_{G1}(t) \quad (6.2)$$

$$q_{G1}(0) = q_{G10} \quad (6.3)$$

$$q_{G2}(0) = q_{G20} \quad (6.4)$$

As described above only a single measurement of the plasma concentration is taken before and after apheresis has been performed. Therefore, the output structure of the

model is simply

$$y_{G1}(\tau) = \frac{q_{G1}(\tau)}{v_1(\tau)} \qquad \tau \in \{t_{s1}, t_{e1}, ..., t_{sN}, t_{eN}\} \tag{6.5}$$

where $v_1(\tau)$ is the plasma volume at time τ, this may not be constant as has been assumed previously; t_{si} and t_{ei} are the time points at the start and end of plasmapheresis, and N is the number of plasmapheresis treatments.

FcRn clearance (k_{01})

It is known that clearance of the IgG protein is mediated through the FcRn receptor [Brambell et al., 1964]. The rate of clearance is dependent upon the quantity of IgG in plasma due to the binding and subsequent recycling via the epithelium, often referred to as the fractional clearance rate (FCR). Unlike the case for Multiple Myeloma intact IgG modelling, seen previously in Chapter 5, there is no distinction between the IgG isotypes produced (Lambda or Kappa); as such the competitive binding to the FcRn receptors modelled previously need not be considered and the clearance reduces to a standard form [Ferl et al., 2005; Kim et al., 2007; Waldman and Strober, 1969]. If pseudo-steady state assumptions are made regarding the receptor dynamics the fractional clearance rate (k_{01G}) can be estimated by

$$k_{01G}(q_{G1}(t)) = \left(\alpha - \frac{V_m}{K_m + q_{G1}(t)}\right) \tag{6.6}$$

where α is the maximum rate of clearance and V_m the maximum recycling rate of IgG and K_m the concentration at which half-maximal recycling occurs, as seen in Chapter 5.

IgA model

IgA is similar in size to IgG (IgA $\approx$ 160 kDa, IgG $\approx$ 146 kDa) and has very similar properties in terms of *in vivo* kinetics and distribution between plasma and EVF [Mehlhorn and Armstrong, 2001, pp. 43 - 44]). IgA does not bind to FcRn receptors and so is not recycled into plasma. Therefore the FCR of IgA is constant and in equation

(6.1) reduces to $k_{01A}(t) = \alpha$ from equation (6.6). As such the system equations can be considered linear and can be described by the following system:

$$\dot{q_{A1}}(t) = -(k_{01A} + k_{pA} + k_{21})q_{A1}(t) + k_{12}q_{A2}(t) + P_A(t) \quad (6.7)$$

$$\dot{q_{A2}}(t) = -k_{12}q_{A2}(t) + k_{21}q_{A1}(t) \quad (6.8)$$

$$q_{A1}(0) = q_{A10} \quad (6.9)$$

$$q_{A2}(0) = q_{A20} \quad (6.10)$$

Concentrations are sampled at pre- and post-apheresis treatment, therefore the output structure is identical to equation 6.5, measuring IgA rather than IgG.

IgM model

IgM is a larger molecule than either IgA or IgG ($\approx$ 900 kDa); it does not flow between plasma and EVF in significant quantities [Mehlhorn and Armstrong, 2001, pp. 43 - 44]). The model constructed therefore consists of a single pool describing the quantity of IgM in plasma:

$$\dot{q_M}(t) = -(k_{01M} + k_{pM})q_M(t) + P_M(t) \quad (6.11)$$

$$q_M(0) = q_{M0} \quad (6.12)$$

$P_M(t)$ is the synthesis of new IgM molecules. k_{01M} can be estimated from the known half-life of IgM ($\approx$ 5 days). $P_M(t)$ and k_{pM} are the parameters to be investigated during this study.

Volume of distribution

The volume of distribution for IgM is the plasma pool, whilst IgG and IgA are distributed between two pools: plasma and EVF. These are calculated from the total body weight given in the clinical data by the approximations given in Malesker and Morrow [2007], as shown in Chapter 4.

6.1.3 Analysis

As has been previously seen, if it is assumed that the patient's immune system is in steady-state prior to treatment, estimates for starting synthesis can be calculated, and if applicable, initial concentration in EVF. Using these initial values, it is then possible to iterate through all given data points generating values for predicted discretised values of $P[n]$, $kp[n]$ and $q_2[n]$, where $[n]$ is the value at the n^{th} data point. For each off-apheresis session the synthesis is calculated, whilst for an on-apheresis data set the clearance is estimated assuming the synthesis is maintained at the pre-apheresis value. This is feasible given the relative timescales of the measurements and the response time of the immune suppression treatment. The pseudo-code for this algorithm is presented below in Algorithm 1. A similar method is suggested by Iga et al. [1986] to determine drug absorption rates for continuous systems, without the added complexity of the switched dynamics introduced by the apheresis treatment.

Algorithm 1: Pseudo-code to estimate clearance and synthesis.

```
Input: all valid data-points
Output: all estimated synthesis and clearance
foreach observation do
    curr ← current observation;
    next ← next observation;
    if (curr of type 'post' or 'no ap') and (next of type 'pre' or 'no ap') then
        Estimate synthesis and EVF concentration at end period;
    else if (curr of type 'pre') and (next of type 'post') then
        Estimate clearance, using previous synthesis estimate, and EVF
        concentration at end period;
    else
        Missing data point. Synthesis and clearance assumed unchanged;
    end
end
```

Equations relating the immune synthesis ($P_*[n]$) and EVF concentration $q_{2i}[n]$ can be found analytically. Unfortunately, the closed form solutions of eqns. (6.1), (6.10) and (6.12) do not allow an algebraic relationship for apheresis clearance ($k_{p*}[n]$) to be calculated. Therefore, a numerical solution for the value of k_{p*} can be found using an

appropriate one-dimensional root-finding algorithm [Press et al., 1992, chap. 9].

As an example, consider the IgM model (eq. (6.12)). If the patient is not on apheresis and the synthesis of IgM is assumed to be constant over the obersvation period, an analytical solution can be found for the quantity of IgM in the patient, given by

$$q_M(t) = \; e^{-k_{01M}t}q_{M0} + \tfrac{P_M}{k_{01M}}(1 - e^{-k_{01M}t}). \tag{6.13}$$

If this equation is then discretised, with the production assumed to be constant over the sample period, equation (6.13) can be re-arranged to give an estimate for the IgM production, given consecutive measurements of IgM ($q_M[n]$ and $q_M[n+1]$)

$$P_M[n] \; = \tfrac{k_{01M}(q_M[n+1]-e^{-k_{01M}\tau}q_M[n])}{1-e^{-k_{01M}\tau}} \quad n = 1, 2, ... \tag{6.14}$$

where n is the sample index, τ time between the measurements and $q_M[i]$ is the measurement at i. If the IgM concentration is assumed to be in steady-state prior to treatment the initial value of the IgM synthesis is given by

$$P_M[0] = \; k_{01}q_M[0] \tag{6.15}$$

The IgM apheresis clearance estimates ($k_{pM}[n]$) can then be found by minimising the difference between the observation and the predicted estimate at time point t_n, e.g.

$$k_{pM}[n] \; = \arg\min_{k_{pM}} d(q_M(t_n; k_{pM}), c_M[n]\; v_p[n]) \quad n = 1, 2, ... \tag{6.16}$$

where $c_M[n]$ is the observation of the patient concentration at time point t_n, $q_M[n]$ the estimated quantity from equation (6.12), and $v_p[n]$ the pre-treatment plasma volume. The metric to be minimised ($d(\cdot)$ in equation (6.16)) can then be chosen to suit the optimisation algorithm, i.e. for the MATLAB *fzero* function the difference between the estimate and measured value is sufficient.

A similar, if slightly more complex, process can be followed for IgG and IgA where the analytical solution at time point t_n is given by

$$q[n] \;=\; e^{A\tau}q[n-1] + \int_0^{\tau} P[n]e^{A(\tau-s)}ds \tag{6.17}$$

where $q[n]$ is a vector of solutions for plasma (q_1) and EVF (q_2); $e^{A\tau}$ is the matrix exponential, A the linear system matrix. As $q_1[n]$ is the given measurement, these equations can be solved simultaneously for $P[n]$ and $q_2[n]$ (see Appendix D). The apheresis clearance (k_p) is again calculated through a numerical minimisation procedure, as in equation (6.16). For the IgG model (equation 6.1) in order to produce a solution the estimate for natural clearance (k_{01G}) is linearised around the observed value ($q_1[n-1]$), and assumed to remain constant through the observation period; giving a linear system to which the solution in equation 6.17 is applicable.

6.1.4 Variation estimates

For the values used to predict the synthesis and clearance two types of uncertainty exist:

- Measurement error in the data provided (e.g. plasma concentration).
- Inter- and intra-patient rate constant variation (e.g. FCR, intercompartmental transfers, natural clearance).

The measurement error for antibody measurements is provided by the clinical staff and an estimate of ±5% has been given by haematological staff through unpublished dilution studies. Patient variability is more difficult to quantify. Each patient will have different parameter estimates than those of the parameter values used; there may also be variation between patients and in the same patient on different days. However, from previous work carried out with regards to FLC clearance (Chapter 4) a coefficient of variation (CoV) of ±10% would appear to give a reasonable estimate.

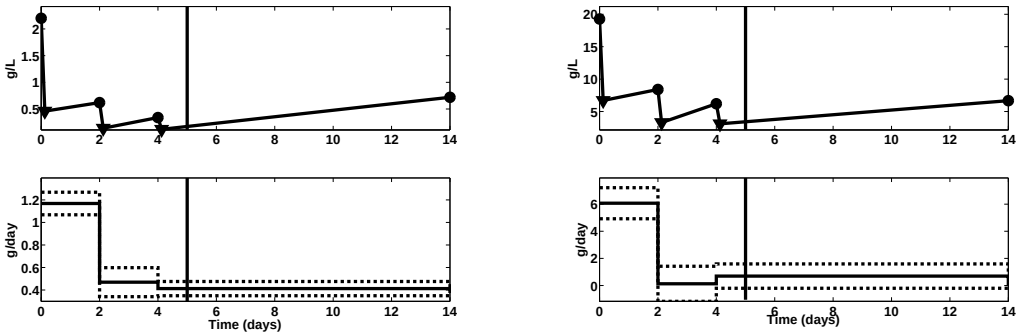
To estimate the deviation in the values a simple iterative Monte Carlo scheme has been implemented. For each value calculated one thousand iterations were performed to re-calculate the estimated values. At each iteration, values for the model parameters and data measurement were selected from a normally distributed parameter set, with mean value (μ_i) equal to the initial estimate and a standard deviation given by $\mu_i \times CoV_i$.

6.1.5 Results

The above algorithm was run for each set of patient data. This results in an estimate for the synthesis of the antibody production over the time course of the treatment. Example sets of results are shown in Figures 6.3 and Figures 6.4. Referring to Figure 6.3, for each antibody type two figures are shown, the lower figure is the estimated production using the above technique; the upper figure shows the simulated output given the estimated input. The circles in the top figure represent pre-apheresis measurements, whilst the triangles denote the post-apheresis observations. The estimated input is shown as a piecewise constant graph to encourage the recognition that a constant production is assumed for the discretisation and the estimate should be considered a mean estimate for the antibody synthesis. For the simulated output the solid line in the top figure represents the concentration of the antibody in the plasma compartment. The vertical line shown in each figure shows the time of organ transplant.

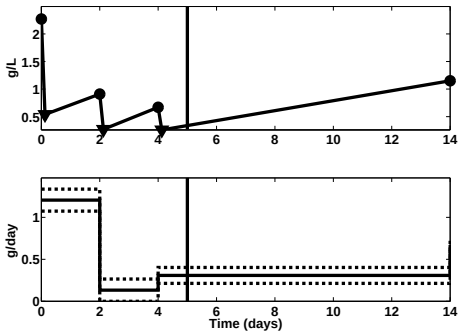
Patient 1, as seen in Figure 6.3, shows a classic, and successful, response to the combination of immune suppression and plasma exchange treatment. Pre-transplant, all immune complexes (IgG, IgA and IgM) decrease within two or three days of treatment. During this time the patient is given a series of plasma exchange treatments to reduce the basal level of antibodies to as low as possible. In some patients an immune response can be generated against the apheresis treatment which would be recognised by an increase in the antibody synthesis; this can often be masked by reductions in concentration in plasma due to the ongoing apheresis treatments. Post-transplant the patient is still on the same amount of immuno-suppressing drug, but little evidence is seen for this from the plots.

An alternative patient response can be seen in Figure 6.4. The patient received extended apheresis treatment sessions prior to surgery, ten sessions in total, indicating that the clinicians did not believe the change in plasma concentration to be great enough for a successful transplant. However, from viewing the underlying synthesis it can be seen that this is probably due to reaching basal levels of concentration due to production. Overall very little change can be seen in production of each of the antibody types, approximately 30% at best, as can be seen in Figure 6.3; in a patient with a successful response this can be considerably higher, often 80% or greater. For IgM a



(a) IgM

(b) IgG

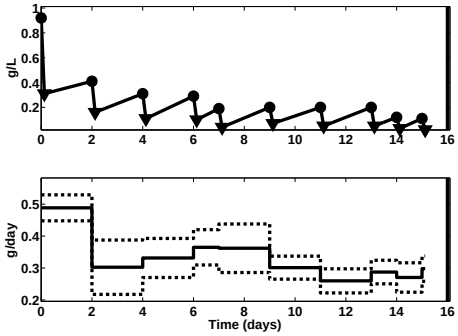


(c) IgA

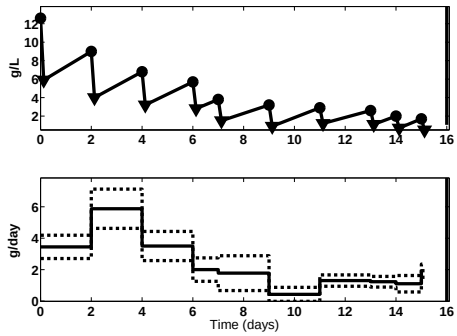
Figure 6.3: *Estimates for synthesis of antibodies from patient 1. Top - simulated plasma concentration and measurements (circles pre-apheresis, triangles post-apheresis. Bottom - solid-line predicted synthesis, dashed-line ± one standard deviation from estimated value.)*

typical response to the immuno-suppressing drugs is evident, whilst IgG rises indicating an increased immune response, and IgA remains relatively constant throughout. Unfortunately, patient data were not available post-operatively.

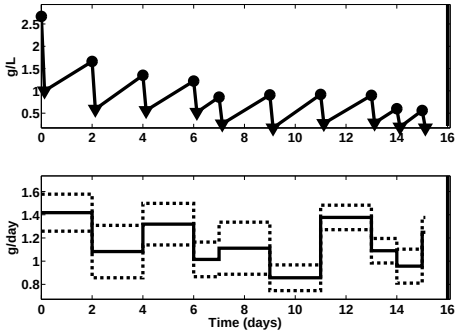
The discrete deconvolution as described presents difficulties that should be noted. Firstly, assuming the production is constant over the discretisation period, e.g. two days for patient two, may be unrealistic as the antibody synthesis is suppressed by drug combinations which are administered throughout the treatment. Whilst the piecewise constant values suggest a trend in the synthesis, it would be beneficial to consider the function as continuous over the treatment range. In addition, there is information regarding the functional form of the synthesis that is known *a priori* which could be used to determine the generation rate more accurately. The input signal cannot, by definition, be negative; whilst this did not occur for the patient data seen it is



(a) IgM



(b) IgG



(c) IgA

Figure 6.4: *Estimates for synthesis of antibodies from patient 2 (pre-transplant only). Top - simulated plasma concentration and measurements (circles pre-apheresis, triangles post-apheresis. Bottom - solid-line predicted synthesis, dashed-line ± one standard deviation from estimated value.*

not prevented by the discrete method. An extension of this problem can be seen in Figure 6.4(b), with the confidence interval generated by the Monte-Carlo method suggesting a negative production as the minimum bound between the second to fourth day. Further, as it is a natural signal, the synthesis will not contain discontinuities and will exhibit a degree of smoothness. This issue will be addressed in Chapter 7, where a more complex deconvolution method is developed.

In addition to offering prediction on a patient's response to immunosuppression the algorithm presented provides information on the different apheresis treatments (plasma exchange, plasma absorption and plasmapheresis). Table 6.1 shows the clearance values obtained for each of the antibody types (IgM, IgG and IgA) for the 13 patients observed during the studies. Each patient was treated with only one of the apheresis methods

Table 6.1: *Estimates for the clearance of IgM, IgG and IgA of different apheresis techniques (DFPP - Double Filtration Plasmapheresis, PE - Plasma Exchange, PA - Plasma Absorption). The estimate for each patient is averaged over all the treatments, as specified in the number column.*

Id	Type	Number	Clearance $mins^{-1}$ (std. dev.)		
			IgM	IgG	IgA
1	PE	11	6.94E-04 (3.71E-05)	6.44E-03 (3.98E-03)	2.94E-03 (1.19E-03)
2	PA	4	6.90E-04 (2.82E-05)	6.31E-04 (3.44E-04)	5.06E-04 (4.84E-04)
3	PA	3	6.75E-04 (7.37E-06)	1.03E-03 (2.20E-04)	9.21E-04 (3.54E-04)
4	DFPP	4	7.03E-04 (4.53E-05)	3.06E-03 (8.74E-04)	3.88E-03 (1.56E-03)
5	DFPP	8	6.87E-04 (3.12E-05)	3.37E-03 (2.44E-04)	4.61E-03 (1.15E-03)
6	DFPP	3	6.84E-04 (2.32E-05)	4.84E-03 (1.14E-03)	6.46E-03 (1.48E-03)
7	DFPP	3	7.16E-04 (5.09E-05)	3.62E-03 (5.43E-04)	3.68E-03 (3.62E-04)
8	DFPP	10	6.85E-04 (2.95E-05)	5.10E-03 (1.09E-03)	6.19E-03 (1.30E-03)
9	DFPP	5	6.81E-04 (4.80E-05)	5.33E-03 (1.90E-03)	6.35E-03 (2.01E-03)
10	DFPP	5	6.81E-04 (2.30E-05)	5.51E-03 (1.46E-03)	6.91E-03 (1.41E-03)
11	DFPP	7	6.90E-04 (3.20E-05)	5.32E-03 (1.07E-03)	5.85E-03 (1.19E-03)
12	DFPP	7	6.77E-04 (4.25E-05)	4.50E-03 (6.43E-04)	5.02E-03 (1.09E-03)
13	DFPP	5	7.05E-04 (6.95E-04)	5.01E-03 (4.93E-03)	6.12E-03 (5.87E-03)

during their entire period, pre- and post-transplant, the values shown in the table are averaged over all sessions. Unfortunately, the treatment used was based on clinical requirements and resources and did not provide equal observations, therefore plasmapheresis measurements predominate. However, it can still be seen from Table 6.1 that all three treatments appear to clear IgM with similar effectiveness, whilst IgG and IgA have much less predictability. As seen with FLC (Chapter 4) clearance of material that flows freely between plasma and EVF is limited by the exchange rate between these two compartments as well as the rate of clearance. It is therefore surprising that IgG and IgA are cleared more rapidly than IgM, even though IgM is limited to plasma pool distribution. This may be down to the size difference between the three isotypes, with IgM being approximately five times larger, which may influence the filter technology

used in the apheresis method.

Apheresis	Clearance mins^{-1} (std. dev.)		
	IgM	IgG	IgA
PA	6.87E-04 (2.51E-05)	7.79E-04 (2.99E-04)	6.62E-04 (4.04E-04)
PE	6.94E-04 (3.71E-05)	6.44E-03 (3.98E-03)	2.94E-03 (1.19E-03)
DFPP	6.90E-04 (3.18E-05)	4.57E-03 (1.46E-03)	5.57E-03 (1.55E-03)

Table 6.2: *Clearance estimates for apheresis treatments, averaged across all patients.*

To allow direct comparison of the apheresis methods the results of Table 6.1 have been further condensed. Table 6.2 shows the average clearance values of each treatment independent of the patient. It should be noted that for plasma exchange only a single patient over eleven sessions is monitored, and for plasma absorption data for two patients from seven sessions were available. This limited sample size should be kept in mind when considering the predicted outcomes, however, from Table 6.2 some interesting conclusions can be drawn. It is believed by clinicians that plasma absorption (PA) is less effective at clearing due to the limited ability of the device to continue absorbing antibodies over the entire treatment period. Whilst this is not proven here, it seems that PA performs relatively consistently across the three isotypes with clearance values in the range 6.62×10^{-4} to 7.79×10^{-4} min^{-1} for IgM, IgG and IgA. Although PE and DFPP clear IgM with similar rates to PA ($\approx 6.9 \times 10^{-4}$ min^{-1}), IgG and IgA have clearance values an order of magnitude greater, with DFPP offering a slight advantage when clearing IgA. This may allow clinical staff to customise the apheresis used depending on the patient's condition, the cost of treatment and the facilities available. For example, it is known (see Chapter 2) that the immune response consists of two phases: primary, which consist of IgM antibodies and secondary which consist predominately IgG; therefore, during the primary phase any of the treatments would be appropriate but for the secondary response PA should be avoided. Alternatively, if the donor-specific response to the implanted organ is measured and contains a majority of a single antibody isotype, the appropriate apheresis method could be used, e.g. PP for IgA, PE for IgG.

6.2 Categorising patients according to immune response

As described in Chapter 2, specific antibodies are produced against invading antigens. After surgery the body perceives the transplanted organ as an antigen and as such produces antibodies specifically against the donated organ; these are referred to as donor specific antibodies, or DSA. Recent measurement techniques have become available that enable DSA to be measured [Higgins, 2007] in isolation from other more general antibodies.

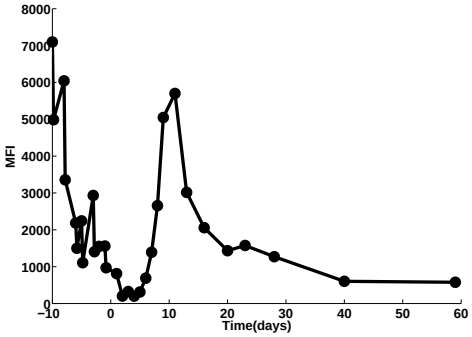
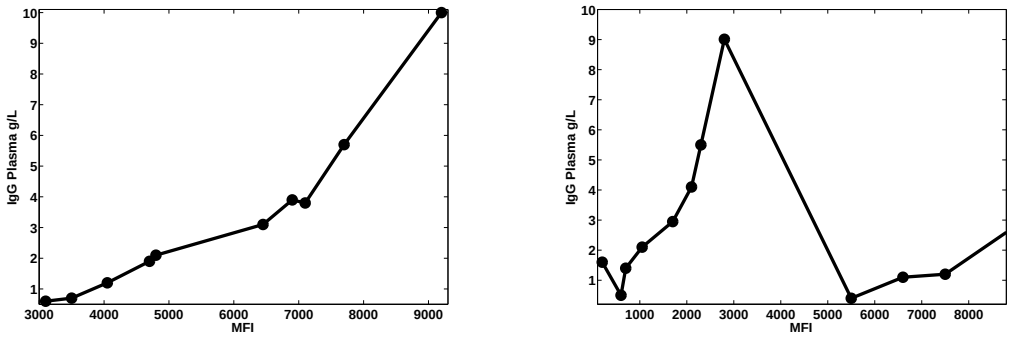


Figure 6.5: *Example Mean Fluorescence Inidicator data.*

The technique uses the Enzyme-Linked Immuno Sorbent Assay, or ELISA, method coupled with photo-reactive agents to provide an indication of the number of antibodies against a particular antigen to be measured. The observation value returned by the measurement technique is the mean fluorescence produced from the application of light to the marker antibodies, known as a Mean Fluorescence Indicator (MFI). An example MFI data set can be seen in Figure 6.5; immediate comparisons can be made with the plasma concentration measurements. The MFI observation exhibits peaks and troughs during the apheresis sessions, with a large increase post surgery indicating an aggressive response to the new organ. However, the MFI scale is not expressed in the required concentration form to allow use of the above process-driven models. This therefore poses a pertinent question: how do you convert between MFI and antibody concentration?

The modelling preferred in this study is that of a process driven approach, i.e. one that inherently relies upon real world states, variables and parameter estimates, for

unknown values or starting estimates for parameter fitting. As such the observations gained must be provided in real world measures (e.g. grams or Mole) in order for parameter estimates and model validation to occur. Unfortunately, the relationship between MFI and quantities of antibodies is difficult to identify. Figure 6.6 shows data sets for two patients, each patient had MFI and concentration measurements taken of antibodies in plasma. The figures plotted are MFI value versus antibody concentration. From these two figures it is evident that the relationship is not linear nor an easily identifiable nonlinear function. In Figure 6.6(a) the relationship can be visualised to be exponential in nature, or linear if a log-transformation is used; however, the data shown in Figure 6.6(b) do not maintain the same relationship between data points. Some of the measurements exhibit a high MFI but a low concentration, with the reverse also holding on several occasions; nor is there a one-to-one mapping between MFI and concentration. The results have been re-tested and neither the MFI nor concentration appear to be incorrect, and it is extremely unlikely to be occurring through noise or measurement error. Efforts to determine this relationship are currently ongoing.



(a) Patient regular relationship between IgG and MFI (b) Patient irregular relationship between IgG and MFI

Figure 6.6: *Patient data showing the unqualified relationship between the MFI values obtained and standard concentration measures from plasma.*

It would, however, seem feasible to make some estimate regarding the patient's concentration in plasma and consider only the percentage change in MFI; this could then simply be converted to a percentage change in concentration. However, the DSA MFI measurement is unable to distinguish between the three antibody types (IgM, IgA and IgG). It is believed that DSA is primarily comprised of IgM and IgG, which as have been seen above, have very different *in vivo* kinetics; thus making a simple

percentage conversion unrealistic.

The MFI data collected are extensive and are believed to carry valuable information that should not be rejected due to loss of conformity with traditional longitudinal analysis methods. Therefore a non-process driven method is used, namely cluster analysis, to investigate the relationship between patients and their response to transplantation.

6.2.1 Cluster analysis

The purpose of the work thus far has been to provide quantitative estimates for biological processes that will allow clinicians to make more accurate predictions on a patient's condition and decisions on future treatment. This requires detailed analysis of the underlying biomedical process and parameters. In this section a more statistical approach is taken, that relies solely upon the measurements without consideration of the *in vivo* kinetics. Data analysis techniques, specifically Cluster Analysis, are used to separate a cohort of patients into categories from which clinicians can infer information regarding the group. Clustering is a technique that allows data sets to be categorised that share common attributes or features as defined by a distance measure. Cluster analysis is a well developed field of statistical analysis with numerous supporting texts (e.g. [Everitt et al., 2001]).

Cluster analysis has been used in the biomedical domain, but predominately in image analysis and segmentation (e.g. [Wismüller et al., 2002] and [Baudelet and Gallez, 2003]). However, using cluster analysis to categorise time series data has been used extensively in economics and engineering for some time (e.g. [Shaw and King, 1992] and [Maharaj, 2000]). In terms of immune systems experiments Spirovskia et al. [2005] used cluster analysis to categorise patients' IgG measurements in cerebrospinal fluid from image data to determine disease conditions. There appears to be little evidence of its use in the analysis of time series patient data in disease or medical complaints that allow periodic samples to taken in plasma. Interestingly, clustering has been used to categorise patient responses to drug treatment in psychology in a similar approach to that which is being taken here [Lipkovich et al., 2008]. Unlike the current study patient response was not based on clinical measurements, but on results of a psychological survey that was taken periodically. From the results the patients were seen to

respond to four distinct cohort groups, from which assessments were made regarding the effectiveness of the treatment.

6.2.2 Method

The processing of time series data by cluster analysis has two distinct phases:

Data set reduction Prior to categorisation, if the data set is large, techniques can be used to reduce the dimension of the data to more manageable dimensions. In this study three reduction techniques are considered:

- Raw Data - biomedical data sets are typically small (less than 100 measurements); as such it is computationally feasible to use the raw data as provided by clinicians, or in the test case, generated *in silico*.
- Principal Component Analysis - reduces the data set by calculating orthogonal components with the effect of maintaining the variance between objects, thus increasing the chance of distinguishing similar patient MFI results.
- Sammon Mapping [Sammon, 1969] - is an algorithm that produces a reduced data set but preserves the distance between objects within the dataset, e.g. the distance between the patient MFI measurements.

Categorisation The data are then collected into categories with similar properties, or clusters, automatically by a clustering algorithm. Clustering algorithms can be separated into two groups: hierarchical and partitioning. Partitioning cluster methods split the group into a predefined number of clusters based on a distance measure. In hierarchical clustering, the data are clustered into cluster trees, or dendograms, separating the data into an increasing number of clusters until a maximum depth is reached.

6.2.3 Implementation

In terms of implementation the clustering was performed in Matlab [Mathworks] and the R statistical package [R Development Core Team, 2008]. Both environments provide robust utilities for clustering, including PCA. The Sammon Mapping algorithm implementation was kindly provided by Cawley and Talbot [2007].

6.2.4 Results

In order to validate the clustering procedures, artificial data sets were generated. After consultation with clinical experts, three hypothetical scenarios were considered:

Response A The patient responds well to immune suppression and plasma exchange and does not produce excess antibodies, post transplant.

Response B The patient has no response to either the immune suppression drugs, or the implanted kidney.

Response C The patient responds well to immune suppression prior to transplant, but post operatively a large response rejecting the kidney is seen that is suppressed with further treatment.

In order to validate the cluster analysis method the IgG model (eqn. (6.1)) was used to generate a sample set of possible observation points (these are shown in Figure 6.7). For simplicity it was assumed that only plasmapheresis was used as the clearance mechanism. To simulate the response to transplant and immune suppression, specified above, three functions were created describing the IgG synthesis. For response A the delayed exponential function seen in Chapter 4 was used; however, unlike chemotherapy it is assumed that immune suppression treatment can reduce the synthesis of antibody to zero, therefore equation (4.12) simplifies to

$$P_G(t) = \begin{cases} P_G(0) & 0 < t \leq t_s \\ P_G(0)e^{-\lambda\, t} & t_s < t \leq t_e \end{cases} . \tag{6.18}$$

where λ is the rate constant required to achieve the half-life of the synthesis decay, t_s the start of effect of the immune suppressive treatment and t_e the end. In the simulated data t_e occurred after the last observation. If the patient has no response to the kidney or immune suppression a simple constant is used ($P_G(t) = P_G(0)$). Describing the patient that has a strong immune reaction to the transplant (response C) an extended version of the logistic decay, described in Chapter 5 (Section 5.3.1) is used, namely:

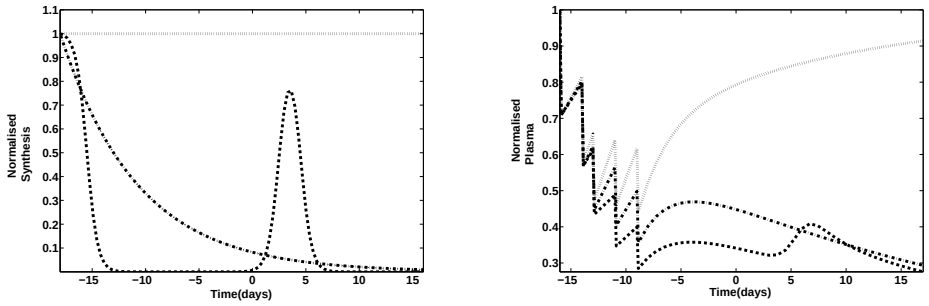
$$\begin{aligned} P_G(t) &= P_G(0) - P_L(t-\tau_1) + P_L(t-\tau_2) - P_L(t-\tau_3) \\ P_L(t) &= \frac{P_G(0)}{1+e^{rt}} \end{aligned} \tag{6.19}$$

with r representing the growth and decay rate due to immune suppression and τ_1, τ_2 and τ_3 describing the times at which the change in synthesis occurs. This function allows for decay, growth and a further decay to be descibed in the IgG synthesis; as is expected of patients responding initially to the tumour then to immune-suppression drugs (for an example please refer to Figure 6.7(a)).

In order to replicate the *in vivo* variability several aspects of the treatment were randomised, using a uniform distribution on the interval shown:

1. Start of treatment (range: 9-20 days before transplant).
2. End of observation (range: 9-20 days after transplant).
3. Number of apheresis sessions (range: 1-7).
4. Initial concentration in plasma, which in turn governs production rate (range: 10 to 30 mg/L).
5. Time that pre- and post-operative effects appear (t_s and τ_1 5-9 days before transplant, τ_2 and τ_3 10-20 days post transplant).

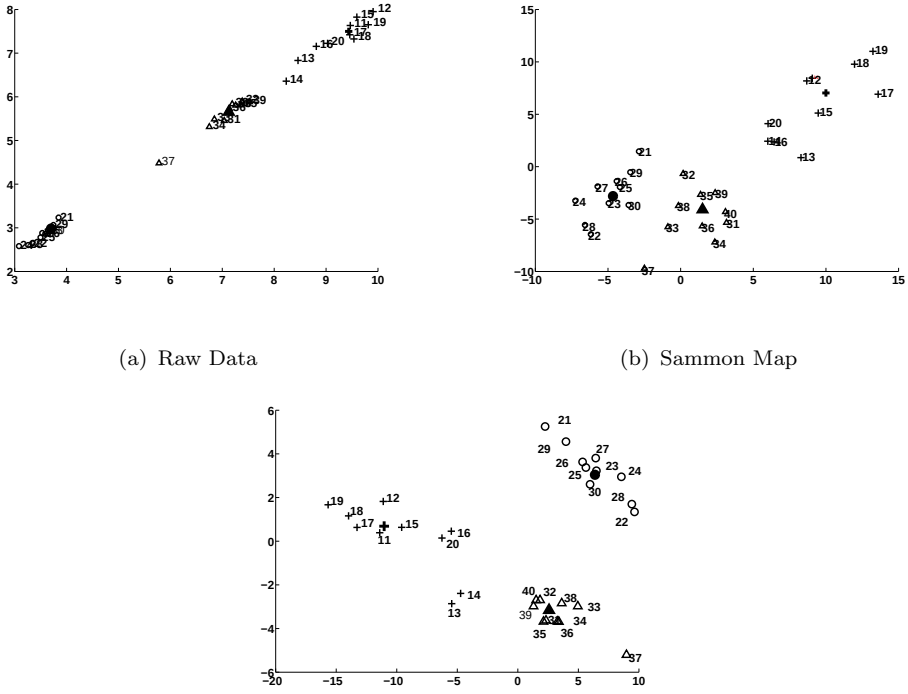
All other parameters are considered constant during the generation of the data sets. Figure 6.7 shows examples of the input functions (Figure 6.7(a)) and resulting patient response (Figure 6.7(b)).



(a) Input functions (b) Example data set generated

Figure 6.7: *Example data sets created with an IgG model using plasmapheresis with randomised treatment parameters (see text). To allow comparison the signal has been normalised to the initial value; chained - patient response A, dotted - response B and dashed - response C.*

In total thirty simulated patient responses were created, ten for each response type (A, B and C as defined above). Once generated, all data sets can be categorised using the cluster algorithm, with the data being pre-processed by one of the dimension reduction techniques. An example of the results of performing this analysis can be seen in Figure 6.8. In each figure the clusters can clearly been seen; although the plots are presented on two dimensional graphs this is purely for simplicity. The raw data have full dimensionality of the measured data, in this case the dimension was 40; the PCA representation has been reduced to two dimensions, whilst the Sammon Mapping is reduced to four.



(a) Raw Data

(b) Sammon Map

(c) PCA

Figure 6.8: *Results from the cluster analysis with generated data sets, only the first two dimensions of each dataset are plotted. At 0% noise, two dimensions for PCA and four for the Sammon Mapping.* ◯ *- patients with response A;* □ *- response B;* △ *- response C.*

In order to test the clustering robustness in the presence of measurement error each data point in the generated datasets was modified with normally distributed noise ranging from 0% to ±20% of the measured value. The clustering process was then performed to determine the number of correctly classified patients, in terms of the three response profiles. This process was repeated one hundred times with randomly added

Table 6.3: *Results of Monte Carlo test of cluster analysis algorithm using generated data. Each cell shows percentage of successful categorisations, with standard deviation expressed in brackets.*

Noise	0%			±10%			±20%		
Raw Data	100 (0)			93.9 (1.2)			53 (3.7)		
Dimensions	**2**	**3**	**4**	**2**	**3**	**4**	**2**	**3**	**4**
Sammon	97 (0.3)	99.7 (0.3)	100 (0)	93.3 (0.8)	94.3 (1.2)	96.3 (0.9)	66.3 (1.0)	71 (2.0)	82.7 (2.7)
PCA	100 (0)	100 (0)	100 (0)	96.7 (0.8)	93.3 (1.4)	90.3 (4.4)	55 (2.2)	54 (1.6)	55.3 (3.0)

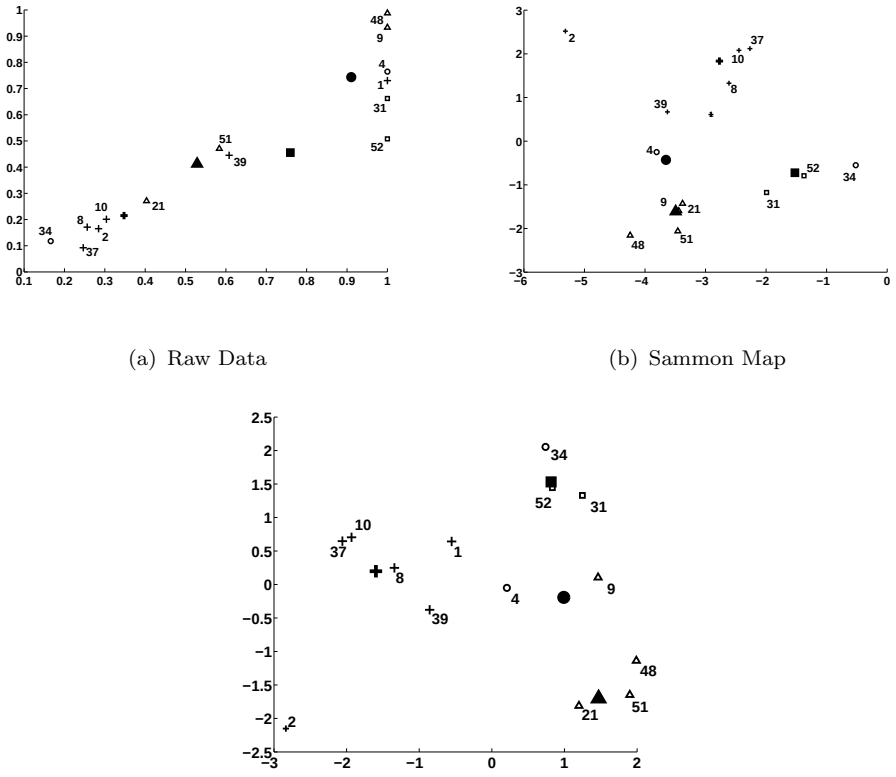
noise. The results of this test can be seen in Table 6.3. All dimension reduction techniques performed extremely well at zero noise level, with even the full dimension data set returning 100% success rates. The Sammon Mapping technique did not produce an exact map until at least four dimensions were used. The PCA approach in general worked better at low dimensions, whilst the Sammon Mapping approach worked well at higher dimensions. At 10% noise levels the three techniques were comparable. However, in the higher noise scenario (20%) the Sammon Mapping shows a surprising resilience in comparison to the other two methods, with both PCA and the raw data techniques showing a large descrease in terms of successful categorisations.

Although the artificial data sets are beneficial in a proof-of-concept exercise, the data used are slightly contrived. In order to validate the cluster analysis against real data a second set was taken from real patient data. Four different types of patients were identified as classic responses to transplantation:

- (LR) Large rises in DSA
- (LF) Large falls in DSA
- (SA) Large quantities of the DSA are absorbed into the implanted organ
- (SM) Modulation of specific DSA post-operatively

Real patient data were selected by clinicians that would be described by one of the above responses, by means other than the MFI data presented. Each category contained between two and four patients. In terms of cluster analysis this experiment is somewhat unusual in that the number of clusters is known *a priori*; therefore, once again the partition clustering algorithm is used, with PCA and the Sammon Mapping dimension reduction emplolyed. Based on the findings in the validation study (Table 6.3) PCA was constrained to two dimensions, and the Sammon Mapping to four.

Figure 6.9 shows a pictorial representation of the clustering when applied to the patient data. The large symbols associated with each cluster (e.g. cross, circle, square or triangle) identify the centre of the cluster, indicating the characteristic of a patient that would be an ideal match in terms of the cluster. The distance from this point indicates how closely the patients response matches the idealised response. The cate-



(a) Raw Data

(b) Sammon Map

(c) PCA

Figure 6.9: *Categorisation the of patients' response, PCA two dimensions and the Sammon Mapping four dimensions.* ◯ *- DSA absorption (SA);* □ *- Larger falls (LF);* △ *- DSA Modulation (SM),* + *- Large Risers (LR).*

gorisation achieved using the raw data can been seen in Figure 6.9(a), the benefit of using dimension reduction can be clearly seen in Figures 6.9(b) and 6.9(c).

Through the graphical representation shown in Figure 6.9 it is difficult to achieve a quantitative description of the results. However, an alternative representation has been provided in Table 6.4, allowing comparison of the cluster analysis categorisation with the separation of patients by clinicians. The underlined numbers outline the incorrectly classified patients.

The results produced by cluster analysis are promising with a high degree of success when categorising the control data and a good degree of accuracy seen in terms of matching the clinicians expert knowledge with real patient data. The necessity for using reduced dimension data is apparent, as can be seen in Table 6.4; using raw data produced five misclassifications compared to two and one with the PCA and Sammon

Table 6.4: *Comparison of categorisations, numbers represent patient ids. The clinical categorisation was provided by experienced medical practitioners. Underlined numbers indicate incorrect categorisation when compared to the clinical diagnosis.*

Category	**Clinical**	**Raw Data**	**PCA**	**Sammon Map**
LR	1 2 8 10 37 39	2 8 10 34 39	1 2 8 10 37 39	1 2 8 10 37 39
LF	31 52	31 <u>34</u> 52	31 <u>34</u> 52	31 <u>34</u> 52
SM	21 48 51 9	<u>1</u> <u>4</u> 9 48	21 48 51	21 48 51 9
SA	4 34	<u>21</u> <u>51</u>	4 <u>9</u>	4

Mapping reduction techniques respectively. However, in this study it has been assumed that each patient fits exclusively into one category, whilst it seems feasible that a patient responds with a large rise in donor specific antibody which ultimately results in high absorption rates onto the transplanted organ.

Unfortunately, the dimension reduction techniques are purely statistical processes, in that there is no relationship with the original system, patient or treatment variables. If this relationship could be formed analytically the cluster analysis would become a more powerful tool in predictive medicine, allowing clinicians to focus treatment on the variables pertaining strongly to a known positive response in patients.

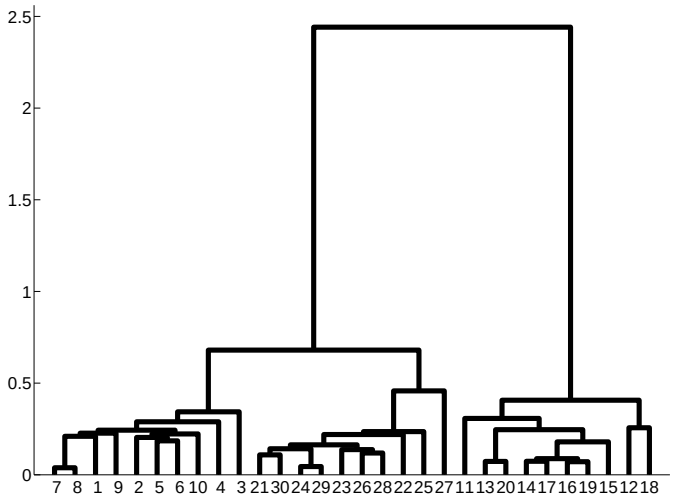


Figure 6.10: *Dendogram of cluster analysis on the generated dataset.*

In the above clustering the number of categories required was given *a priori*; in a true clinical setting, this would not be the case. However, as mentioned earlier an alternative to partition clustering is hierarchical clustering. In hierarchical clustering a tree is generated that splits the data into an increasing number of clusters, until all data sets are categorised into their own cluster. This is implemented by Partition Clustering at each branch in the tree until the terminal, unique, clustering is reached. Hierarchical clusters can be succinctly displayed by a dendrogram, an example of which can be seen in Figure 6.10, from which the breakdown of the generated data sets into an increasing number of clusters can be seen. In a clinical setting, if a large data set was constructed of patients undergoing treatment, combined with patients that have received a transplant and the outcome were known, it may be possible to predict the current patient's long term response, based on their categorisation. For example, if patients were located in a branch of the hierarchical cluster tree in which many patients have rejected the organ it may suggest that the patient's prognosis is poor. However, without the advantages of process-driven modelling and its benefits in identifying a causal relationship between the treatments and the underlying system, such an approach would need significant testing.

6.3 Summary

In this chapter prediction of the immune response of patients undergoing transplants has been addressed. The models developed in the previous section were used to evaluate and validate the clustering techniques, through the generation of datasets with feasible anatomical features, as seen in patients undergoing kidney transplants. Finally, clustering technique were used to classify data collected from kidney transplant patients according to categorisation specified by clinical collaborators. Using cluster analysis an expert system has been suggested that may allow clinicians to use the data to predict patients' long term prognosis through the categorisation of a cohort study of patient response.

In addition, it has been shown how discretised deconvolution techniques can allow clinicians and laboratory staff to estimate the synthesis of antibody production in response to kidney transplants. This required the development of three models to

describe the *in vivo* kinetics, one for each antibody type of interest. Each model had slightly different dynamics, but through assumptions and linearisation, analytical solutions are available to enable estimates for concentrations in the immeasurable EVF compartment and the synthesis. The simplified deconvolution method enables an uncomplicated implementation, that is suitable in both time and complexity, to be used by medical staff without extensive training or non-standard tools. The mechanics of deconvolution will be explored further in the following chapter; however, rather than assuming that the synthesis is discrete over a fixed time period, to ensure linear solutions to the system a more flexible method will be used, namely that of regularisation, which allows the synthesis to be predicted non-parametrically over the entire treatment period.

Chapter 7

Non-parameteric estimation of model parameters

In the previous chapters several models have been presented that describe immunological kinetics *in vivo*; two types of diseased conditions have been described: Free Light Chain (FLC) and Intact IgG secreting tumours. In both cases the synthesis of these proteins was assumed to be known; for FLC a constant production was assumed, whilst for intact IgG it was described by a logistic equation. However, while these conditions may hold for short periods, the long-term prospect of the synthesis obeying these assumptions is unlikely.

A question naturally arises: what if the functional form of the system input is not known? In this situation a parametric form for the synthesis cannot be provided or assumed. This was indeed the case for the transplant patients, as seen in Chapter 6. In that case, a simplistic method was described that assumed synthesis was constant over a known period yielding a piecewise constant function that could be estimated, given a known form for the kinetics of the antibody observed. The process of determining the input in these cases is referred to as deconvolution, and the piecewise function approximation used is known as discrete deconvolution [Iga et al., 1986; Sparacino et al., 2001]. In this chapter a method is presented that assists in the identification of model inputs using a non-parametric approach. Although the focus of this work is to determine the effects of chemotherapy on plasma tumours, the deconvolution techniques were found to be suitable to several other applications. Subsequently, the

same method will be used to recover other unknown non-linear aspects of models in a non-parametric fashion.

An interesting extension to deconvolution is a process known as 'Blind Deconvolution' [Ayers and Dainty, 1988; Holmes, 1992], a method which attempts to recover the input signal when the output is measured, but the system response is also unknown. The technique used in Chapter 6 can be considered a simple Blind Deconvolution, as elements of the system response and input signal were determined in parallel. However in general, such techniques would allow a more general approach than that described in this chapter, which requires that model structure and parameters are assumed known. Blind deconvolution methods have been pioneered by the optics and astronomical community, where strong assumptions can be made about the form of the system response; in most cases a point-spread distribution of blur on an image is assumed. Unfortunately, for biomedical systems such strong assumptions are often not valid, therefore a more process-driven model-based approach to the system description is employed; namely, that of a compartmental structure as has been seen previously.

7.1 Deconvolution

Given a known system response ($h(t)$) and a known input ($u(t)$) the output of the system ($y(t)$), with zero initial conditions, can be given by the convolution integral

$$y(t) = l(h(t), u(t)) \quad = \int_0^t h(t-\tau)u(\tau)d\tau \tag{7.1}$$

Conversely, if the output ($y(t)$) and system response ($h(t)$) are known, and an inverse function ($l^{-1}(\cdot,\cdot)$) is available, the input function can be calculated. To derive the input in this fashion is known as deconvolution. This process fits in naturally with the models generated in the previous three chapters. As the output is measured the system dynamics are identified through a process model, but the input is unknown due to the perturbations enforced by the drugs used to control the immune system generation (e.g. chemotherapy and immuno-suppression drugs). However, for the immunological conditions under consideration for this study additional constraints are forced upon the deconvolution process. The following must be considered:

1. The system has non-zero initial conditions. Any immunological protein to be measured will already exist in the patient prior to observation.

2. The system is not linear or time-invariant. Most patients will either be undergoing an apheresis treatment which is non-linear and switched, either periodically or aperiodically, during a typical treatment, e.g. three times a week in addition to ad-hoc treatments if the concentration levels in plasma become dangerously large.

Several deconvolution techniques have been considered for biomedical systems [Cutler, 1978; Hovorka et al., 1998; Pedersen, 1980]. All of these assume a linear time-invariant system with non-zero initial conditions. Furthermore, they assume a known functional form for the input function (e.g. polynomial [Cutler, 1978] or quadratic [Hovorka et al., 1998]) which may be appropriate for the drug synthesis but is not applicable for the *in vivo* synthesis of antibody proteins. However, all of these techniques are based on the closeness of the simulated output to the data in a least-squares sense, a concept that is expanded below (see Section 7.2).

A deconvolution technique that has been successfully employed with the above constraints in mind is Maximum Entropy [Charter and Gull, 1991; Hattersley et al., 2008; Madden et al., 1995]. This approach will be described below, but Maximum Entropy can be constructed as a regularisation technique that can function under the criteria of high noise, sparse and irregularly sampled data that are commonplace in biomedical signals, without the need to assume a functional form for the input signal. In addition, it will be shown how a time-variant and non-linear compartmental model can be integrated into the Maximum Entropy framework to produce simulation results that show good comparison with patient data and produce clinically viable immune complex synthesis. The Maximum Entropy method is a subset of a large group of deconvolution methods known as Regularisation [Engl et al., 1996].

7.2 Regularisation

Simply stated regularisation is a method of translating an ill-posed problem into a more tractable well-posed one by approximation or constraints. Consider a discrete

linear system, with zero initial conditions,

$$y = Fu, \quad F \in \mathbb{R}^{m \times n} \tag{7.2}$$

where y is an output vector, length m, F a convolution matrix and u the vector of the unknown input, of length n. If a noisy output of the system ($\hat{y}$) is available it would be expected that the predicted output y be sufficiently close to the measured output, accounting for noise (ϵ)

$$||y - \hat{y}||_2 < \epsilon. \tag{7.3}$$

A naive solution to this, if F is square and invertible, is to ignore the presence of noise in the measured signal ($\hat{y}$) and to invert the convolution matrix to give an estimate for the input

$$\hat{u} = F^{-1}\hat{y}; \tag{7.4}$$

however, due to the noise present the solutions of this process are ill-conditioned, leading to negative values and oscillations. Alternatively, to recover the vector u a least squares minimisation approach could be employed

$$\min_u ||\hat{y} - Fu||_2, \, F \in \mathbb{R}^{m \times n} \quad m > n \tag{7.5}$$

where u are parameters relating to the input function, not system parameters as would be seen in a standard parameter estimation procedure. This may still result in an ill-posed and ill-conditioned problem [Sparacino et al., 2001] as the optimisation process is free to alter the parameters, possibly into unfeasible regions, to obtain the closest fit to the measured output. To alleviate this problem *a priori* knowledge of the input function is used to form additional constraints on the solution in order to stabilise the deconvolved input. The constraints are formed by an additional function, that will provide a mapping from the input vector u to a scalar value which encapsulates the *a priori* information regarding the input signal, referred to as $g(u)$ in the following text.

Using regularisation is beneficial when recovering biomedical signals due to the existence of two obvious constraints: Firstly, due to the distributive nature of drug kinetics the *in vivo* immune system response in terms of the synthesis of the immune complex would be 'smooth' over the measured response time. Secondly, the input signal is by definition positive, and a negative synthesis is already accounted for in the model by clearance.

In order to deconvolve the input vector u from noisy data both the constraint $g(u)$ and distance metric between the observered and simulated output $f(u)$ need to be minimised, which is a constrained optimisation problem.

$$\min_u f(u) \quad \text{subject to} \quad g(u) = c. \tag{7.6}$$

It should be noted at this point that the object of this work is not to evaluate constrained optimisation, nor evaluate the performance of algorithms to solve the problems specified. A successful outcome is a working prototype that demonstrates the use of deconvolution to highlight possible synthesis and chemotherapy effects. Nonetheless, the question remains as to the form the distance metric ($f(u)$) and constraint function ($g(u)$) should take. These will be explored in the following section.

7.2.1 Maximum Entropy signal recovery

The use of Maximum Entropy to recover signal data originates in the field of image analysis [Cornwell and Evans, 1985; Skilling and Bryan, 1984], where it is used to filter images to remove high-levels of noise. However, it has also been used in the recovery of biomedical signals [Charter and Gull, 1987]. An excellent review and comparison of methods can be found in Madden et al. [1995]. The 'entropy' in the title refers to the amount of uncertainty present in a signal. The concept behind Maximum Entropy is to produce an input signal that maximises the uncertainty, thus creating a signal that has the minimal assumptions whilst maintaining an acceptable fit to an observed output. If the input signal ($u(t)$) is discretised into piecewise function, the entropy of the signal can be used [Skilling and Bryan, 1984]

$$x_i = \frac{u_i}{\Sigma u_i}, \quad S(x) = -\sum_{i=1}^{N} x_i \ln(\frac{x_i}{r_i}). \tag{7.7}$$

where x_i presents the normalised value of the input at the i^{th} sample, and r_i is a baseline value that the production should take in the presence of no other information. The values of x_i are assumed to be positive, since a negative production is unfeasible. In the above equations r_n is calculated using a nearest neighbour average $((x_{i-1} + x_{i+1})/2)$. This discourages adjacent elements of the input function from large variation, thus smoothing the recovered input signal. At the sample points $i = 0$ and $i = N$ the average is taken of adjacent samples, e.g. $(x_0 + x_1)/2$.

During the Maximum Entropy signal reconstruction the entropy of the input signal is maximised under the constraint that the output of the system should 'match' the real data measurements. A common form used to model this constraint is the χ^2 metric. This is given by:

$$\chi^2 = \sum_{i=1}^{N} \frac{(D_1[i] - c_1[i])^2}{\sigma_i^2}, \quad E(\chi^2) = N \tag{7.8}$$

where E denotes the expected value, N the number of samples, and $D_1[i]$ and $c_1[i]$ are the measurements taken and the predicted value at the i^{th} sample, $t = \tau_i$. This can be seen as a weighted least squares estimator with weights of $\frac{1}{\sigma_i}$.

7.3 Interfacing to compartmental models

In the methods discussed above it has been assumed that the simulation results can be expressed in the standard convolution form, equation (7.4). All the modelling conducted thus far, in terms of patient kinetics, has been implemented through compartmental modelling which ultimately results in a state-space model, i.e. a system of ODEs, of the form

$$\begin{aligned} \dot{q}(t,p) &= f(q(t,p), u(t,p), p) \\ y(t,p) &= h(q(t,p), p) \\ q(0,p) &= q_0(p) \end{aligned} \tag{7.9}$$

where $q(t,p)$ are the state-variables, $y(t,p)$ are system outputs, $u(t,p)$ the inputs to the system and p a vector of constant parameters. If a single-input-single-output linear time-invariant model structure with a linear observation function is first considered, the system can be described in the familiar form given by:

$$\begin{aligned} \dot{q}(t) &= Aq(t) + Bu(t) \\ y(t) &= Cq(t) \\ q(0) &= q_0 \end{aligned} \tag{7.10}$$

where A, B and C are matrices of the parameter vector p only. Using the matrix exponential form an analytical solution of the system can be found

$$y(t) = Ce^{At}q_0 + C\int_0^t e^{A(t-s)}Bu(s)\,ds. \tag{7.11}$$

If the effect of the initial conditions is removed from the measured system output ($\hat{y}(t)$), to give a translated observation function ($\bar{y}(t)$)

$$\bar{y}(t) = \hat{y}(t) - Ce^{At}q_0 = C\int_0^t e^{A(t-s)}Bu(s)\,ds \tag{7.12}$$

then the standard convolution integral as expressed in equation (7.1), is obtained. Further, if it is assumed that the input function $u(t)$ is piecewise constant over the period of sampling, equation (7.12) is equivalent to the discrete convolution

$$\bar{y}[n] = \sum_{i=1}^{n} h[n-j]u[j]\Delta_j, \tag{7.13}$$

where $h[j]$ is the discretisation of the impulse response, $u[j]$ the discretised input function and Δ_j the change in time between samples. An interesting feature of the convolution matrix is that due, to the causality of the system (i.e. future states depend only on the current state and past inputs, the matrix must be Toeplitz [Sparacino et al., 2001]). The elements of the convolution matrix can in turn be expressed by [**?**, Hovorka1998]

$$F_{ij} = \begin{cases} 0 & \text{if } i > j \\ C \int_{\tau_{j-1}}^{t_i} e^{A(\tau_j - s)} B \, ds & \text{elsewhere} \end{cases} . \tag{7.14}$$

where τ_j and t_j are the sample points for the input and output signals respectively. This enables a linear compartmental model to be expressed in a convolution matrix form, equation (7.4). If the convolution matrix can be constructed there are non-iterative approaches available [Hansen, 1998] to deconvolve the input signal; a toolbox that is freely available which has implemented several of these methods is described in Hansen [2007].

Unfortunately, this approach has several limitations that have prevented it from being used for the deconvolution of antibody synthesis:

- If the system is non-linear, an analytical solution in the form of equation (7.11) is not available.
- For an accurate solution, a small time step (Δ_j) between samples is required, or the approximation of the input $u(t)$ being constant over the period will be invalid. For biomedical *in vivo* measurements this is not feasible, taking, for example FLC patients on dialysis, during treatment a sample every 30 minutes for a 6 hour treatment would be too resource-intensive for clinical staff.
- Creation of the zero initial conditions system (eqn. 7.12) requires the use of measured data, which contain noise; from this it is possible to generate observations ($\bar{y}(t)$) that are negative.
- In a time-variant system, e.g. a switched system in the case of apheresis, the convolution matrix will not be continuous and a new matrix is required for each period.
- The construction of the design matrix is a non-trivial process in all but the simplest of compartmental models; for example, in models with more than a single compartment the analytical solution of equation (7.11) leads to a state-transition matrix that uses the Matrix Exponential, rather than the exponential function. This quickly leads to difficulty in implementing the convolution matrix and is a process prone to error and difficult to adapt in the light of model changes.

- This method is applicable to single-input single-output systems; however, it is not applicable for multi-input or multi-output systems (for an example, see Section 7.5.2).

To account for these issues an obvious, but effective, alternative is available. The solution to the system of ODEs (eqn. (7.16)) can be evaluated through existing robust numerical solutions (e.g. Runge-Kutta, Gear's algorithm, Predictor-corrector methods) [Chapra and Canale, 2002, Ch. 25-26]. Using standard packages to implement the ODE enables the non-linearity and time-variant nature of the models under consideration to be easily accounted for. In addition, the output of the ODE solver can be sampled at the same time as the measured data, even accounting for non-equally spaced time intervals.

The use of variable step numerical integration algorithms does present some problems in terms of evaluating the input function, in that, at any time point within the sampled period, a value for production must be available to the solver. This can be solved using interpolation techniques to approximate the input function at the required time point. Care should be taken in the choice of interpolation method to ensure the constraints of smoothness and positivity are not violated during calculation of the interpolated input value; for the work presented here simple linear interpolation proved sufficient.

7.4 Optimisation implementation

The purpose of this work is not to determine the best optimisation routine, nor to investigate the different characteristics of the available optimisation algorithm. The goal of the present work was to arrive at a more pragmatic methodology, searching for a useable solution to assist in the clinical difficulties related to therapeutic effects of chemotherapy treatment. However, it is necessary for clarity to review the basic approach taken and the algorithms and libraries used to solve the deconvolution problems.

In order to implement the input recovery, the process can be viewed as a constrained optimisation problem, given by:

$$\min_{u} -S(u) \quad \text{subject to } \chi^2(u) = N \qquad u_i \geq 0 \tag{7.15}$$

where $-S(u)$ is the negative of the regularisation function, to allow for minimising, as opposed to maximising the entropy. There are several algorithms available for solving this problem (see for example [Charter and Gull, 1987; Cornwell and Evans, 1985; Fletcher, 1987; Skilling and Bryan, 1984]) the majority of which require the derivative of the χ^2 metric to be known. Unfortunately, due to the non-linearity present within the systems this is not available. The implementation is therefore restricted to derivative free optimisation techniques. In all the algorithms used in this study a numerical approximation is used to estimate the derivative of the function to be optimised. Good results have been obtained using a Sequential Quadratic Programming (SQP) technique. The SQP algorithm is implemented in the *fmincon* method of MATLAB, which uses a Quadratic Programming sub-problem coupled with calculation of the Hessian of the Lagrangian of the objective function ($S(u)$) and the constraints ($\chi^2(u)$ and positivity of u) via the Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula [Mathworks]. For validation, alternative methods have been used to perform the derivative free optimisation (e.g. COBYLA [Powell, 1998]). Each of the constrained optimisation algorithms used produced similar results with computational performance being seen when compiled code was used, as opposed to the interpreted methods. An implementation of the Maximum Entropy routine can be found in Appendix F.

It should be noted that the optimisation techniques discussed are all local in nature, in that they do not attempt to search for the global minimum of the solution. As such, in each example shown in Section 7.5 the deconvolution was tried from a range ($\pm 10\%$) of initial conditions around an estimated constant input value. Global optimisation techniques are available [Mongeau et al., 2000; Ratschek and Rokne, 1988] but have not been considered due to computational complexity and the availability of *a priori* knowledge regarding the input function giving a feasible start point.

Prior to deconvolving the synthesis on patient data several validation checks were performed. The results produced were comparable with those seen in Madden et al. [1995]. Madden et al. [1995] define a series of test functions, and an example model, that

can be used to assess the performance of different deconvolution algorithms. However, in the Madden work constraints were placed regarding the initial and final values of the input signal, i.e. it was assumed that the signal at $t = 0$ and for several points around the final sampled point were known to be zero. This restriction was not required to produce equivalent results using the Maximum Entropy implementation. In addition, the results were not 'smoothed' via post-processing (e.g. averaging) as in the Madden et al. [1995] study. Further details on the validation of the Maximum Entropy implementation can be seen in Appendix E.

7.5 Deconvolution results

The deconvolution methods were initially developed to investigate the effect of chemotherapy on Multiple Myeloma patients, the results of which are shown in the first part of this section. Both FLC and Intact IgG Myeloma patients are considered. In the last section a series of results are described that are a consequence of interaction with other projects that required deconvolution to esimate non-chemotherapy related functions, and in one instance a non-biomedical problem, showing the versatility of the simple implementation of the regularisation methods.

7.5.1 Patient response to chemotherapy

As has been discussed in previous chapters, patients with gammopathic conditions received medication to reduce the antibody synthesis. For kidney transplant patients these are drugs to suppress the entire immune system; whilst for Multiple Myeloma patients, combined chemotherapy is used. Irrespective of the mechanics of the drugs course the effect is simply to modulate the underlying production. In the following two sections the Maximum Entropy method is used to investigate the effect of treating Multiple Myeloma patients with FLC and Intact IgG secreting tumours. Although the use of deconvolution is used to elicit input signals, the definition of the 'input' allows some flexibility, as is shown in Section 7.5.2, where the impact of increased proteins levels on the viscosity of plasma are investigated.

Free Light Chain patients

During the Maximum Entropy deconvolution process the input signal ($u(t)$) is modified to allow $E(\chi^2) \rightarrow N$ as the iterations increase (see Section 7.4). At each iteration the plasma concentration ($c_1(t)$) is estimated using a numerical solver for the new values of $u(t)$. This is then sampled at the appropriate points for the concentration measurements (e.g. $D_1[1], \ldots, D_1[N]$) and χ^2 is calculated. To produce the simulated output ($c_1(t)$) the FLC model described in Chapter 4 is used; as all the patients considered below are FLC-Kappa, the parameter estimates associated with this type of FLC are provided in Table 4.5. The initial estimate for the FLC generation rate is chosen to be constant over the treatment period to ensure minimal assumptions regarding the input form. The input function is sampled at the same time points as the data given by clinicians; however, there are no limitations on the number or location of input samples beyond that of computing resource. The initial values for the input function are calculated from the steady-state input required to maintain the concentration of the first measurement, without haemodialysis.

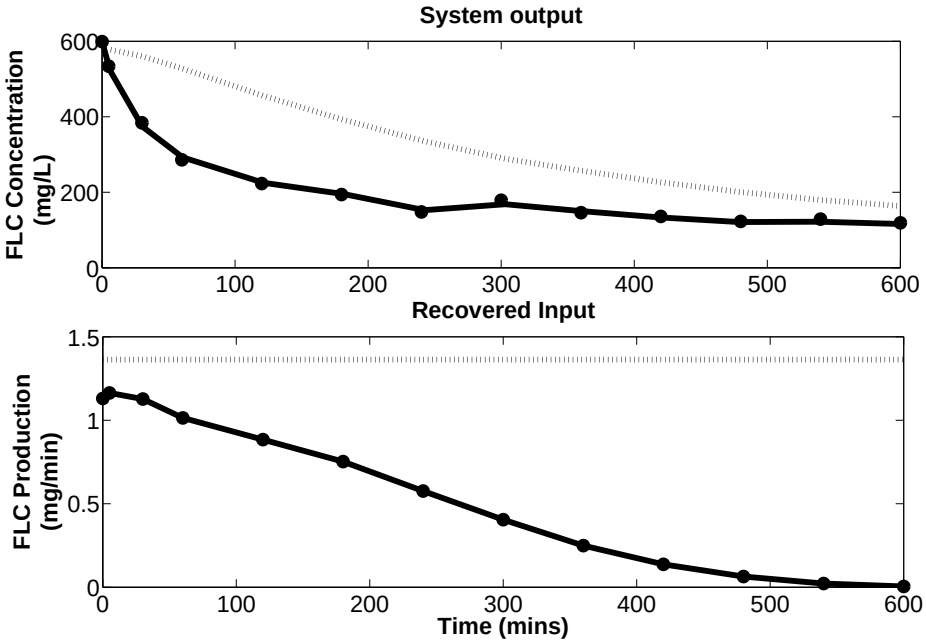


Figure 7.1: *Recovered FLC input signal for patient one (Pat1). Top (solid - simulated FLC concentration in plasma, dashed - FLC concentration in EVF, dots - measured FLC concentration in plasma). Bottom (solid - deconvolved input, dotted - initial estimate, dots - input sample points)*

The results of the Maximum Entropy reconstruction can be seen in Figures 7.1 to 7.4. Each figure consists of two graphs. The top graph shows the results of applying the recovered input signal to the system equations and the simulated output of plasma (solid line) and EVF (dashed-line) concentrations . The circles in the top graph indicate the measured values in plasma, as provided by the clinicians.

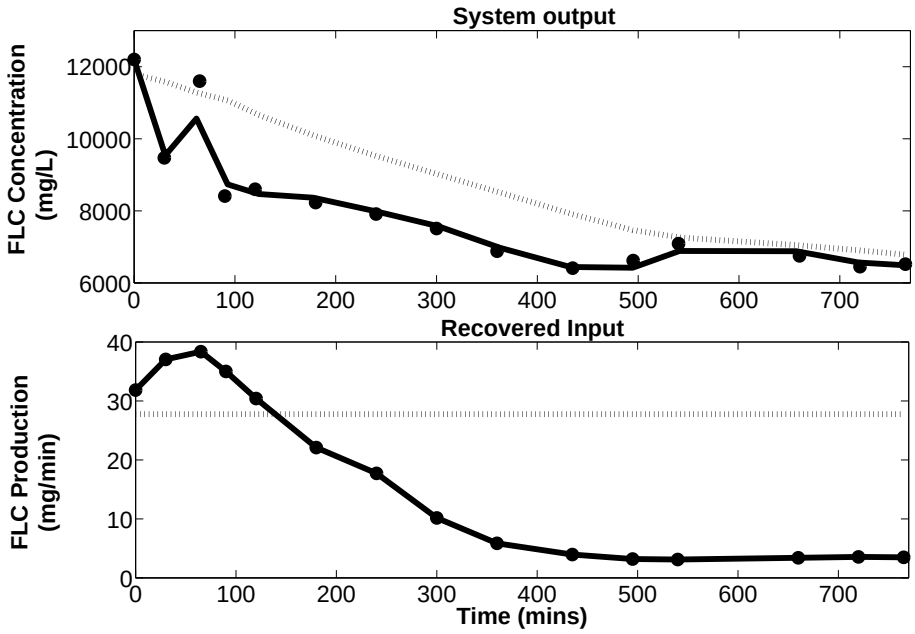


Figure 7.2: *Recovered FLC input signal patient two (Pat2). Top (solid - simulated FLC concentration in plasma, dashed - FLC concentration in EVF, dots - measured FLC concentration in plasma). Bottom (solid - deconvolved input, dotted - initial estimate, dots - input sample points)*

In the bottom graph the deconvolved input signal is shown (solid-line). The dotted line denotes the FLC production calculated from the initial plasma measurement, assuming the production is constant and the system is in steady-state at time $t = 0$. This is used as the starting value for production ($f(t)$) during the Maximum Entropy optimisation.

It is of interest to note the point at which the patient received chemotherapy in relation to the measured haemodialysis session. For all the patients shown the dialysis session monitored is the first session received since the previous chemotherapy treatment. Patients three (Figure 7.3) and four (Figure 7.4) received chemotherapy within an hour of the dialysis treatment; whilst patients one (Figure 7.1) and two (Figure

7.2) both had chemotherapy over 24 hours before dialysis treatment. Unfortunately information regarding the particular chemotherapy drugs used is not available.

As can be seen in Figures 7.1 and 7.2, the results suggest that chemotherapy treatment is still effectively reducing the FLC generation rate in the two patients. In patient Pat1 a gradual decline in synthesis occurs over the entire treatment period of 10 hours, even though the chemotherapy was applied 24 hours previously. It should also be noted that FLC synthesis has reached approximately zero in this patient, indicating an excellent response to the treatment. The overall trend for patient two is similar, in that the trajectory over the measured period is generally downward, with a basal level reached after 500 mins; however, patient Pat2 did not reach a synthesis rate that was below the FLC production expected in a healthy individual (0.22 mg/min see Chapter 4). This suggests that the chemotherapy has not completely destroyed all the tumourous plasma cells. Patient Pat2 exhibits some unusual behaviour in regard to the FLC concentration seen in plasma; the third observation, rather than falling (as would be expected when the patient is on haemodialysis) shows a significant increase, almost reaching the level of the initial concentration. According to the clinical notes there was no break in haemodialysis during this period nor was the filter changed at any point; in addition, the sample was re-measured suggesting it is a genuine increase in the FLC generation. The reason for this increase is unclear; it has been noted by clinicians in other patients, but unfortunately no data were available for comparison. However, the deconvolution technique shows the synthesis trajectory that would be required to achieve such a rise. It is interesting to note that the deconvolution method does not predict a simulated output that is as close to the measured value as would be expected. This is due to the 'smoothness' constraint and the number of input sample points chosen. If it were imperative that this measurement were more closely simulated the constraint could be relaxed or the number of sample points increased around the point of greatest curvature, to allow for 'peaks' in the input signal. These issues are covered in more detail in Chapter 8.

As mentioned previously patient Pat3 had chemotherapy immediately before dialysis; the effect of this on synthesis can be clearly seen in Figure 7.3. During the first thirty minutes of dialysis the deconvolved synthesis shows a significantly elevated rate of reduction to that seen in patient Pat1. It is, however, interesting to note that after

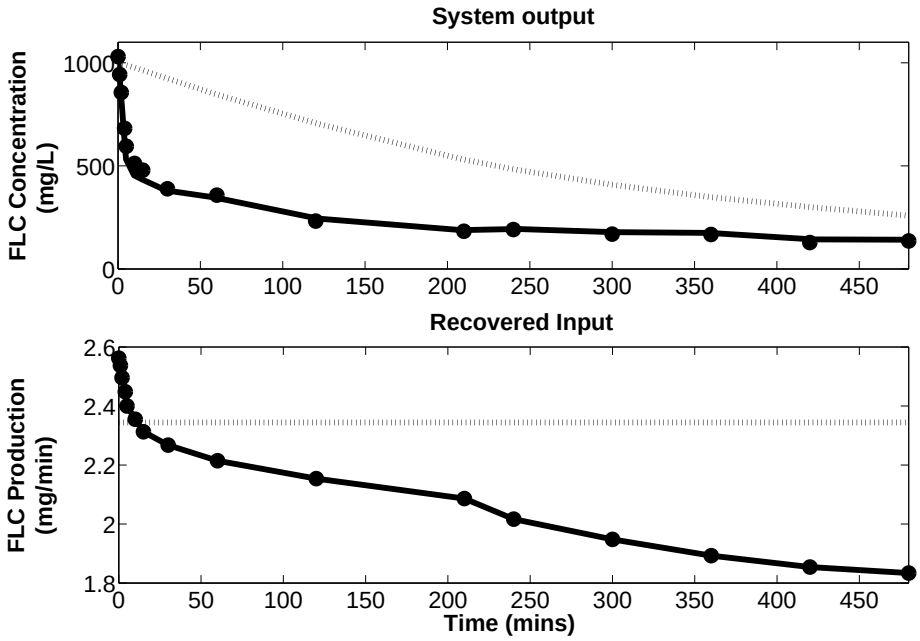


Figure 7.3: *Recovered FLC input signal patient three (Pat3). Top (solid - simulated FLC concentration in plasma, dashed - FLC concentration in EVF, dots - measured FLC concentration in plasma). Bottom (solid - deconvolved input, dotted - initial estimate, dots - input sample points)*

this initial period (first 30 mins) the synthesis declines more gradually.

For the final patient Pat4 (Figure 7.4) plasma concentration data are available for an extended 16 hour period, including a 6 hour dialysis session, which started at $t = 0$, and 10 hours off dialysis. Patient Pat4 also had chemotherapy within an hour of haemodialysis and as with patient three the synthesis reduction is quickest early in the treatment ($t < 100$ mins), continuing to fall until approximately two hundred minutes. However, unlike the previous three FLC patients the synthesis does not show a decline over the entire treatment and even before the end of dialysis the deconvolved synthesis suggests an increasing generation rate of FLC, indicating a poor response to chemotherapy. As in patient Pat2 there is an unexpected increase in plasma concentration between the second ($t = 30$ mins) and third ($t = 60$ mins) measurements; however, after referring to clinical notes there was a break in dialysis which would account for this rise. This change in clearance is described by a discontinuity in the solution to the system equations.

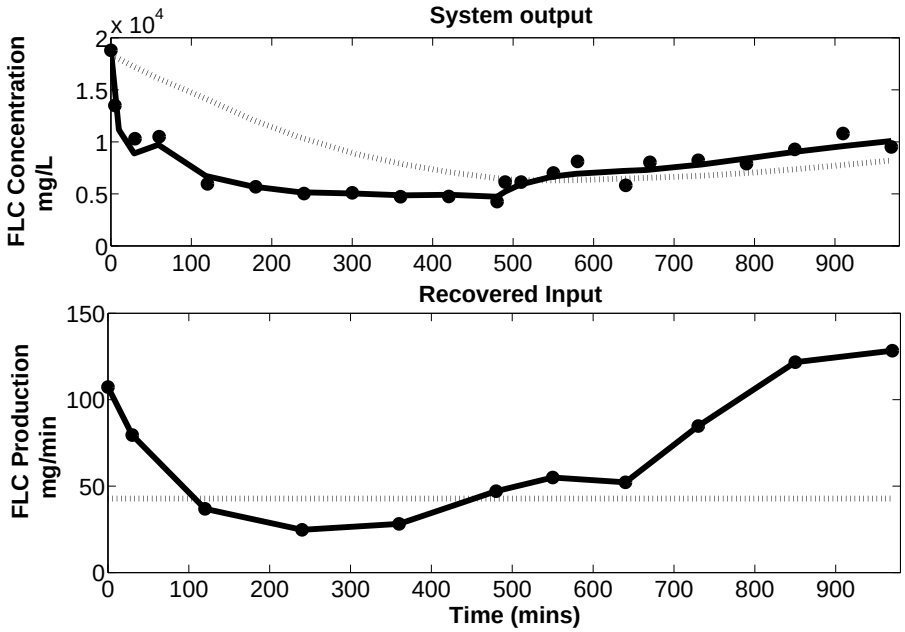


Figure 7.4: *Recovered FLC input signal patient four (Pat4). Top (solid - simulated FLC concentration in plasma, dashed - FLC concentration in EVF, dots - measured FLC concentration in plasma). Bottom (solid - deconvolved input, dotted - initial estimate, dots - input sample points)*

Intact IgG MM patients

As stated in Chapter 5, patients with intact IgG also undergo aggressive chemotherapy in order to reduce the number of tumourous plasma cells that are secreting intact IgG. As such a similar method of deconvolution as that described for the FLC producing Myeloma patients is used to predict the change in synthesis. The simulated output $c_1(t)$ for the χ^2 metric is produced by solving the quasi steady-state model, as described in Section 5.3 of Chapter 5. The model was initialised with the parameters defined in Table 5.3. After discussion with clinical staff it is believed that chemotherapy affected only the tumourous synthesis; as such the non-tumourous production of normal antibodies was assumed to be constant over the treatment period, and only the tumourous IgG (IgG-Kappa for both the patients considered) was estimated through deconvolution.

The results of applying the deconvolution process to predict the tumourous IgG-Kappa synthesis for two patients can be seen in Figures 7.5 and 7.6. The figures are displayed with the same format as above, i.e. simulated output and measurements in

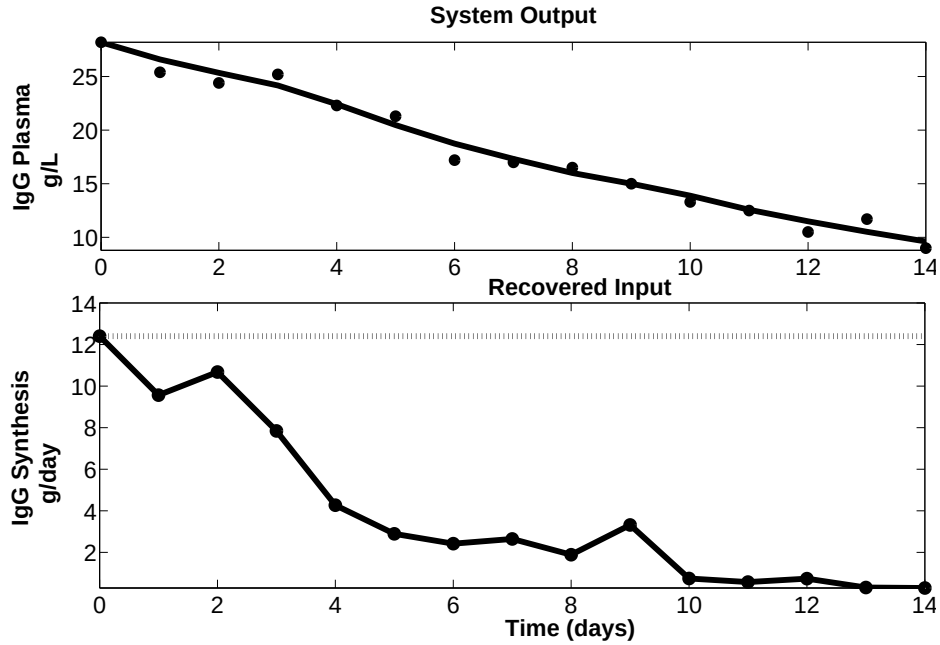


Figure 7.5: *Recovered synthesis for Intact IgG-Kappa patient 1. Top (solid - simulated IgG-Kappa concentration in plasma, dashed - IgG-Kappa concentration in EVF, dots - measured IgG-Kappa concentration in plasma). Bottom (solid - deconvolved input, dotted - initial estimate)*

the top-pane, with deconvolved synthesis and starting values shown in the bottom-pane. Patient 1, as seen in Figure 7.5, was observed over a short period; as this is intact IgG secreting Myeloma no apheresis was used, therefore measurements were only available on a timescale of days, rather than minutes as has been seen previously. The deconvolved input, as with the FLC patients, shows a noticeable decline in the synthesis rate over the entire treatment period, with two slight increases at days two and nine; this is most likely due to measurement noise. Over the time period measured it would seem plausible when considering a parametric form for the input function, to assist in predictive analysis of treatment, that a simple exponential function may be appropriate as was used for FLC synthesis in Chapter 4.

However, if we consider an alternative patient (see Figure 7.6) the necessity for using the delayed logistic equation becomes apparent. Patient two was observed over a much longer period, approximately 2 years. In this case although the patient received chemotherapy prior to the start of the observation period ($t = 0$) the predicted synthesis does not show a significant reduction until approximately 3 months later, at which point

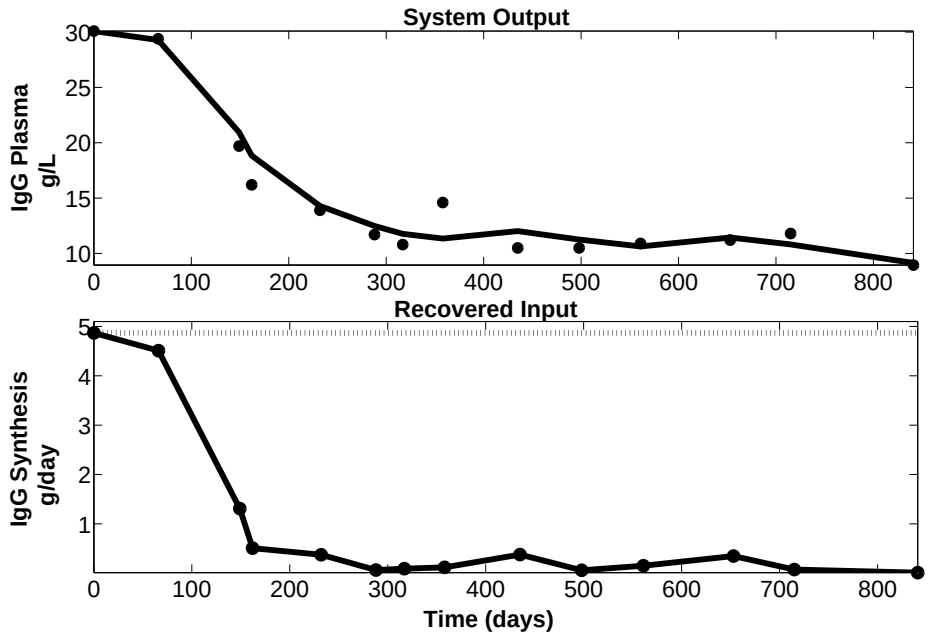


Figure 7.6: *Recovered synthesis for Intact IgG-Kappa patient 2. Top (solid - simulated IgG-Kappa concentration in plasma, dashed - IgG-Kappa concentration in EVF, dots - measured IgG-Kappa concentration in plasma). Bottom (solid - deconvolved input, dotted - initial estimate)*

synthesis falls dramatically down to a baseline level less than 1 g/day. From discussion with clinical collaborators, whilst this is not an uncommon feature of the chemotherapy effects on intact IgG secreting MM patients, it is not fully understood. Analysing this would require additional datasets to be made available for a variety of patients. It is therefore the focus of further research and is discussed further in Chapter 8.

7.5.2 Volume changes in IgG patients

Patients with IgG secreting Myeloma often present with an increased viscosity of plasma due to the elevated concentration of proteins generated by tumourous cells [Tuddenham and Bradley, 1974]. This results in increased plasma volumes that evolve as the tumour synthesis changes. However, in the clinical study for this project the facilities were not available to measure the viscosity of each sample. In some patients unexplained changes in the non-tumourous IgG concentration were evident that were unlikely to be due to synthesis or clearance, as has been previously encompassed in this study. The following deconvolution was performed to determine whether such changes

in plasma volume could account for this phenomena.

Whilst this is a departure from the concept of deconvolving the input in a classic sense, with a simple reformulation of what is considered as an input the same deconvolution methods described previously can be used. As has been done previously with synthesis, the volume was discretised and treated as an input to the deconvolution process. The output of the system can be described by

$$\begin{aligned} \dot{q}(t,p) &= f(q(t,p), u[t], p) \\ y(t,p) &= \frac{q(t,p)}{v[t]} \\ q(0,p) &= q_0(p) \end{aligned} \tag{7.16}$$

where $v[t]$ represents the discretised plasma volume and $u[t]$ the discretised tumourous input function for IgG-Kappa. The non-tumourous synthesis is considered constant during the treatment. The input vector for the deconvolution process is a combination of the volume and tumourous synthesis, both sampled at the output sample periods, each element of which is modified by the optimisation routine at each iteration. The input vector ($u = [v(t_1), ..., v(t_n), u(t_1), ..., u(t_n)]^T$) is split in the numerical integration routine, interpolated to provide a continuous representation, and used in the generation of the system output for the χ^2 estimate.

The results of deconvolving the tumourous synthesis can be seen in Figure 7.7; the patient initially responds to chemotherapy, but then relapses towards the end of treatment. However, if the chemotherapy has no effect on the non-tumourous IgG production the synthesis of the normal IgG should be constant, assuming the patient does not have an unknown immune response. As such there should be a rise, then subsequent fall, due to the change in the fractional catabolic rate in opposition to the change in the tumourous concentration.

The modulation in the non-tumourous IgG-Lambda can be seen in the data shown in Figure 7.8. If the volume is not modified the simulated output does not match the measurements well, with the central peak (at approximately 800 days) being poorly estimated and two troughs at the beginning and the end of treatment not well described by the simulation. However, if the deconvolution is permitted to also modify the volume during optimisation, much better results are obtained. The top graph of Figure 7.9

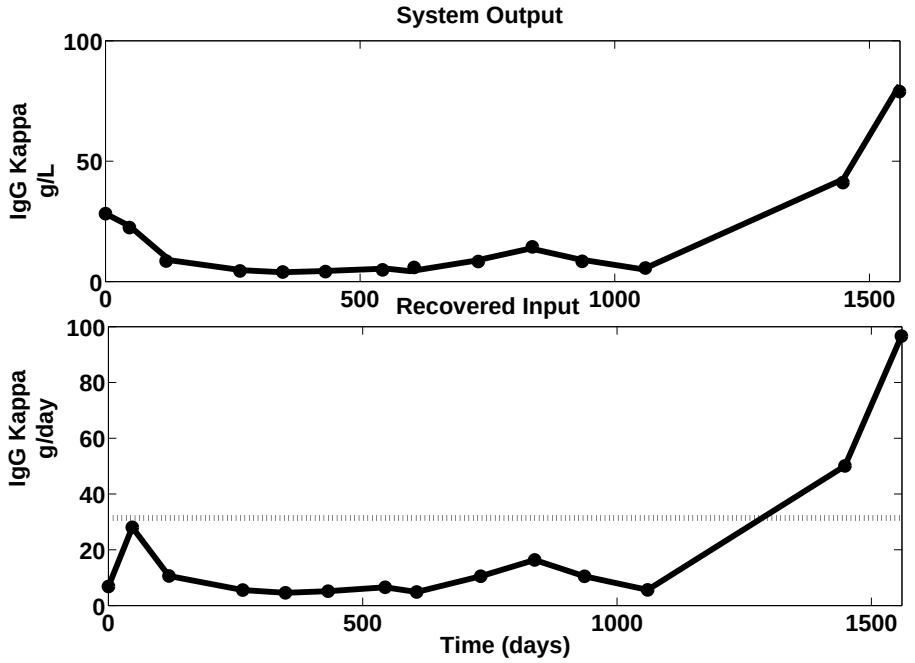


Figure 7.7: *Tumour synthesis (IgG-Kappa). Top (solid - simulated IgG-Kappa concentration in plasma, dashed - IgG-Kappa concentration in EVF, dots - measured IgG-Kappa concentration in plasma). Bottom (solid - deconvolved input, dotted - initial estimate)*

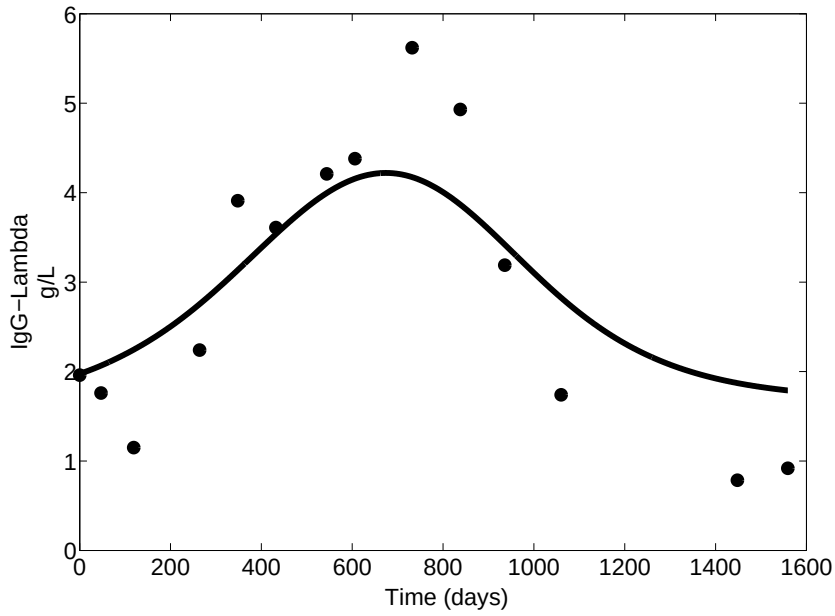


Figure 7.8: *Simulated output of the IgG-Lambda with constant volume (solid - simulated output of non-tumourous IgG-Lambda concentration in plasma, dots - measured IgG-Lambda concentration in plasma.*

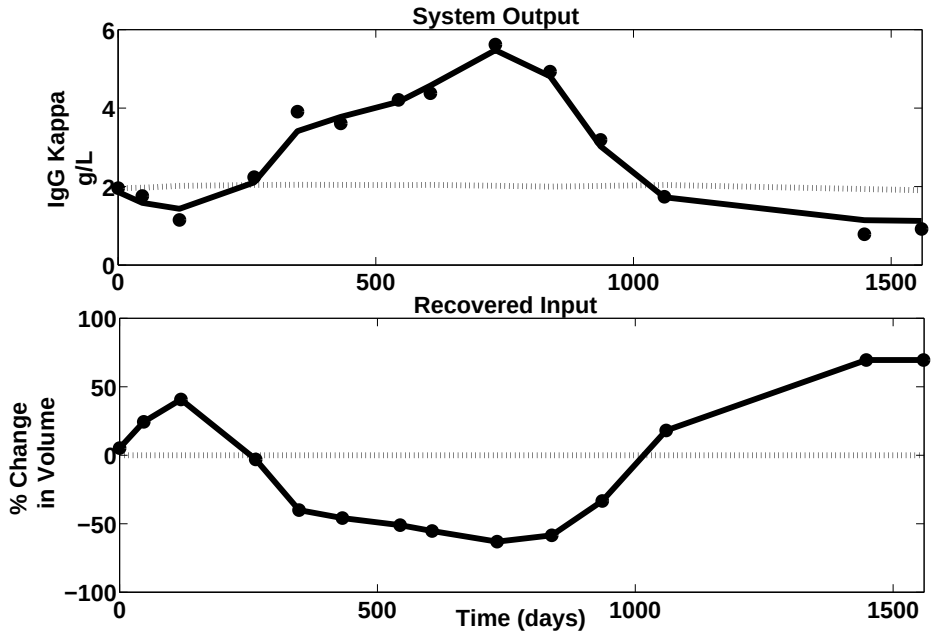


Figure 7.9: *Estimation of volume change due to excess protein viscosity and the effect on the non-tumourous concentration. Top (solid - simulated IgG-Lambda concentration in plasma, dots - measured IgG-Lambda concentration in plasma). Bottom (solid - deconvolved percentage change in volume, dotted - initial estimate)*

shows the simulation results (dashed-line) against those for the measured data (∘).

Whilst the simulation results suggest that the volume change could account for the difference between the observed behaviour and that expected from the model, the results may be misleading. In Tuddenham and Bradley [1974] patients with Macroglobulinaemia with increased concentrations of IgM and IgA certainly exhibit volume increases within the range shown; in patients with MM secreting intact IgG it is less evident. IgG patient volumes in Tuddenham and Bradley [1974] showed an increase in the range of 5-50%. However, all but one patient in the study produced monoclonal IgG-Kappa, the sole patient with IgG-Lambda Myeloma exhibited a 3% increase in volume over that expected.

7.6 Additional applications

Two different applications of deconvolution are described below. Although neither of these methods relate to chemotherapy of tumour synthesis, they do show important

facets of the deconvolution process and highlight the simplicity and flexibility of the methodology used. In addition, these applications also show how, with altering the perspective in which the problem is viewed, it may be recast in a classic input-output relationship applicable for the deconvolution process.

Firstly, in a series of experiments conducted by Guymer et al. [2005] the relationship between the dimensions of a drainage manhole structure and the transport of particulate matter was investigated. Such systems have the capacity to maintain pollutants through fluid storage for extended periods; this ability has implications upon the dispersion of pollutants and the retention of chemicals hazardous to the surrounding environment. In order to analyse such systems the residence distribution time (RDT) is calculated. Previously it was estimated using computational fluid dynamic techniques [Stovin et al., 2008]. These are, however, computationally expensive and time consuming, and prone to convergence errors. The RDT is analogous to the system impulse response and it has been suggested that the response generation could be accomplished through deconvolution [Levenspiel, 1962]. To investigate this, data were provided of upstream and downstream concentration of markers of fluid flowing through a manhole from which the RDT was efficiently estimated. Using Maximum Entropy deconvolution the RDT for a range of drainage structures were calculated simply and accurately, a process which could not previously be done. This work is presented in detail in Stovin et al. [2009].

Secondly, nonlinear compartmental models have been developed that describe the *in vivo* kinetics of insulin and glucose during glucose intolerance tests [Cobelli and Pacini, 1988]. In a collaboration between the University of Warwick and AstraZeneca UK the deconvolution method was used to estimate the insulin secretion in fasting rats, enabling a more improved marker of insulin secretion to be defined (please see Watson et al. [2007] for details).

7.6.1 Confidence estimates

Point estimations, as shown so far, are beneficial when considering trends or overall trajectories of the underlying system behaviour. However, when conducting any form of parameter estimation it is important to have a measure of how confident the experi-

menter is in the results obtained. In statistics this will generally result in identification of the variance for each estimate; in the case of deconvolution, the variance of the deconvolved signal at each point estimate. Two techniques have been used to investigate the variance of the point estimates generated in the deconvolution:

- Approximation of Hessian of the unconstrained optimisation

  If the surface of the cost function is considered, from a geometric perspective, the curvature of the solution around the minimal value relates to the amount of variance in the parameters [Jacquez, 1996, p. 187-190]. If the surface around the minimal value is of high curvature a small change in the parameter values would account for a large change in the cost function, therefore the confidence with which each parameter is defined should be high. If the cost function surface were broad and flat (low curvature) a large change in the parameter may result in very little change in the cost function and hence the confidence in that parameter estimate would be low.

  If it is assumed that the true solution of the input signal were known (u) and the deconvolved signal ($\hat{u}$) had been estimated using the above methods, the variance can be found through the covariance matrix [Fletcher, 1987, p. 110-112]

$$V = var(u) \quad = E((u - \hat{u})(u - \hat{u})^T), \tag{7.17}$$

  where the diagonal elements of V are the variance of elements ($\hat{u}_i$) of the recovered input signal.

  It can be shown [Fletcher, 1987, p. 111] that the covariance matrix is equal to the inverse of the residual function Hessian matrix; further, if the solution is close enough to the 'real' underlying solution, the hessian can be linearised around the estimated values and the Hessian matrix (H) can be approximated by the square of the Jacobian Matrix (J). Therefore V can be approximated by

$$V = H^{-1}\sigma^2 \quad \approx (JJ^T)^{-1}\sigma^2, \tag{7.18}$$

where $J_{ij} = \frac{\partial r_i}{\partial u_j}$ and r_i are the residuals between simulated and measured output evaluated at the sample point i, and σ^2 is the variance of the error. Once again, due to the non-linearity in the solution the Jacobian matrix must be calculated numerically [Press et al., 1992].

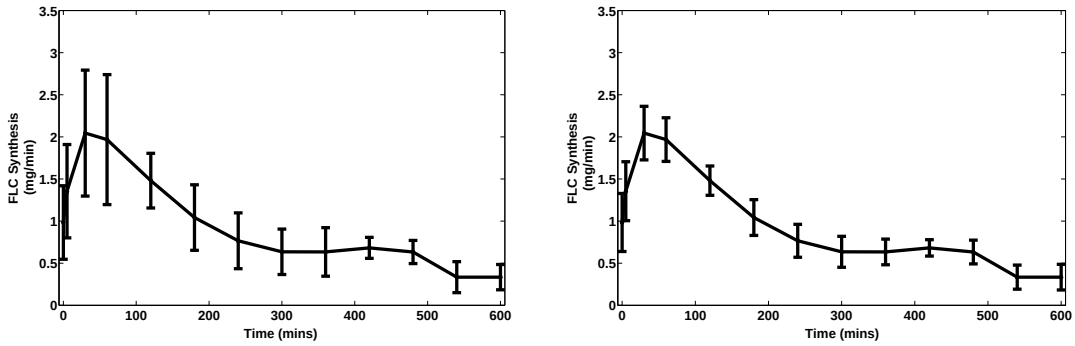
The above analysis assumes that the optimisation is unconstrained, this however is not the case. Several authors have suggested that this results in an over-estimate in the confidence of the variance [De Nicolao and Liberati, 1993; Sparacino et al., 2001]. To accomodate this Monte-Carlo methods are recommended [De Nicolao et al., 1995].

- **Monte-Carlo method**

  A Monte-Carlo technique is a brute-force approach that relies on numerical randomisation and repetition of the simulation to identify the statistical parameters. The techniques used stems from the 'bootstrapping' methods invented by Efron and Tibshirani [1993], where datasets are sampled, or divided, to generate new datasets which can be used in the regularisation procedure to produce an ensemble of input estimates, allowing variance estimates to be generated. Unfortunately, the number of samples available for the biomedical systems model is relatively small, and therefore an alternative method known as residual resampling is used ([Efron and Tibshirani, 1993, p. 111-121] and [Davison and Hinkley, 1997, p. 353-355]). In this method a new data set is generated at each iteration by adding randomly selected residual errors to the simulated results ($\hat{y}$) at the optimal solution ($\hat{u}$), giving an new dataset that can be used to deconvolve the input signal.

An example of estimating the variance of the signal is shown in Figure 7.10, where Figure 7.10(a) shows the variance as estimated by the Jacobian approximation to the Hessian, Figure 7.10(b) the variance using the Monte-Carlo method. Both figures show a single standard deviation. It should be noted that the examples shown here can be considered as a 'best case' and were selected due to the stable behaviour of the optimisation process, and the consistent calculation of the Hessian of the residuals.

Whilst the results seem promising both methods suffer from problems which need to be addressed before the variation in the estimate can be reliably given. Firstly,



(a) Hessian Calculation

(b) Residual Bootstrap Calculation

Figure 7.10: *Variance estimates for deconvolved signal, error bars indicate a single standard deviation from the deconvolved point estimate.*

the bootstrap method, from visual inspection, appears to produce consistent results; however, this method is computationally very expensive. For each iteration a full deconvolution is required, with a minimum of two hundred iterations, preferably more, being suggested by Efron and Tibshirani [1993], which, depending on the machine resource, the number of sample points and the optimisation routine, could be unacceptable.

The Hessian method requires numerical differentiation to be performed, which in the presence of noise is a notoriously unstable process. In addition, if the Hessian matrix is ill-conditioned the inverse can prove unstable. Further research is required in order to successfully recover confidence intervals for the deconvolved signals.

7.7 Summary

In this chapter convolution and deconvolution are presented in the context of biomedical systems analysis. Maximum Entropy regularisation was implemented to allow signals from biomedical processes to be constrained using *a priori* knowledge regarding the processes to be recovered. The regularisation was reformulated as a constrained optimisation problem that can be solved by readily available optimisation routines, without the need for extensive development of novel minimisation methods, as presented by Skilling and Bryan [1984]. As compartmental models are ubiquitous in the field of biomedical modelling, two methods of interfacing the optimisation process to compartmental models were shown, for non-linear and for linear models.

The deconvolution method can successfully be used to provide point estimates for immeasurable quantities in patients. Synthesis of cancerous proteins produced by patients with Multiple Myeloma were deconvolved; for patients with FLC producing tumours different responses can be seen to different regimes. Some patients responded extremely quickly, over hours, during the dialysis treatments. In patients with intact IgG secreting tumours the response to chemotherapy appeared over a more gradual timescale, months rather than hours. However deconvolving the input signal in a non-parametric form permitted a parametric function to be found that was then subsequently fed back to the standard modelling processes (see Chapter 5).

In addition to the original objective of estimating the therapeutic effects of chemotherapy on patients, Maximum Entropy regularisation offered solutions to unexpected problems. Ranging from estimating changes in plasma volume, insulin secretion in response to glucose models, and to non-biomedical problems of estimating the system response in urban drainage systems.

Chapter 8

Conclusions

In this final chapter, the significant outcomes of this thesis will be identified and discussed, both in terms of their contribution to the body of knowledge regarding immune kinetics analysis and the wider field of biomedical systems modelling. The general conclusion to be drawn from the totality of the work presented here is the facility of physiological modelling to directly influence clinical decisions; supporting or discouraging hypotheses that are generated by medical staff during the treatment of patients, through the use of validated compartmental models and mathematical analysis.

As described in the introduction, the aim of the research presented is to assist clinicians in the diagnosis, treatment and development of new treatments for patients with three immune conditions: two types of Multiple Myeloma (FLC and Intact Secreting IgG) and the immune response to organ transplantation. This aim has been successfully met through the creation of several new mathematical models (Chapters 4-6) that describe the kinetics of measurable antibodies relating to the patient's condition. The models generated are, where possible, validated against patient data to ensure their applicability in describing known dynamic responses. They are then used to investigate alternative treatments or assist in the diagnosis of patients currently undergoing treatment. An iterative approach to the model development has been presented starting with linear models to describe the kinetics of FLC under the influence of haemodialysis; these simple models were then extended to include non-linear ligand and receptor binding kinetics to describe the *in vivo* kinetics of whole IgG antibodies. The non-linear receptor models were subsequently used, with minor modification, to

investigate the underlying response of the patient to implanted kidneys.

The models developed enable suggestions to be made regarding the optimal form of apheresis to be used for the clearance of different immune complexes. For Free Light Chains (FLCs) it was shown (Chapter 4) that haemodialysis, with high-flux dialysis membranes, is more effectively able to clear FLC than plasma exchange. Further, through parameter estimation applied to patient data, it was shown that the clearance rate for the two apheresis methods (haemodialysis and plasma exchange) was approximately the same. However, due to the ability to use haemodialysis for a longer duration than plasma exchange greater whole-body removal can be achieved. From the FLC model it was also possible to analyse the effect of haemodialysis in increasing the patient's chance of recovery of renal function under different chemotherapy effectiveness and treatment schedules; suggesting that prolonged and early use of high-flux haemodialysis throughout the chemotherapy treatment has the potential to aid in kidney function recovery. It was also shown that even low-level recovery of the natural renal function had a surprising impact on clearance of monoclonal FLC. In order to predict the long term effect of haemodialysis, a simplified model was presented that utilised the pseudo periodic nature of haemodialysis treatment to yield estimates for minimum and maximum plasma concentrations, without the need for iteratively solving the system equations; however, although periodicity in treatment is hypothetically feasible due to resource constraints (e.g. dialysis equipment, trained staff, etc.) dialysis is often performed a-periodically, thus requiring iterative procedures to be used to solve the system equations. The work on FLC haemodialysis has encouraged the development of a larger follow-on study [Hutchison et al., 2008]; it has also stimulated manufacturers of the membranes to invest resources into further development of the membrane technologies with regards to the FLC clearance.

Comparison of clearance methods is also important when considering kidney transplant patients (Chapter 6) as three apheresis methods are available: plasma exchange, plasmapheresis and plasma absorption. Unfortunately, unlike FLC Myeloma patients the clearance estimates for these methods cannot readily be calculated using standard parameter estimation. A discretisation technique was developed that simultaneously estimates the underlying synthesis and apheresis clearance for a series of transplant patients, from which clearance rates for a range of patients and antibody types (IgM,

IgG and IgA) were estimated. The results of this suggest that IgM is cleared equally well by all methods, whilst IgG and IgA clearance is improved with plasma exchange and plasmapheresis. It should be noted that the data supplied for the clearance estimates showed a bias towards patients undergoing plasmapheresis, due to the facilities available at the University Hospital Coventry and Warwick. Ideally a larger cohort study should be performed, with data from more patients undergoing both plasma exchange and plasma absorption.

For Intact Secreting Myeloma patients an improved tumour marker has been proposed (Chapter 5) that relies on the competitive binding of IgG-Kappa and IgG-Lambda antibodies. A mathematical study has been performed to demonstrate the effectiveness of using combinations of IgG variant markers to optimally determine chemotherapy effects on IgG tumours. It has been shown that in the absence of noise, measuring the ratio of monoclonal tumourous IgG over the non-tumourous IgG will give the most effective measure of tumour synthesis reduction. Further, through simulated results, if the monoclonal IgG cannot be measured independently the ratio of total IgG tumour over non-tumourous IgG will be a better indicator than measuring IgG concentration changes. As resource allocation is always an issue within any clinical environment, it is prudent to rank the markers shown in terms of accuracy of the measurement technique, thus enabling informed decisions to be made regarding which marker to employ. From the analysis presented an appropriate ranking of marker effectiveness would be:

1. Ratio of tumourous IgG over non-tumour IgG.
2. Ratio of total IgG type over non-tumourous IgG.
3. Tumourous IgG only.
4. Total Tumour IgG type inclusive of tumour and non-tumour production.

Based upon the model presented it is anticipated that the ratios discussed could be used either in an ‘on-line’ diagnostic mode during treatment to assess chemotherapy, or post-treatment, to assist in categorising patients to determine future treatment. The ratio under observable events (changes in initial conditions) will offer, at worst, a

comparable estimate of tumour kill as measurement of concentration changes and at best, a significantly more accurate estimate. Once again, a caveat exists that additional clinical validation of the model is required with a greater range of patients and with different responses to ensure the dynamics of the model are sufficient for clinical consideration.

In addition to the clearance of the immune complex, to reduce their impact on the patients, a key feature of treatment is the use of drugs to reduce the production of new proteins. Whilst the long-term effects of the drugs are known, i.e. if the patient responds successfully to the drug this leads to a reduction in the concentration of the immune complex occurs over time; unfortunately, the short-term behaviour and functional descriptions of the direct relationship between drug induction and synthesis dynamics are generally not available. To assist in this, deconvolution methods have proven to be particularly useful (Chapter 7). A simple deconvolution method was used to describe the synthesis of three antibody classes (IgM, IgG and IgA) in patients undergoing immune suppression and apheresis pre- and post-kidney transplants (Chapter 6) . A more complex deconvolution method was used to investigate the effects of chemotherapy on the synthesis of monoclonal antibodies (IgG) and antibody fragments (FLC) in Multiple Myeloma (MM) patients. Through the use of Maximum Entropy deconvolution, a regularisation technique, estimates were provided for several patients. In the case of Intact Secreting MM patients the deconvolved IgG synthesis suggested the use of the Logistic function to describe the dynamics of the immune complex generation rate. For FLC MM patients the short term (hours rather than days) impact of chemotherapy was described, with the effects declining over a twenty four hour period. However, of greater interest was the detection of the tumour recovering from the drug treatment within eight hours; this was discovered even in the presence of apheresis, which, through clearance of FLC, would normally prevent clinicians from detecting such an increase. Deconvolution was also used to suggest, non-parametrically, the form of other time-varying signals, e.g. the relationship between viscosity and plasma volume in MM patients, and the production of insulin in an *in vitro* diabetic study.

Finally, two prototype applications have been suggested that, although having limited validation, show promise in the analysis of patient data. For haemodialysis, an extendable multi-filter model was introduced (Chapter 4) that may allow clinicians

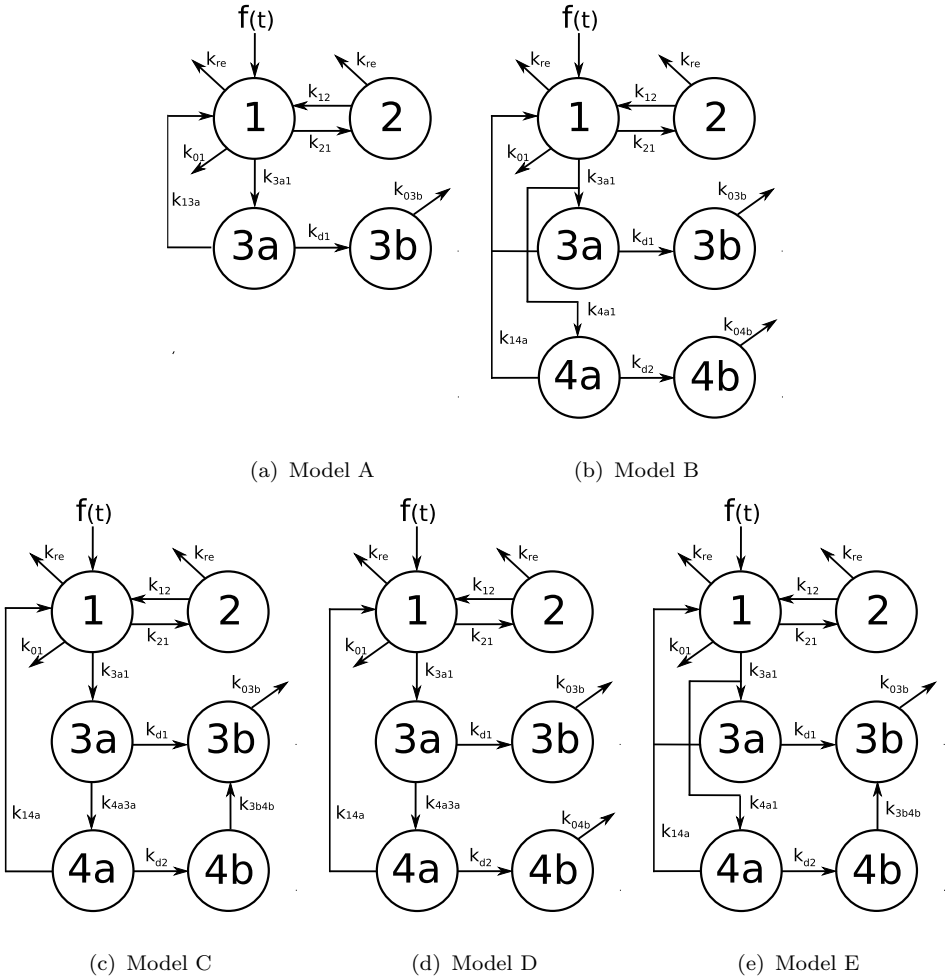
to predict the effects of using multi-filters to reduce increased amounts of FLC in a single session. The model also highlighted the effects of ultrafiltration on middle-molecule clearance. In Chapter 6 cluster analysis was used on longitudinal data of patients undergoing kidney transplants, using dimension reduction techniques (PCA and Sammon Mapping) and partition clustering, patients were categorised into groups, which closely matched the suggested categorisation given by clinical staff. However, both methods require further validation and testing before they can be used for the analysis of real patient data to assist clinical decision making.

8.1 Future work

As with all research, it is possible to define alternative directions that may be taken and further refinement that should be made. The patients studied for the FLC validation (Chapter 4) had no renal function for clearing FLC and were still able to maintain fluid levels in the body. This resulted in a zero ultrafiltration rate for the entire treatment period and a constant volume level during the treatment. However, this is not always true and volume should be introduced as a state-variable, as described in Ward et al. [2006]. Also relating to ultrafiltration, it is known that proteins bind to the surface of a filter membrane during dialysis. This is a process known as membrane fouling [Pedrini et al., 2006]. Fouling negatively affects the membrane's ability to clear material due to restrictive porosity, which becomes worse with both an increase in ultrafiltration and the time that the surface is in contact with the fluid. In order to increase the accuracy of the estimates presented, and determine the impact of membrane fouling on FLC clearance, it might be beneficial to include additional features in the model that would account for this reduction in clearance.

As mentioned in Chapter 4, it is theoretically possible for multiple filters to be connected in a variety of formats, through modification of the tubes or lines connecting the filters. Schematics describing how the extended FLC model could be used to create five different filter configurations can be seen in Figure 8.1. It may be possible using alternative configurations to increase the amount of FLC cleared in a single session. For instance, if Model D (Fig 8.1(d)) is considered, by having the plasma and dialysate sides of the filter connected in parallel an increased concentration gradient across the

inlet of both filters is possible. However, it would be premature to suggest this would result in an increased diffusive transfer across the filter, without further information on the effects this may have on ultrafiltration rates and maximum flow rates in a clinical setting. It would therefore require an *in vitro* validation experiment to be conducted prior to suggesting modification in patient treatment using the extended filter model.



(a) Model A (b) Model B

(c) Model C (d) Model D (e) Model E

Figure 8.1: *Possible extended filter configurations. Model A - single filter. Model B - two filters with plasma and dialysate flows in parallel. Model C - two filter plasma and dialysate flows in series. Model D - two filters, plasma in series and dialysate flow in parallel. Model E - plasma flow in parallel with dialysate in series.*

In addition to clearing the protein of interest, in this case FLC, all apheresis techniques remove elements from the bloodstream that are beneficial to the homeostasis of the patients [Kim et al., 2007], e.g. albumin. If measurements of other particles were taken alongside the FLC plasma concentration it may be possible to investigate the

clearance of these substances in addition to FLC and analyse the correlation between differing protein clearance and patient recovery long term.

The cluster analysis approach of Chapter 6 may offer early indicators for patients that may not be responding successfully to an implanted organ, even though measurements taken throughout treatment indicate otherwise. If a database of patients' data and categorisation is maintained over a period of time, it seems plausible that these data could be used to reference current patients' response to those seen previously, offering an indicator to assist in the prediction of the future response.

The deconvolution process discussed in Chapter 7 allows immeasurable inputs to the system to be analysed. However, two areas of this method which require further investigation are the placement and number of samples for the deconvolved signal. The 'sample points' are defined as the points in the input function around which the interpolation is produced. To determine the synthesis of cancerous cells in MM patients, due to the assumed 'smoothness' of the synthesis function, the choice of sample points was relatively simple in that choosing sample points equal to the time points at which the observation were taken was adequate. When the deconvolution method was applied to alternative experiments, namely, man-hole flow reconstruction and insulin deconvolution, this was inappropriate. In the man-hole problem the output was measured every millisecond for several minutes, resulting in thousands of sample points, using all of these as input sample points is computationally infeasible. In the diabetes experiment spikes were observed at the start of the observation period, which should be matched in the deconvolved input. Both these problems were solved through manual manipulation of the time at which the input signal was sampled. In order to determine the optimal node placement automatically several techniques were investigated including:

1. **Genetic Algorithms** [Goldberg, 1989] with each individual in the population representing a different distribution of sample points.

2. **Nested Optimisation** where both the number and time points are used as parameters for the outer minimisation problem, and magnitude of the input is minimised in the inner loop.

3. **Sensitivity Analysis of the Output to sample point**. D-optimal sampling

techniques [Jacquez, 1996, ch. 16] were used to determine the output sensitivity to changes in the placement of the input samples on the timeline.

Unfortunately, all these methods were computationally expensive with poor convergence and repeatability.

When conducting parameter estimation it is important to have a measure of how confident the experimenter is in the results obtained. In statistics this generally results in identification of the variance for each estimate, or in the case of deconvolution, the variance of the deconvolved signal at each point estimate. Two techniques have been used to estimate the variance of the point estimates generated in the deconvolution, Hessian approximation and Monte Carlo simulations. Both of these require more robust algorithms to be developed.

The models presented in this thesis could be expanded to other immune disorders, e.g. Multiple Sclerosis and Rheumatoid Arthritis, offering similar insight into the underlying dynamics and allowing simulated prediction on patient response to treatment. Due to the exploratory nature of the treatments being undertaken, a limited number of data sets were available for this work. It would be interesting to gather data from a large cohort study and investigate the physical variation within a population in terms of both the *in vivo* parameters (e.g. renal clearance of FLC, plasma and EVF transfer of IgG) in addition to their response to the treatment. For the FLC modelling this may be possible as the use of haemodialysis in treating Multiple Myeloma FLC patients is now being tested in an extended European trial, referred to as EuLite [Hutchison et al., 2008]. Data from these trials were not available during the writing of this thesis.

References

J.J. Abel, L.G. Rowntree, and B.B. Turner. On the removal of diffusible substances from the circulating blood of living animals by dialysis. *Journal of Pharmacology And Experimental Therapeutics*, 5(3):275–316, 1914.

R. Alexanian, S. Salmon, J. Bonnet, E. Gehan, A. Haut, and J. Weick. Combination therapy for multiple myeloma. *Cancer*, 40(6):2765–2771, Dec 1977.

D.G. Altman and J.M. Bland. Missing data. *British Medical Journal*, 334(7590): 334–424, February 2007.

G.R. Ayers and J.C. Dainty. Interative blind deconvolution method and its applications. *Optic Letters*, 13(7):547–555, 1988.

Y. Bard. *Nonlinear Parameter Estimation*. Academic Press, New York and London, 1974.

W.F. Barth, R.D. Wochner, T.A. Waldmann, and J.L. Fahby. Metabolism of human gamma macroglobulins. *Journla of Clinical Investigation*, 43:1036–104, 1964.

R. Bataille. Management of myeloma with bisphosphonates. *New England Journal of Medicine*, 334(8):529–530, Feb 1996.

C. Baudelet and B. Gallez. Cluster analysis of bold fmri time series in tumors to study the heterogeneity of hemodynamic response to treatment. *Magnetic Resonance in Medicine*, 49(6):985–990, 2003.

R. Bellman and K.J. Aström. On structural indentifiability. *Mathematical Bioscience*, 7:329–339, 1970.

G. Bellu, M.P. Saccomani, S. Audoly, and L. D'Angi. Daisy: a new software tool to test global identifiability of biological and physiological systems. *Computer Methods and Programs in Biomedicine*, 88(1):52–61, Oct 2007.

E.A. Bender. *An Introduction to Mathematical Modeling*. Hamilton Printing, 1978.

J.M. Bland and D.G Altman. Statistics notes: Measurement error proportional to the mean. *British Medical Journal*, 313:106–110, 1996.

P.L. Bonate. *Pharmacokinetic-Pharmacodynamic Modeling and Simulation*. Springer, New York, 2006.

A.R. Bradwell, N.D. Evans, M.J. Chappell, P. Cockwell, S.D. Reid, J. Harrison, C. Hutchison, and G.P. Mead. Rapid removal of free light chains from serum by hemodialysis for patients with myeloma kidney. *Blood*, 106(11–13):3482, 2004.

A.R. Bradwell, G.P. Mead, and H.D. Carr-Smith. *Serum Free Light Chain Analysis*. The Binding Site Ltd, The Binding Site Ltd, PO Box 11712, Birmingham, B14 4ZB, UK, 3^{rd} edition, 2005.

F.W.R Brambell, W.A. Hemmings, and I.G. Morris. A theoretical model of gamma-globulin catabolism. *Nature*, 203:1352–1355, 1964.

M. Brandli. Antibody — wikipedia, the free encyclopedia, 2008. URL `http://en.wikipedia.org/w/index.php?title=Antibody`. [Online; accessed 17-December-2008].

I.D. Bross. Mathematical models vs. animal models. *Medical Research Moderization Society*, 1:83–108, 1989.

R.F. Brown. *Biomedical Systems Analysis via Compartmental Concepts*. Abacus, 2^{nd} edition, 1985.

P.M. Campbell and P.F. Halloran. *Solid Organ Transplant Rejection: Mechanisms, Pathology and Diagnosis*, chapter Antibody Mediated Rejection, pages 29–59. Informa Healthcare, 1996.

Cancer-Research(UK). Uk multiple myeloma statistics. 2007. URL http://info.cancerresearchuk.org/cancerstats/types/multiplemyeloma/. [Online; accessed 23-June-2007].

E.R. Carson, C. Cobelli, and L. Finkelstein. *The Mathematical Modeling of Metabolic and Endocrine Systems*. John Wiley and Sons, UK., 1983.

G.C. Cawley and N.L.C. Talbot. *Simple vectorised MATLAB implementation of Sammon's non-linear mapping algorithm*. Computational Biology Lab, University of East Anglia, 2007. URL http://theoval.sys.uea.ac.uk/matlab/. [Online; accessed 1-March-2009].

G. Cevat. Stability analysis of periodically switched linear systems using floquet theory. *Mathematical Problems in Engineering*, 1:1–10, 2004.

M.J. Chapman and K.R. Godfrey. Nonlinear compartmental model indistinguishability. *Automatica*, 32(3):419–422, 1996.

M.J. Chappell, K.R. Godfrey, and S.Vajda. Global identifiability of the parameters of nonlinear systems with specified inputs: A comparison of methods. *Mathematical Bioscience*, 102:41–73, 1990.

M.J. Chappell, N.D. Evans, K.R. Godfrey, and M.J. Chapman. Structural identifiability of controlled state space systems: a quasiautomated methodology for generating identifiable reparameterisations of unidentifiable systems. *IEE Seminar Digests*, 1999 (88):8–18, 1999.

S.C. Chapra and R.P. Canale. *Numerical Methods for Engineers: with Software and Programming Applications*. McGraw-Hill, 4^{th} edition, 2002.

M.K. Charter and S.F. Gull. Maximum entropy and its application to the calculation of drug absorption rates. *Journal of Pharmacokinetics and Pharmacodynamics*, 15 (6):645–655, 1987.

M.K. Charter and S.F. Gull. Maximum entropy and drug absorption. *Journal of Pharmacokinetics and Pharmacodynamics*, 19(5):497–520, 1991.

W.F. Clark, A.K. Stewart, G.A. Rock, M. Sternbach, D.M. Sutton, B.J. Barrett, A.P. Heidenheim, A.X. Garg, and D.N. Churchill. Plasma exchange when myeloma presents as acute renal failure: a randomized, controlled trial. *Annals of Internal Medicine*, 143(11):777–784, Dec 2005.

W.R. Clark, W.L. Macias, B.A. Molitoris, and N-H. L. Wang. Membrane adsorption of bold beta2-microglobulin: Equilibrium and kinetic characterization. *Kidney International*, 46:1140–1146, 1994.

C.M. Close, D.H. Frederick, and J.C. Newell. *Modeling and Analysis of Dynamic System*. Hamilton Printing, New Jersey, 3^{rd} edition, 2002.

C. Cobelli and G. Pacini. Insulin secretion and hepatic extraction in humans by minimal modeling of c-peptide and insulin kinetics. *Diabetes*, 37(2):223–231, Feb 1988.

J. Coresh, D. Byrd-Holt, B.C. Astor, J.P. Briggs, P.W. Eggers, D.A. Lacher, and T.H. Hostetter. Chronic kidney disease awareness, prevalence, and trends among u.s. adults, 1999 to 2000. *Journal of the American Society of Nephrology*, 16(1):180–188, Jan 2005.

T.J. Cornwell and K.F. Evans. A simple maximum entropy deconvolution algorithm. *Astronomy and Astrophysics*, 143:77–83, 1985.

D.J. Cutler. Numerical deconvolution by least squares: use of polynomials to represent the input function. *Journal of Pharmacokinetics and Pharmacodynamics*, 6(3):243–263, Jun 1978.

J.T. Daugirdas. *Physiological Principle in Urea Kinetic Modeling*, chapter 3, pages 25–58. Oxford Science Publications, London, 4^{th} edition, 2006.

J.T. Daugirdas, T. Greene, T.A. Depner, F.A. Gotch, and R.A. Star. Relationship between apparent (single-pool) and true (double-pool) urea distribution volume. *Kidney International*, 56(5):1928–1933, Nov 1999.

H. Davies. *Introductory immunobiology*. Chapman Hall, London, 2^{nd} edition, 1997.

A.C. Davison and D.V. Hinkley. *Bootstrap Methods and their Applications*. Chapman Hall, 1997.

G. De Nicolao and D. Liberati. Linear and nonlinear techniques for the deconvolution of hormone time-series. *IEEE Transactions on Biomedical Engineering*, 40(5):440–455, May 1993.

G. De Nicolao, D. Liberati, and A. Sartorio. Deconvolution of infrequently sampled data for the estimation of growth hormone secretion. *IEEE Transactions on Biomedical Engineering*, 42(7):678–687, July 1995.

S. Diop and M. Fliess. Nonlinear observability, identifiability, and persistent trajectories. *Proceedings of the 30th Conference on Decision and Control*, 1:714–719, Dec 1991.

P.O. Droz. Quantification of biological variability. *Annals of Occupational Hygiene*, 36 (3):295–306, Jun 1992.

B. Efron and R.J. Tibshirani. *An introduction to the Bootstrap*. Chapman Hall, Florida, 1993.

S. Eloot, J-Y De Vos, F. De Vos, R. Hombrouckx, and P. Verdonck. Middle molecole removal in low-flow polysulfone dialyzers: Impact of flows surface area on whole-body and dialyzer clearances. *Hemodialysis International*, 9:399–408, 2005.

H.W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Springer, London, 1996.

N.D. Evans, M.J. Chapman, M.J. Chappell, and K.R. Godfrey. Identifiability of uncontrolled nonlinear rational systems. *Automatica*, 38(10):1799–1805, 2002.

N.D. Evans, M.J. Chappell, M.J. Chapman, and K.R. Godfrey. Structural indistinguishability between uncontrolled (autonomous) nonlinear analytic systems. *Automatica*, 40(11):1947–1953, 2004.

N.D. Evans, J. Hattersley, C. Hutchison, Y. Hu, K.R. Godfrey, A.R. Bradwell, G.P. Mead, and M.J. Chappell. Modelling of haemodialysis in limiting serum free light chains in patients with renal failure. In J. Zaytoon D. Feng, O. Dubois and E. Carson, editors, *Proceedings of the 6th IFAC Symposium on Modelling and Control in Biomedical Systems*, pages 75–80. Elsevier, Oxford, September 2006.

B.S. Everitt, S.Landau, and M.Leese. *Cluster analysis*. Arnold, London, 4^{th} edition, 2001.

F.D.A. Challenge and opportunity on the critical path to new medical products. Technical report, U.S. Department of Health and Human Services, 2004. URL `http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html`. [Online; accessed 2-January-2008].

G.Z. Ferl, A.M. WU, and J.J. Distefano III. A predictive model of therapeutic monoclonal antibody dynamics and regulation by the neonatal fc receptor (fcrn). *Annals of Biomedical Engineering*, 33(11):1640–1652, 2005.

R. Fielding. *Emergency Consequences of Acute Renal Failure*. Acute Renal Failure in Practice. Imperial College Press, 2002.

R. Fletcher. *Practical Methods of Optimization*. Wiley, London, 2^{nd} edition, 1987.

J.C. Foreman and T. Johansen. *Textbook of Receptor Pharmacology*. CRC Press, London, 2^{nd} edition, 2002.

U. Forys and A. Marciniak-Czochra. Logistic equations in tumour growth modelling. *International Journal of Applied Mathematics and Computer Science*, 13(3):317–325, 2003.

R.L. Fournier. *Basic Transport Phenomena in Biomedical Engineering*. Taylor and Francis, London, 2^{nd} edition, 2007.

M. Galach, A. Ciechanowska, S. Sabalinska, J. Waniewski, J. Wojcicki, and A. Werynski. Impact of convective transport on dialyzer clearance. *Journal of Artificial Organs*, 6:42–48, 2003.

D.A.G. Galton and F. Brito-Babapulle. *Myelomatosis*, chapter 37, pages 737–752. Treatment of Cancer. Chapman and Hall Medical, 2^{nd} edition, 1990.

Gambro Dialysatoren GmbH. Company Website, 2009. URL `http://www.gambro.com/`. [Online; accessed 17-December-2008].

A.J.T. George and C.E. Urch. *Diagnostic and Therapeutic Antibodies*. Humana Press, New Jersey, 2000.

V. Ghetie and E.S. Ward. Trancytosis and catabolism of antibody. *Immunological Research*, 25(2):97–113, 2002.

K.R. Godfrey. *Compartmental Models and their Application*. Academic Press, New York, 1983.

K.R. Godfrey and M.J. Chapman. The problem of model indistinguishability in pharmacokinetics. *Journal of Pharmacokinetics and Pharmacodynamics*, 17(2):229–267, Apr 1989.

D.E. Goldberg. *Genetic Algorithms in search optimization and machine learning*. Addison Wesley Longman, 1989.

F. Gotch, N. Levin, G. Zasuwa, and J. Tayeb. Kinetics of beta-2-microglobulin in hemodialysis. *Contributions to nephrology*, 74:132–138, 1989.

F.A. Gotch. Evolution of the single-pool urea kinetic model. *Seminars in Dialysis*, 14 (4):252–256, 2001.

F.A. Gotch and M.A. Keen. *Kinetic Modeling in Hemodialysis*. Clinical Dialysis. McGraw-Hill Professional, New York, 2005.

B.M. Gurbaxani and S.L. Morrison. Development of new models for the analysis of fc-fcrn interactions. *Molecular Immunology*, 43:1379–1389, 2006.

I. Guymer, P. Dennis, R. O'Brien, and C. Saiyudthong. Diameter and surcharge effects on solute transport across surcharged manholes. *Journal of Hydraulic Engineering*, 131(4):312–321, 2005.

P.F. Halloran. Immunosuppressive drugs for kidney transplantation. *New England Journal of Medicine*, 351(26):2715–2729, Dec 2004.

P.C. Hansen. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1998. ISBN 0-89871-403-6.

P.C. Hansen. Regularization tools version 4.0 for matlab 7.3. *Numerical Algorithms*, 46:189–194, October 2007.

J.G. Hattersley, N.D. Evans, M.J. Chappell, G. Mead, C. Hutchison, and A.R. Bradwell. Nonparametric prediction of free-light chain generation in multiple myeloma patients. In Myung Jin Chung and P. Misra, editors, *Proceedings of the 17th World Congress The International Federation of Automatic Control Seoul, Korea, July 6-11, 2008*, 2008.

K.M. Heal, M.L. Hansen, and K.M. Rickard. *Maple V Learning Guide (Release 5).* Springer: Berlin, New York., 1998.

A.C. Hearn. *REDUCE Users Manual 3.6.* RAND, Santa Monica, 1995.

W.L. Henrich. *Principles and Practice of Dialysis.* Lippincott Williams & Wilkins, Philidelphia, 2004.

I. Herzum, H. Renz, and H.G. Wahl. Immunochemical quantification of free light chains in urine. *Clinical Chemistry*, 51(6):645–655, 2005.

R. Higgins. Antibody incompatible kidney transplantation. In *Conference Proceeding CIN 2007 - 4th Congress of Nephrology In Internet.* Society of Nephrology, 2007.

T.J. Holmes. Blind deconvolution of quantum-limited incoherent imagery: maximum-likelihood approach. *Journal of the Optical Society of America*, 9(7):1052–1570, 1992.

R. Hovorka, M.J. Chappell K. R. Godfrey, F.N. Madden, M.K. Rouse, and P.A. Soons. Code: a deconvolution program implementing a regularization method of deconvolution constrained to non-negative values. description and pilot evaluation. *Biopharmaceutics & Drug Disposition*, 19(1):39–53, 1998.

E. Hubert. Differential algebra for derivations with nontrivial commutation rules. *Journal of Pure and Applied Algebra*, 200(1-2):163–190, 2005.

C.A. Hutchison, P. Cockwell, S. Reid, K. Chandler, G.P. Mead, J. Harrison, J. Hattersley, N.D. Evans, M.J. Chappell, M. Cook, H. Goehl, M. Storr, and A.R. Bradwell.

Efficient removal of immunoglobulin free light chains by hemodialysis for multiple myeloma: In vitro and in vivo studies. *Journal of the American Society of Nephrology*, 18:886–895, 2007.

C.A Hutchison, M. Cook, N. Heyne, K. Weisel, L. Billingham, A.R. Bradwell, and P. Cockwell. European trial of free light chain removal by extended haemodialysis in cast nephropathy (eulite): A randomised control trial. *Trials*, 9:55–56, 2008.

K. Iga, Y. Ogawa, T. Yashiki, and T. Shimamoto. Estimation of drug absorption rates using a deconvolution method with nonequal sampling times. *Journal of Pharmacokinetics and Pharmacodynamics*, 14(2):213–225, Apr 1986.

J.A. Jacquez. *Compartmental Analysis in Biology and Medicine*. Biomedware Press, Ann Arbor, 3^{rd} edition, 1996.

C.A. Janeway, P. Travers, M. Walport, and J.D. Capra. *Immunobiology: the immune system in health and disease*. Elsevier Science, 4^{th} edition, 1999.

S.C. Jordan, A. Vo, S. Bunnapradist, M. Toyoda, A. Peng, D. Puliyanda, B. Kamil, and D. Tyan. Intravenous immune globulin treatment inhibits crossmatch positivity and allows for successful transplantation of incompatible organs in living-donor and cadaver recipients. *Transplantation*, 76(4):631–636, 2003.

R.P. Junghans and C.L. Anderson. The protection receptor for igg catabolism is the beta2-microglobulin-containing neonatal intestinal transport receptor. *Proceedings of the National Academy of Science of the United States of America*, 93(11):5512–5516, 1996.

A.S. Karkach. Trajectories and models of individual growth. *Demographic Research*, 15(12):347–400, 2006.

T. Kato, N. Kubo, H. Shimizu, and P.M. Mineshima. Clinical evaluation of dual-dialyzer hemodialysis (ddhd). *Journal of Artificial Organs*, pages 42–46, 2007.

O. Kedem and A. Katchalsky. Thermodynamic analysis of the permeability of biological membranes to nonelectrolytes. *Biochimica et biophysica Acta*, 27:229–246, 1958.

J. Kim, W.L. Hayton, J.M. Robinson, and C.L. Anderson. Kinetics of fcrn-mediated recycling of igg and albumin in human: Pathophysiology and therapeutic implications using a simplified mechanism-based model. *Clinical Immunology*, 122:146–155, 2007.

D.J. King. *Applications And Engineering Of Monoclonal Antibodies*. CRC, Philidelphia, 1998.

W.J. Kolff. The artificial kidney and its effect on the development of other artificial organs. *Nature Medicine*, 8:1063–1065, 2002.

V. Krishna. *Textbook of Pathology*. Orient Longman, London, 2004.

V.A. Kumar, J.Y. Yeun, T.A. Depner, and B.R. Don. Extended daily dialysis vs. continuous hemodialysis for icu patients with acute renal failure: A two-year single center report. *International journal of artificial organs*, 27(5):371–379, 2004.

N. Leung, M.A. Gertz, S.R. Zeldenrust, S.V. Rajkumar, A. Dispenzieri, F.C. Fervenza1, S. Kumar, M.Q. Lacy, J.A. Lust, P.R. Greipp, T.E. Witzig, S.R. Hayman, S.J. Russell, R.A. Kyle, and J.L. Winters. Improvement of cast nephropathy with plasma exchange depends on the diagnosis and on reduction of serum free light chains. *Kidney International*, 73:1282–1288, 2008.

O. Levenspiel. *Chemical Reaction Engineering; an Introduction to the Design of Chemical Reactors*. Wiley, New York, 1962.

D.G. Levitt. Pkquest: a general physiologically based pharmacokinetic model. introduction and application to propranolol. *BMC Clinical Pharmacology*, 2002.

D.G. Levitt. Physiologically based pharmacokinetic modeling of arterial - antecubital vein concentration difference. *BMC Clinical Pharmacology*, 4:2–10, Feb 2004.

J.K. Leypoldt. Kinetics of b2-microglobulin and phosphate during hemodialysis: Effects of treatment frequency and duration. *Seminars in Dialysis*, 18(5):401–408, 2005.

J.K. Leypoldt, A.K. Cheung, and R.B. Deeter . Single compartment models for evaluating beta 2-microglobulin clearance during hemodialysis. *American Society for Artificial Internal Organs*, 43(6):904–909, 1997.

J.K. Leypoldt, B.L. Jaber, M.J. Lysaght, J.T. McCarthy, and J. Moran. Kinetics and dosing predictions for daily haemofiltration. *Nephrology Dialysis Transplant*, 18: 769–776, 2003.

D. Liberzon and A.S. Morse. Basic problems in stability and design of switched systems. *IEEE Control Systems Magazine*, 19(5):59–70, 1999.

I. Lipkovich, J. Houston, and J. Ahl. Identifying patterns in treatment response profiles in acute bipolar mania: a cluster analysis approach. *BMC Psychiatry*, 8(1):1–8, 2008.

F.N. Madden, K.R. Godfrey, M.J. Chappell, R. Hovorka, and R.A. Bates. A comparison of six deconvolution techniques. *Journal of Pharmacokinetics and Pharmacodynamics*, 24(4):283–299, 1995.

E.A. Maharaj. Cluster of time series. *Journal of Classification*, 17:297–314, 2000.

M.A. Malesker and L.E. Morrow. *Fluids and Electrolytes*, chapter 24, pages 403–417. Pharmacotherapy principles & practice. McGraw-Hill Professional, 2007.

M.J. Manning and R.J. Turner. *Comparative Immunobiology.* Blackie, Glasgow, 2^{nd} edition, 1976.

G. Margaria, E. Riccomagno, M.J. Chappell, and H.P. Wynn. Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences. *Mathematical Bioscience*, 174(1):1–26, Nov 2001.

G. Mariani and W. Strober. *Immunoglobulin Metabolism*, chapter 8, pages 74–93. Fc Receptors and the Action of Antibodies. American Society of Microbiology, 1990.

The Mathworks. Matlab optimization toolbox user's guide. URL `http://www.mathworks.co.uk/access/helpdesk/help/optim/`. [Online; accessed 12-Sept-2009].

M.C.P.A. Facsimile. Company Website, 2009. URL `http://www.mcpa-software.com`.

G. Mead, C. Beardsmore, S. Reid, J. Hattersley, P. Moss, G. Pratt, A. Macwhannal, A. Jacob, S. Handa, C. Craddock, M. Cook, and S. Basu. Kinetics of tumour kill during induction chemotherapy for multiple myeloma using frequent free light chain measurements. In M. Cazzola and R. Fo, editors, *Proceedings of the 11th Congress of the European Hematology Association Amsterdam, the Netherlands.*, pages 440–441. Ferrata Storti Foundation, Pavia, Italy, June 15 - 18 2006.

H. Mehlhorn and P.M. Armstrong. *Encyclopedic Reference of Parasitology: Diseases, Treatment, Therapy*. Springer, New York, 2001.

H. Metzer. *Fc Receptors and the Action of Antibodies*. American Society of Microbiology, Massacheusetts, 1990.

C.D. Meyer. *Matrix Analysis and Applied Linear Algebra: Solutions Manual*. SIAM, Philadelphia, 2000.

M. Mongeau, H. Karsenty, V. Rouz, and J.B. Hiriart-Urruty. Comparison of public-domain software for black box global optimization. *Optimization Methods and Software*, 13(3):203–226, 2000.

R.A. Montgomery, A.A. Zachary, L.C. Racusen, M.S. Leffell, K.B. King, J. Burdick, W.R. Maley, and L.B. Ratner. Plasmapheresis and intravenous immune globulin provides effective rescue therapy for refractory humoral rejection and allows kidneys to be successfully transplanted into cross-match-positive recipients. *Tranplantation*, 70(6):887–95, 2000.

Y. Mrabet. Hemodialysis-en.svg, January 2008. URL `http://en.wikipedia.org/wiki/File:Hemodialysis-en.svg`. Image reproduced under GNU and Creative Commons License. [Online: accessed 24-March 2009].

A.R. Nissenson and R.N. Fine. *Clinical dialysis*. McGraw-Hill Professional, 2005.

G. Oster. Berkley madonna. Company Website, 2009. URL `http://www.berkeleymadonna.com`. [Online: accessed 4-April 2008].

M.J. Owen and J.R. Lamb. *Immune recognition*. Oxford University Press, Oxford, 1988.

P.V. Pedersen. Novel deconvolution method for linear pharmacokinetic systems with polyexponential impulse response. *Journal of Pharmaceutical Sciences*, 69(3):312–318, Mar 1980.

L.A. Pedrini, G. Cozzi, P. Faranna, A. Mercieri, P. Ruggiero, S. Zerbi, A. Feliciani, and A. Riva. Transmembrane pressure modulation in high-volume mixed hemodiafiltration to optimize efficiency and minimize protein loss. *Kidney International*, 69 (3):573–579, 2006.

T. Petitclerc. Estimation of mass transfer through a hemodialyzer: Theoretical approach and clinical applications. *Artificial Organs*, 22(7):601–607, 1998.

V.N. Phat and S. Pairote. Global stabilization of linear periodically time-varying switched systems via matrix inequalities. *Journal of Control Theory and Applications*, 4(1):26–31, 2006.

H. Pohjanpalo. Systems identifiability based on the power series of the solution. *Mathematical Bioscience*, 41:21–33, 1978.

M.J.D. Powell. Direct search algorithms for optimization calculations. *Acta Numerica*, 7(1):287–336, 1998.

K.M. Powers, M.J. Wilkowski, A.W. Helmandollar, K.G. Koenig, and W.K. Bolton. Improved urea reduction ratio and kt/v in large hemodialysis patients using two dialyzers in parallel. *American Journal of Kidney Diseases*, 35(2):266–274, 2000.

C. Pozzi. Prognostic factors and effectiveness of treatment in acute renal failure due to multiple myeloma: a review of 50 cases. report of the italien renal immunopathology group. *Clinical Nephrology*, 28(1):1–9, 1987.

W.H. Press, S.A. Teukolsky W.T. Vetterling, and B.P. Flannery. *Numerical Recipes*. Cambridge University Press, New York, 2^{nd} edition, 1992.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. URL `http://www.R-project.org`. ISBN 3-900051-07-0.

H. Ratschek and J. Rokne. *New Computer Methods fo Global Optimization*. Ellis Horwood Ltd, Chichester, UK, 1988.

F.J. Richards. A flexible growth function for empirical use. *Journal of Experimental Botany*, 10(39):290–300, 1959.

W.A. Ritschel and G.L. Kearns. *Handbook of Basic Pharmacokinetics*. American Pharmacist Assoc., New York, 6^{th} edition, 2004.

W.A. Ritshel and G.L. Kearns. *Handbook of Basic Pharmacokinetics including Clinical Applications*. American Pharmacist Association, New York, 2004.

J.F. Ritt. *Differential Algebra*. American Mathematical Society, New York, 1950.

G.N. Rogentine, D.S. Rowe, J. Bradley, T.A. Waldmann, and J.L. Fahey. Metabolism of human immunoglobulin d (igd). *Journal of Clinical Investigation*, 45(9):1467–1478, 1966.

S. Rohatagi, N.E. Martin, and J.S. Barrett. Pharmacokinetic/pharmacodynamic modeling in drug development. In K. Krishna, editor, *Applications of Pharmacokinetics Principles in Drug Development*, part 8, pages 333–372. Klumer Academic Press, Boston, 3^{rd} edition, 2004.

M. Saccomani, S. Audoly, G. Bellu, and L. D'Angio. A new differential algebra algorithm to test identifiability of nonlinear systems with given initial conditions. *IEEE Conference on Decision and Control*, 4:3108–3113, 2001.

M.P. Saccomani, L. D'Angio, S. Audoly, and C. Cobelli. *A Priori Indentifiability of Physiological Parameteric Models*, chapter 4, pages 77–106. Modelling Methodology for Physiology and Medicine. Academic Press, London, 2002.

J.W. Sammon. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, C-18(5):401–409, 1969.

J.A. Sargent and F.A. Gotch. *Principles and Biophysics of Dialysis*, chapter 2, pages 35–102. Replacement of renal function by dialysis. Springer, New York, 4^{th} edition, 1996.

C.T. Shaw and G.P. King. Using cluster analysis to classify time series. *Physica D: Nonlinear Phenomena*, 58(1-4):288–298, 1992.

J. Skilling and R.K. Bryan. Maximum entropy image reconstruction: general algorithm. *Monthly Notices of the Royal Astronomical Society*, 211:111–124, 1984.

L.A. Sklar, J. Sayre, V.M. McNeil, and D.A. Finney. Competitive binding kinetics in ligand-receptor-competitor systems: Rate parameters for unlabeled ligands for the formyl peptide receptor. *Molecular Pharmacology*, 28:323–330, 1985.

G. Sparacino, G. De Nicolao, and C. Cobelli. *Deconvolution*, chapter 3, pages 107–151. Modelling Methodology for Physiology and Medicine. Academic Press, London, 2001.

D.M. Spiegel, P.L. Baker, S. Babcock, R. Contiguglia, and M. Klein. Hemodialysis urea rebound: the effect of increasing dialysis efficiency. *American Journal of Kidney Disease*, 25(1):26–29, 1995.

F. Spirovskia, K. Stojanoskia, and A. Mitrevskib. Cluster analysis method in electrophoretic patterns classification. *Magnetic Resonance in Medicine*, 2(1):26–31, 2005.

V. Stovin, L. Shing-Tak, and I. Guymer. A scaling methodology to characterize the mixing effects of surcharged manholes. *Submitted to Journal of Hydraulic Engineering*, 2008.

V. Stovin, I. Guymer, M. Chappell, and J. Hattersley. The use of deconvolution techniques to identify the fundamental mixing characteristics of urban drainage structures. In *Proceedings of the 8th International Conference on Urban Drainage Modelling*, 2009.

W. Strober, R.D. Wochner, M.H. Barlow, D.E. McFarlin, and T.A. Waldmann. Immunoglobulin metabolism in ataxia telangiectasia. *Journal of Clinical Investigation*, 47(8):1905–1915, 1968.

P.W. Sullivan and S.E. Salmon. Kinetics of tumor growth and regression in igg multiple myeloma. *Journal of Clinical Investigation*, 51:1697–1708, 1972.

K. Takahashi, K. Saito, S. Takahara, A. Okuyama, K. Tanabe, H. Toma, K. Uchida, A. Hasegawa, N. Yoshimura, and Y. Kamiryo. Excellent long-term outcome of abo-incompatible living donor kidney transplantation in japan. *American Journal of Transplantation*, 4 (7):1089–1096, 2004.

J. Tattersall. *Adequacy of Dialysis*. Oxford Textbook of Clinical Nephrology. Oxford University Press, Oxford, 2005.

R. Tomovic and M. Vukobratovic. *General Sensitivity Theory*. Elsevier North-Holland, New York, 1972.

Toray. Filtryzer bk-f series: Hollow fiber dialyzer. Product documentation, Toray Industries, Inc, Tokyo, 2006.

Toray Industries Inc. Company Website, 2009. URL `http://www.toray.com/`. [Online; accessed 17-December-2008].

E.G.D Tuddenham and J. Bradley. Plasma volume expansion and increased serum viscosity in myeloma and macroglobulinaemia. *Clinical and Experimental Immunology*, 16:169–176, 1974.

S. Vajda, P. Valko, and T. Turnyi. Principal component analysis of kinetic models. *International Journal of Chemical Kinetics*, 17(1):55–81, 1985.

R. Vanholder, M. Burgelman, R. De Smet, P. Voogeleere, and S. Ringoir. Two-pool versus single-pool models in the determination of urea kinetic parameters. *Blood Purification*, 14(6):437–450, 1996.

T.A. Waldman and W. Strober. Metabolism of immunoglobulins. *Progress in Allergy*, 13:1–110, 1969.

T.A. Waldmann and W.D. Terry. Familial hypercatabolic hypoproteinemia - a disorder of endogenous catabolism of albumin and immunoglobulin. *Journal of Clinical Investigation*, 86:2093–2098, 1990.

T.A. Waldmann, A. Iio, M. Ogawa, O.R. McIntyre, and W. Strober. The metabolism of ige: Studies in normal individuals and in a patient with ige myeloma. *The Journal of Immunology*, 117:1139–1144, 1976.

E. Walter. *Indentifiability of State Space Models*. Berlin:Spring-Verlag, 1982.

G.G. Walter and M. Contreras. *Compartmental Modeling with Networks*. Birkhauser, Boston, 1999.

J. Waniewski. Linear approximations for the description of solute flux through permselective membranes. *Joumal of Membrane Science*, 95:179–184, 1994.

R.A. Ward, T. Greene, B. Hartmann, and W. Samtleben. Resistance to intercompartmental mass transfer limits b2-microglobulin removal by post-dilution hemodiafiltration. *Kidney International*, 69(8):1431–1437, 2006.

E. Watson, M.J. Chapell, S.M. Poucher, F. Ducrozet, R. Macdonald, A. Yu, and J. Hattersley. Deconvolution of c-peptide to estimate insulin secretion and clearance in han wistar rats. In *Proceedings of the 2007 BioMedMath 2008: Stochastic Differential Equation Models with Applications to the Insulin-Glucose System and Neuronal Modelling*, Middelfart, Denmark, August 2007.

M.C. Williams. *Mass Transfer*, chapter 4, pages 171–201. Introduction to Bioengineering. Oxford University Press, Oxford, 1996.

J. F Winchester, J.A. Salsberg, and N.W. Levin. Beta-2 microglobulin in esrd: an in-depth review. *Advances in Renal Replacement Therapy*, 10(4):279–309, Oct 2003.

A. Wismüller, O. Lange, D.R. Dersch, G.L. Leinsinger, K. Hahn, B. Pütz, and D. Auer. Cluster analysis of biomedical image time-series. *International Journal of Computer Vision*, 46(2):103–128, 2002.

S. Wolfram. *The Mathematica Book*. Wolfram Media/Cambridge University Press, Cambridge, UK, 1999.

P. Zucchelli, S. Pasquali, L. Cagnoli, and G. Ferrari. Controlled plasma exchange trial in acute renal failure due to multiple myeloma. *Kidney International*, 33(6): 1175–1180, Jun 1988.

Appendix A

Identifiability analysis

The identifability analysis for each of the models presented in this thesis can be found below. Each section contains models pertaining to a single chapter.

A.1 Chapter 4

A.1.1 Warwick FLC model

For completeness both the Mathematica [Wolfram, 1999] and Daisy [Bellu et al., 2007] code are presented.

On/Off dialysis model analysis using Taylor Series in Mathematica

The system equations for an 'on' and 'off' dialysis FLC model

$$
\begin{aligned}
&\text{q1}'[\text{t\_}]:= -(\text{kre}+\text{k12})\text{q1}[t] + \left(\tfrac{\text{v1}}{\text{v2}}\text{k12}\right)\text{q2}[t] + \text{f},\\
&\text{q2}'[\text{t\_}]:=\text{k12q1}[t] - \left(\text{kre} + \left(\tfrac{\text{v1}}{\text{v2}}\text{k12}\right)\right)\text{q2}[t],\\
&\text{q1d}'[\text{t\_}]:= -(\text{kd}+\text{kre}+\text{k12})\text{q1d}[t] + \left(\tfrac{\text{v1}}{\text{v2}}\text{k12}\right)\text{q2d}[t] + \text{f},\\
&\text{q2d}'[\text{t\_}]:=\text{k12q1d}[t] - \left(\text{kre} + \left(\tfrac{\text{v1}}{\text{v2}}\text{k12}\right)\right)\text{q2d}[t],\\
&\text{yd}[\text{t\_}]:=\text{q1d}[t]/\text{v1},\\
&\text{yd}[\text{t\_}]:=\text{q1}[t]/\text{v1};
\end{aligned}
$$

the initial conditions,

$\mathrm{q1}[0] = \mathrm{q10}; \mathrm{q2}[0] = \mathrm{q20}; \mathrm{q1d}[0] = \mathrm{q1d0}; \mathrm{q2d}[0] = \mathrm{q2d0}.$

Defining the alternative parameterisation set

$\mathrm{subst} = \{\mathrm{q10} \to \mathrm{q10b}, \mathrm{q1d0} \to \mathrm{q1d0b}, f \to \mathrm{fb},$
$\mathrm{k12} \to \mathrm{k12b}, \mathrm{q2d0} \to \mathrm{q2d0b}, \mathrm{q20} \to \mathrm{q20b},$
$\mathrm{kd} \to \mathrm{kdb}, \mathrm{v1} \to \mathrm{v1b}\}.$

Create the Taylor series coefficients

$\mathrm{a0} = \{y[0], \mathrm{yd}[0]\};$
$\mathrm{y1} = D[y[t], t]; \mathrm{y1d} = D[\mathrm{yd}[t], t]; \mathrm{a1} = (\{\mathrm{y1}, \mathrm{y1d}\}/.t \to 0);$
$\mathrm{y2} = D[\mathrm{y1}, t]; \mathrm{y2d} = D[\mathrm{y1d}, t]; \mathrm{a2} = (\{\mathrm{y2}, \mathrm{y2d}\}/.t \to 0);$
$\mathrm{y3} = D[\mathrm{y2}, t]; \mathrm{y3d} = D[\mathrm{y2d}, t]; \mathrm{a3} = (\{\mathrm{y3}, \mathrm{y3d}\}/.t \to 0);$
$\mathrm{y4} = D[\mathrm{y3}, t]; \mathrm{y4d} = D[\mathrm{y3d}, t]; \mathrm{a4} = (\{\mathrm{y3}, \mathrm{y3d}\}/.t \to 0);$

Solve the equations for the parameterisation, using the first and second Taylor series co-efficients

$\mathrm{eqn0} = \mathrm{a0} - (\mathrm{a0}/.\mathrm{subst});$
$\mathrm{eqn1} = \mathrm{a1} - (\mathrm{a1}/.\mathrm{subst});$
$\mathrm{eqns} = \{\mathrm{eqn0}, \mathrm{eqn1}\};$
$\mathrm{newSoln} = \mathrm{Simplify}[\mathrm{Solve}[\mathrm{eqns} == 0, \{\mathrm{q10b}, \mathrm{q10db}, \mathrm{fb}, \mathrm{k12b}, \mathrm{q2d0b}, \mathrm{q20b}, \mathrm{kdb}, \mathrm{v1b}\}]]$

$\mathrm{newSoln}[[1]] =$

$\mathrm{kdb} \to \frac{\mathrm{kdq1d0v2+k12(q20v1-q2d0v1-q10v2+q1d0v2)}}{\mathrm{q1d0v2}}, \mathrm{fb} \to \frac{\mathrm{q1d0b(k12q20v1+fv2-k12q10v2)}}{\mathrm{q1d0v2}},$
$\mathrm{q10b} \to \frac{\mathrm{q10q1d0b}}{\mathrm{q1d0}}, \mathrm{k12b} \to 0, \mathrm{v1b} \to \frac{\mathrm{q1d0bv1}}{\mathrm{q1d0}}$

$\mathrm{newSoln}[[2]] =$

$\text{kdb} \to \text{k12} + \text{kd} - \frac{f}{\text{q1d0}} + \frac{\text{fb}}{\text{q1d0b}} + \text{k12b}\left(-1 + \frac{\text{q2d0bv1}}{\text{q1d0v2}}\right) - \frac{\text{k12q2d0v1}}{\text{q1d0v2}}$,
$\text{q20b} \to \frac{\text{k12q1d0bq20v1}-\text{fbq1d0v2}+\text{fq1d0bv2}-\text{k12q10q1d0bv2}+\text{k12bq10q1d0bv2}}{\text{k12bq1d0bv1}}$,
$\text{q10b} \to \frac{\text{q10q1d0b}}{\text{q1d0}}, \text{v1b} \to \frac{\text{q1d0bv1}}{\text{q1d0}}$

k12b cannot be zero, we therefore reject the first solution

soln = newSoln[[2]];

The following two solutions are separated for computational issues. If this was not done first the 'Solve' would not return a value but consume all available computer resources (e.g. memory and CPU).

eqn = (a2 − (a2/.subst))/.soln;

newSoln = Simplify[Solve[eqn == 0, {q10db, fb, q2d0b}]];

soln = Simplify[soln/.newSoln[[1]]];
soln = Union[soln, newSoln[[1]]];

eqn = (a3 − (a3/.subst))/.soln;

newSoln = Simplify[Solve[eqn == 0, {q1d0b, k12b}]]

{{q1d0b → q1d0, k12b → k12}}

The only solution is $p = \bar{p}$; therefore,the system is SGI.

soln = Simplify[soln/.newSoln[[1]]];
soln = Union[soln, newSoln[[1]]]

{fb → f, k12b → k12, kdb → kd, q10b → q10, q1d0b → q1d0,

q20b $\rightarrow$ q20, q2d0b $\rightarrow$ q2d0, v1b $\rightarrow$ v1}

On/Off dialysis model using Differential Algebra in Daisy

The daisy input file for the on/off dialysis model.

```
WRITE "Two Compartment Warwick Model - On + Off dialysis."
% B_ IS THE VARIABLE VECTOR
B_:={y1,y1d,q1,q2,q1d,q2d}
FOR EACH EL_ IN B_ DO DEPEND EL_,T
%B1_ IS THE UNKNOWN PARAMETER VECTOR
B1_:={q10,q10d,q20,q20d,k12,kd,P,v1}
LET kre=beta,v2=alpha
%NUMBER OF STATES
NX_:=4
%NUMBER OF OUTPUTS
NY_:=2
%MODEL  EQUATIONS
C_ := { df(q1,t) = -(kre+k12)*q1 + k12*(v1/v2)*q2 + P,
df(q2,t) = k12*q1-(k12*(v1/v2)+kre)*q2,
df(q1d,t) = -(kre+k12+kd)*q1d + k12*(v1/v2)*q2d + P,
df(q2d,t) = k12*q1d-(k12*(v1/v2)+kre)*q2d,
q1*v1=y1,
q1d*v1=y1d}
SEED_:=150
DAISY()
% INITIAL CONDITIONS
IC_:={q1=q10,q2=q20,q1d=q10d,q2d=q20d}
CONDINIZ()
```

Below is the resulting output using the above input and Daisy 1.4 [Bellu et al., 2007] under Reduce 3.8 [Hearn, 1995]. Only the characteristic set, exhaustive summary and identifiability results are shown.

```
CHARACTERISTIC SET
aa_(1) := df(y1,t,2)*v2 + df(y1,t)*(k12*v1 + k12*v2 + 2*kre*v2)
+ y1*kre*(k12*v1 + k12*v2 + kre*v2)
- p*v1*(k12*v1 + kre*v2)
aa_(2) := df(y1d,t,2)*v2
+ df(y1d,t)*(k12*v1 + k12*v2 + kd*v2 + 2*kre*v2)
+ y1d*(k12*kd*v1
+ k12*kre*v1 + k12*kre*v2 + kd*kre*v2 + kre**2*v2)
- p*v1*(k12*v1 + kre*v2)
aa_(3) := q1*v1 - y1
aa_(4) := df(y1,t)*v2 - q2*k12*v1**2
+ y1*v2*(k12 + kre) - p*v1*v2
aa_(5) := q1d*v1 - y1d
aa_(6) := df(y1d,t)*v2 - q2d*k12*v1**2
+ y1d*v2*(k12 + kd + kre) - p*v1*v2

RANDOMLY CHOSEN NUMERICAL PARAMETER VECTOR

b2_ := {q10=120,q10d=25,q20=117,q20d=83,k12=38,kd=130,p=55,v2=135}

THE ALGEBRAICALLY OBSERVABLE STATE(S)
flist1_ := {(135*k12*v1 + 135*k12*v2 - 38*v1*v2 - 5130*v2)/(135*v2),
(kre*(135*k12*v1 + 135*k12*v2 - 38*v1*v2 - 5130*v2))/(135*v2),
(v1*( - 27*k12*p*v1 - 27*kre*p*v2 + 1485*kre*v2 + 418*v1*v2))/(27*v2),
(v1*( - 27*k12*p*v1 - 27*kre*p*v2 + 1485*kre*v2 + 418*v1*v2))/(27*v2),
(135*k12*v1 + 135*k12*v2 + 135*kd*v2 - 38*v1*v2 - 22680*v2)/(135*v2),
(135*k12*kd*v1 + 135*k12*kre*v1 + 135*k12*kre*v2 + 135*kd*kre*v2
 - 38*kre*v1*v2 - 22680*kre*v2 - 4940*v1*v2)/(135*v2),
```

```
v1*(q10 - 120),
(v1*( - 15*k12*q20*v1 + 1800*k12*v2 - 15*p*v2
 + 494*v1*v2 - 67575*v2))/15,
v1*(q10d - 25),
(v1*( - 135*k12*q20d*v1 + 3375*k12*v2 + 3375*kd*v2
- 135*p*v2 + 3154*v1*v2 - 559575*v2))/135}

MODEL PARAMETER SOLUTION(S)
gi_ := {{k12=38,kd=130,p=55,q10=120,q10d=25,q20=117,q20d=83,v2=135}}

SYSTEM GLOBALLY IDENTIFIABLE
```

On dialysis model using Differential Algebra in Daisy

The daisy input file for the dialysis model when measurements are only taken when the patient is on dialysis.

```
WRITE "Two Compartment Warwick Model - On dialysis only."
% B_ IS THE VARIABLE VECTOR
B_:={y1,q1,q2}
FOR EACH EL_ IN B_ DO DEPEND EL_,T
%B1_ IS THE UNKNOWN PARAMETER VECTOR
B1_:={q10,k12,kd,P}
let kre=alpha,q20=beta,v1=gamma,v2=delta
%NUMBER OF STATES
NX_:=2
%NUMBER OF OUTPUTS
NY_:=1
%MODEL  EQUATIONS
C_ := { df(q1,t) = -(kre+k12+kd)*q1 + k12*(v1/v2)*q2 + P,
df(q2,t) = k12*q1-(k12*(v1/v2)+kre)*q2,
y1=q1/v1}
SEED_:=150
```

```
DAISY()
% INITIAL CONDITIONS
IC_:={q1=q10,q2=q20}
CONDINIZ()
```

The resulting output showing a globally identifiable parameter set.

```
CHARACTERISTIC SET

aa_(1) :=  - df(y1,t,2)*delta*gamma + df(y1,t)*gamma*( - 2*alpha*delta
   - delta*k12 - delta*kd - gamma*k12) - y1*gamma*(alpha**2*delta
   + alpha*delta*k12 + alpha*delta*kd + alpha*gamma*k12
   + gamma*k12*kd) + p*(alpha*delta + gamma*k12)
...
EXHAUSTIVE SUMMARY INCLUDING INITIAL CONDITION(S) OF
THE ALGEBRAICALLY OBSERVABLE STATE(S)

flist1_ := {alpha*delta*p - 83*alpha*delta + gamma*k12*p
    - 2075*gamma, - delta*k12 - delta*kd + 142*delta
    - gamma*k12 + 25*gamma, - alpha*delta*k12
    - alpha*delta*kd + 142*alpha*delta - alpha*gamma*k12
    + 25*alpha*gamma -gamma*k12*kd + 2925*gamma,-q10
    + 120,- beta*gamma*k12 + 25*beta*gamma
    + 120*delta*k12 + 120*delta*kd - delta*p
    - 16957*delta}
MODEL PARAMETER SOLUTION(S)
gi_ := {{k12=25,kd=117,p=83,q10=120}}
SYSTEM GLOBALLY IDENTIFIABLE
```

Two filter model without ultrafiltration

Input commands to determine the identifiability of the two filter model.

```
% B_ IS THE VARIABLE VECTOR
```

```
B_:={y1,y2,y3,y4,q1,q2,q3a,q3b,q4a,q4b}
FOR EACH EL_ IN B_ DO DEPEND EL_,T
%B1_ IS THE UNKNOWN PARAMETER VECTOR
B1_:={q10,v1,v3a,v3b,v4a,v4b,k12,k21,k1e,P,
      kpout,kre,kd1,kd2,kdout1,kpret1,kdout2,kpret2}
let q20=a, q30=b, q40=c
%NUMBER OF STATES
NX_:=6
%NUMBER OF OUTPUTS
NY_:=4
%MODEL  EQUATIONS
C_ := { df(q1,t)  = -(kre + k12 + k1e + kpout)*q1 + k21*q2
  + kpret2*q4a +  P,
df(q2,t)  = k12*q1-(k21+kre)*q2,
df(q3a,t) = -kd1*(q3a*v3a-q3b*v3b) - kpret1*q3a
  + kpout*q1,
df(q4a,t) = -kd2*(q4a*v4a-q4b*v4b) - kpret2*q4a + kpret1*q3a,
df(q3b,t) =  kd1*(q3a*v3a-q3b*v3b) - kdout1*q3b + kdout2*q4b,
df(q4b,t) =  kd2*(q4a*v4a-q4b*v4b) - kdout2*q4b,
y1=q1*v1,
y2=q3a*v3a*2 - q1*v1,
y3=q4a*v4a*2 - q3a*v3a*2 - q1*v1,
y4=q3b*v3b*2 - 2*q4b*v4b}
SEED_:=150
DAISY()
% INITIAL CONDITIONS
IC_:={q1=q10,q2=q20, q3a=q30, q3b=0,  q4a=q40, q4b=0}
CONDINIZ()
```

The resulting output showing the global identifiability is

```
...
THE SYSTEM IS ALGEBRAICALLY OBSERVABLE
```

```
...
MODEL PARAMETER SOLUTION(S)
gi_ := {{k12=39,k1e=73,k21=146,kd1=124,kd2=34,kdout1=111,
 kdout2=115,kpout=4,kpret1=147,kpret2=81,kre=75,
 p=17,q10=48,v1=42,v3a=92,v3b=58,v4a=131,v4b=69}}
SYSTEM GLOBALLY IDENTIFIABLE
```

Two filter model with ultrafiltration

```
WRITE "2 Filter FLC Model with UF included (2FilterUF.txt)."
% B_ IS THE VARIABLE VECTOR
B_:={y1,y2,y3,y4,q1,q2,q3a,q3b,q4a,q4b}
FOR EACH EL_ IN B_ DO DEPEND EL_,T
%B1_ IS THE UNKNOWN PARAMETER VECTOR
B1_:={v1,k12,k21,P,kre,kd,UF1,UF2}
let kpout=d,kdout1=e,kpret1=f,kdout2=g,kpret2=h,v3a=i,v3b=j,v4a=k,v4b=l
%NUMBER OF STATES
NX_:=6
%NUMBER OF OUTPUTS
NY_:=4
%MODEL  EQUATIONS
C_ := { df(q1,t)  = -(kre + k12 + kpout)*q1 + k21*q2 + kpret2*q4a +  P,
df(q2,t)  = k12*q1-(k21+kre)*q2,
df(q3a,t) = -kd*(q3a*v3a-q3b*v3b) - kpret1*q3a + kpout*q1
  - UF1*(q3a*v3a + q3b*v3b),
df(q3b,t) =  kd*(q3a*v3a-q3b*v3b) - kdout1*q3b + kdout2*q4b
  + UF1*(q3a*v3a + q3b*v3b),
df(q4a,t) = -kd*(q4a*v4a-q4b*v4b) - kpret2*q4a + kpret1*q3a
  - UF2*(q4a*v4a + q4b*v4b),
df(q4b,t) =  kd*(q4a*v4a-q4b*v4b) - kdout2*q4b
  + UF2*(q4a*v4a + q4b*v4b),
y1=q1*v1,
```

```
y2=2*q3a*v3a - y1,
y3=2*q4a*v4a - y2,
y4=2*q3b*v3b-2*q4b*v4b}

SEED_:=100
DAISY()
```

The resulting output showing local identifiability is

```
...
THE SYSTEM IS ALGEBRAICALLY OBSERVABLE
...
MODEL PARAMETER SOLUTION(S)
gi_ := {{k12=75,k21=67,kd=30,kre=88,p=83,uf1=5,uf2=35,v1=20},
{k12=-67,k21=-75,kd=30,kre=88,p=83,uf1=5,uf2=35,v1=20}}
SYSTEM LOCALLY IDENTIFIABLE
```

However, the second parameter set contains negative values for the rate constants k_{12} and k_{21} that are, by definition, positive. Therefore as the parameter estimatation process is restricted to positive values a globally identifiable solution is available.

A.2 Chapter 5

Identifiability analysis relating to IgG and FcRn recycling.

A.2.1 Hypoproteinemia patient

Define the system equations of model for patient with Hypoproteinemia,

$$\text{IgGT1}'[\text{t\_}]:= -(\text{k01}+\text{k21})\text{IgGT1}[t]+\text{k12IgGT2}[t],$$
$$\text{IgGT2}'[\text{t\_}]:=\text{k21IgGT1}[t]-\text{k12IgGT2}[t],$$

and specify the observations, plasma concentration and total tracer,

y1[t_]:=IgGT1[t],

y2[t_]:=IgGT1[t] + IgGT2[t].

Set the initial conditions, tracer is administered as a bolus injection into plasma,

IgGT1[0] = 1

IgGT2[0] = 0.

Calculate the Taylor series co-efficients,

y11 = D[y1[t], t];

y12 = D[y11, t];

y21 = D[y2[t], t];

and define the alternate parameter vector and Taylor series coefficients,

subst = {k01 → k01b, k12 → k12b, k21 → k21b},

coeffs = {y11, y12, y21}/.t → 0,

coeffsPbar = coeffs/.subst,

eqns = coeffs − coeffsPbar,

soln = Solve[eqns == 0, {k01b, k12b, k21b}].

The system is not locally identifiable as there is only a single solution

Length[soln]

1

and all the parameters are globally identifiable

Simplify[soln[[1]]] {k12b → k12, k01b → k01, k21b → k21}.

Appendix B

Analytical asymptotic prediction

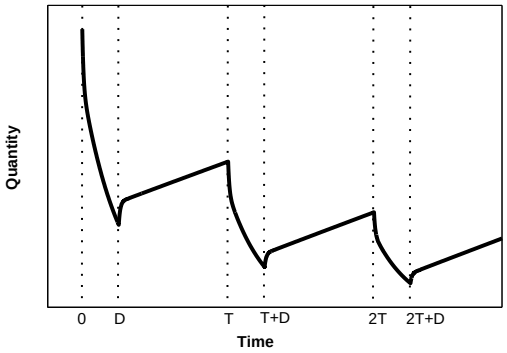


Figure B.1: *Example of apheresis treatment.*

Proposition B.0.1. *Given definition 4.23, the asymptotic minimum and maximum values of a linear single-switch*[1] *state-space system with constant input are given by*

$$\begin{aligned} \lim_{t\to\infty} \max c(t) &= (I_d - e^{A_1\tau_1}e^{A_2\tau_2})^{-1}(e^{A_2\tau_2}\alpha + \beta) \\ \lim_{t\to\infty} \min c(t) &= (I_d - e^{A_2\tau_2}e^{A_1\tau_1})^{-1}(\alpha + \beta e^{A_1\tau_1}) \end{aligned} \tag{B.1}$$

where e^{Ai} is the matrix exponential, $A_1 \in \mathbb{R}^{d\times d}$ and $A_2 \in \mathbb{R}^{d\times d}$ are the two possible system matrices, τ_1 is the duration system one is operational, τ_2 system matrix two, I_d is identity matrix of dimension d; with α and β defined as

$$\alpha(\tau_1) := \int_0^{\tau_1} e^{A_1(\tau_1 - s)}\, P_\infty\, ds \tag{B.2}$$

[1] A system with only two possible system matrices

$$\beta(\tau_1) := \int_0^{\tau_2} e^{A_2(\tau_2 - s)} P_\infty \, ds \tag{B.3}$$

where P_∞ is a vector of the asymptotic baseline of the input functions; for FLC patients this is $P_\infty = [f_{min}, 0]^T$.

Proof. As each system matrix is linear, an analytical solution is available for each switching phase, using the matrix exponential solution the quantities of the material in each compartment can be expressed as

$$q(t) = e^{A_i t} q_0 + \int_0^t e^{A_i(t-s)} P_\infty \, ds, \quad t \in [\tau_i, \tau_{i+1}] \tag{B.4}$$

where A_i is the matrix for switching phase i and q_0 is vector of starting conditions, either the initial conditions if $0 > t > D$ or terminal conditions from the previous switching phase if $t > D$. With reference to Figure B.1, taking $t = 0$ as the start of the first apheresis treatment, the quantity in the system after the treatment (with duration D) can be calculated as

$$q(D) = e^{A_1 D} q_0 + \int_0^D e^{A_1(D-s)} P_\infty \, ds. \tag{B.5}$$

The patient is then left off apheresis for a fixed period, after which the FLC quantity can then be found from

$$q(T) = e^{A_2(T-D)} q(D) + \int_0^{T-D} e^{A_2((T-D)-s)} P_\infty \, ds \tag{B.6}$$

where T is the time between the first and second apheresis treatments. If the treatment is then repeated with period T, using a recursive process (often referred to as Picard's iterative process [Jacquez, 1996, pp. 81–85]) by substituting equation (B.5) into equation (B.6) results in

$$\begin{aligned} q(T) &= e^{A_2(T-D)}(e^{A_1 D} q_0 + \int_0^D e^{A_1(D-s)} P_\infty \, ds) \\ &\quad + \int_0^{T-D} e^{A_2((T-D)-s)} P_\infty \, ds. \\ &= e^{A_2(T-D)} e^{A_1 D} q_0 + e^{A_2(T-D)} \int_0^D e^{A_1(D-s)} P_\infty \, ds \\ &\quad + \int_0^{T-D} e^{A_2((T-D)-s)} P_\infty \, ds. \end{aligned} \tag{B.7}$$

It should be noted that in general the on and off apheresis matrices do not commute, $A_2A_1 \neq A_1A_2$, therefore the matrix exponential can not be manipulated as in the scalar case. Repeating this process for two more treatment cycles yields

$$\begin{aligned} q(T+D) &= e^{A_1D}q(T-D) + \int_0^D e^{A_1(D-s)}P_\infty\, ds \\ &= e^{A_1D}e^{A_2(T-D)}e^{A_1D}q_0 + e^{A_1D}e^{A_2(T-D)}\int_0^D e^{A_1(D-s)}P_\infty\, ds \\ &+ e^{A_1D}\int_0^{T-D} e^{A_2((T-D)-s)}P_\infty\, ds. \\ &+ \int_0^D e^{A_1(D-s)}P_\infty\, ds \end{aligned} \tag{B.8}$$

$$\begin{aligned} q(2T) &= e^{A_2(T-D)}q(T+D) + \int_0^{T-D} e^{A_2((T-D)-s)}P_\infty\, ds \\ &= (e^{A_2(T-D)}e^{A_1D})^2q_0 \\ &+ e^{A_2(T-D)}e^{A_1D}e^{A_2(T-D)}\int_0^D e^{A_1(D-s)}P_\infty\, ds \\ &+ e^{A_2(T-D)}e^{A_1D}\int_0^{T-D} e^{A_2((T-D)-s)}P_\infty\, ds. \\ &+ e^{A_2(T-D)}\int_0^D e^{A_1(D-s)}P_\infty\, ds \\ &+ \int_0^{T-D} e^{A_2((T-D)-s)}P_\infty\, ds \end{aligned} \tag{B.9}$$

Noting that the contribution of $q(0) \to 0$ as $t \to \infty$, repeating this process infinity yields a recursive formula for the on and off dialysis points. Collecting the term $e^{A_2(T-D)}\int_0^D e^{A_1(D-s)}P_\infty\, ds + \int_0^{T-D} e^{A_2((T-D)-s)}P_\infty\, ds$ from equation (B.9) yields a peak, off dialysis, value of

$$\lim_{t\to\infty} \max q(t) = (I_n + e^{A_1\tau_1}e^{A_2\tau_2} + (e^{A_1\tau_1}e^{A_2\tau_2})^2 + \ldots)(e^{A_2\tau_2}\alpha + \beta). \tag{B.10}$$

Whilst collecting $\int_0^D e^{A_1(D-s)}P_\infty\, ds + e^{A_1(D)}\int_0^{T-D} e^{A_2((T-D)-s)}P_\infty\, ds$ from the recursion of equation (B.8), gives the following formula for the minimum, on dialysis, measurement

$$\lim_{t\to\infty} \min q(t) = (I_n + e^{A_2\tau_2}e^{A_1\tau_1} + (e^{A_2\tau_2}e^{A_1\tau_1})^2 + \ldots)(\alpha + \beta e^{A_1\tau_1}). \tag{B.11}$$

In both equation (B.10) and (B.11), α and β are as defined by equations (B.2) and (B.3). Using the Neumann Series [Meyer, 2000], if A is a square matrix, such that $||A|| < 1$ then

$$(I_n + A)^{-1} = I_n + A + A^2 + ... = \sum_{m=0}^{\infty} A^m \tag{B.12}$$

where A is matrix in $R^{n\times n}$, I_n is a d dimensional identity matrix. Using this identity equations (B.9) can be succinctly expressed as

$$\lim_{t\to\infty} \max q(t) = (I_n - e^{A_1\tau_1}e^{A_2\tau_2})^{-1}(e^{A_2\tau_2}\alpha + \beta) \tag{B.13}$$

$$\lim_{t\to\infty} \min q(t) = (I_n - e^{A_2\tau_2}e^{A_1\tau_1})^{-1}(\alpha + \beta e^{A_1\tau_1}) \tag{B.14}$$

If attention is now turned to the effect of the initial conditions on the early phase of treatment. Comparing coefficients of the initial conditions

$$\begin{aligned} q(0) &= 1 \\ q(T) &= e^{A_2(T-D)}e^{A_1D} \\ q(2T) &= (e^{A_2(T-D)}e^{A_1D})^2 \\ &\dots \\ q(nT) &= (e^{A_2(T-D)}e^{A_1D})^n \end{aligned}$$

from which it is evident that the initial conditions are decaying by an exponential function. It seems therefore feasible to approximate the discrete values of the maximum, pre dialysis, measurements by a simple function

$$q_{env}[n] = (e^{A_2(T-D)}e^{A_1D})^n(q(0) - \gamma) + \gamma \tag{B.15}$$

$$\tag{B.16}$$

where $n \in \mathbb{Z}^+$ is the number of complete dialysis treatments performed (on dialysis plus interdialytic rest period) and γ is the maximum asymptotic value as defined in equation (B.1). It should be noted that equation (B.15) is only an approximation to

the peak values and as such has inherent error in the calculation. This occurs due to the use of the asymptotic values of $q(t)$ used as an approximation to the effects of the input integral on the overall FLC levels. Comparing equation (B.1) with the integral terms in equation (B.7), it can be seen the asymptotic maximum contains higher order terms that are not present at this stage in the recursion. Therefore as n, or t, increase the error will be reduced. An analytical bound on the error has not yet been found, however, in the examples presented it is significantly smaller ($O(10^{-3})$) than the numerical error in ODE solvers used. To assist in visual representation, and comparison of treatments, equation (B.15) can be converted into a continuous function

$$q_{env}(t) \quad = (e^{A_2(T-D)}e^{A_1 D})^{\frac{t}{T}}(q(0)-\gamma)+\gamma; \tag{B.17}$$

however, this function should only be interpreted where analytical at time $t = nT$.

□

Appendix C

Calculation of initial conditions for the FcRn model

If IgG-Kappa, IgG-Lambda and the tumourous IgG are measured, and the system is in steady-state prior to treatment, the system can be described by the following equations:

$$-\left(\beta - \frac{\gamma}{\zeta + (K_1(0) + L_1(0) + T_1(0))}\right) K_1(0) + P_K = 0 \quad \text{(C.1a)}$$

$$-\left(\beta - \frac{\gamma}{\zeta + (K_1(0) + L_1(0) + T_1(0))}\right) L_1(0) + P_L = 0 \quad \text{(C.1b)}$$

$$-\left(\beta - \frac{\gamma}{\zeta + (K_1(0) + L_1(0) + T_1(0))}\right) T_1(0) + P_T(0) = 0 \quad \text{(C.1c)}$$

For the patient simulations all parameters are known with the exception of $P_T(0)$, $K_1(0)$ and $L_1(0)$. Equations (C.1a) can be re-arranged to yield three quadratic Equations, each with two solutions of opposing sign:

$$P_T(0) = \frac{T_1(0)\ \alpha\ \delta}{\zeta} \quad \text{(C.2a)}$$

$$K_1(0) = \frac{P_K\ \delta}{2P_L(P_K+P_L)\alpha} \quad \text{(C.2b)}$$

$$L_1(0) = \frac{\delta}{2(P_K+P_L)\alpha} \quad \text{(C.2c)}$$

where

$$\begin{aligned}\delta = {} & P_K P_L + P_L^2 + P_L T_1(0)\alpha - P_L\beta + P_L\alpha\gamma \\ & \pm\sqrt{P_L^2\left(4(P_K+P_L)\alpha(T_1(0)+\gamma) + (P_K + P_L - T_1(0)\alpha + \beta - \alpha\gamma)^2\right)}\end{aligned}$$

and

$$\begin{aligned}\zeta &= P_K P_L + P_L^2 + P_L T_1(0)\alpha + P_L\beta + P_L\alpha\gamma \\ &\quad + \sqrt{P_L^2\left(4(P_K+P_L)\alpha(T_1(0)+\gamma) + (P_K+P_L-T_1(0)\alpha+\beta-\alpha\gamma)^2\right)}.\end{aligned}$$

However, *in vivo* IgG concentration and synthesis must be positive, therefore any negative solution can ignored. As all parameters are positive, for Equations (C.2a) and (C.2c) to be negative δ must be negative.

$$\begin{aligned}0 &> P_K P_L + P_L^2 + P_L T_1(0)\alpha - P_L\beta + P_L\alpha\gamma \\ &\quad - \sqrt{P_L^2\left(4(P_K+P_L)\alpha(T_1(0)+\gamma) + (P_K+P_L-T_1(0)\alpha+\beta-\alpha\gamma)^2\right)}\end{aligned}$$

which reduces to

$$P_L^2(P_K+P_L)\alpha(T+\alpha) > 0 \tag{C.3}$$

As all parameters are positive, this must be true, yielding a single positive solution

$$\begin{aligned}\delta &= P_K P_L + P_L^2 + P_L T_1(0)\alpha - P_L\beta + P_L\alpha\gamma \\ &\quad + \sqrt{P_L^2\left(4(P_K+P_L)\alpha(T_1(0)+\gamma) + (P_K+P_L-T_1(0)\alpha+\beta-\alpha\gamma)^2\right)}\end{aligned}$$

Appendix D

Equation to support transplant immune response

The following equations describe the synthesis ($P[n]$) and quantity of Ig in EVF ($q2[n]$) for the IgG and IgA models.

$$P[n] = \frac{\lambda_1 \lambda_2 \alpha}{\beta} \tag{D.1}$$

where

$$\begin{aligned} \alpha = \quad & e^{t\lambda_2}\left((q_1[n-1]+q_2[n-1])k_{21} - q_1[n-1]\lambda_1\right) + e^{t(\lambda 1+\lambda_2)} q_1[n]\left(\lambda_1-\lambda_2\right) + \\ & e^{t\lambda_1}\left(q_1[n-1]\lambda_2 - (q_1[n-1]+q_2[n-1])k_{21}\right) \\ \beta = \quad & e^{t(\lambda_1+\lambda_2)} k_{21}\left(\lambda_1-\lambda_2\right) + e^{t\lambda_2}\left(k_{21}-\lambda_1\right)\lambda_2 - e^{t\lambda_1}\lambda_1\left(k_{21}-\lambda_2\right) \end{aligned}$$

$$q2[n] = \quad \frac{e^{-t(\lambda_1+\lambda_2)} - e^{t\lambda_2}(A+B) + \frac{k_{12}C}{D}}{\lambda_1-\lambda_2} \tag{D.2}$$

where

$$
\begin{aligned}
A := \quad & (q_1[n-1] + q_2[n-1])k_{12} + q_2[n-1]\,(k_{1e} + k_p - \lambda_1) \\
B := \quad & e^{t\lambda_1}\left((q_1[n-1] + q_2[n-1])k_{12} + q_2[n-1]\,(k_{1e} + k_p - \lambda_2)\right) \\
C := \quad & \left(-e^{t\lambda_1}\lambda_1 + e^{t(\lambda_1+\lambda_2)}\,(\lambda_1 - \lambda_2) + e^{t\lambda_2}\lambda_2\right) E \\
D := \quad & -e^{t\lambda_1}\lambda_1\,(k_{21} - \lambda_2) + e^{t(\lambda_1+\lambda_2)}k_{21}\,(\lambda_1 - \lambda_2) + e^{t\lambda_2}\,(k_{21} - \lambda_1)\,\lambda_2 \\
E := \quad & e^{t\lambda_2}\left((q_1[n-1] + q_2[n-1])k_{21} - q_1[n-1]\lambda_1\right) + \\
& e^{t(\lambda_1+\lambda_2)}q_1[n]\,(\lambda_1 - \lambda_2) + \\
& e^{t\lambda_1}\left(q_1[n-1]\lambda_2 - (q_1[n-1] + q_2[n-1])k_{21}\right)
\end{aligned}
$$

λ_1 and λ_2 are the eigenvalues of characteristic polynomial of the system matrix defined by

$$
A = \begin{pmatrix} -(k_{01} + k_p + k_{21}) & k_{12} \\ k_{21} & -k_{21} \end{pmatrix}.
$$

Appendix E

Validation of Maximum Entropy method

The results shown in Figure E.1 are reproductions of the Madden et al. [1995] equations which are used to test the Maximum Entropy procedure. The system was based on a zero initial condition linear system with a bi-exponential impulse response; the impulse and state-space representation can be seen in equations E.3 and E.2.

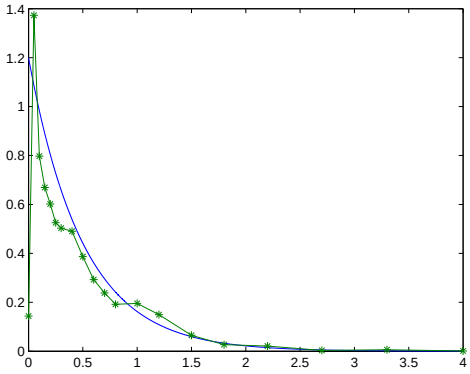
$$\begin{aligned} f_1(t) &= 1.2e^{-2t} \\ f_3(t) &= 20te^{-5t} \\ f_4(t) &= 15te^{-9t^{1.7}} + t^3e^{\frac{t^4}{2}} \end{aligned} \tag{E.1}$$

Each input function was convolved with the system (E.2), the output was 'blurred' with 15% randomly distributed noise and sampled at the following points $t = [0\ 0.05\ 0.1\ 0.15\ 0.2\ 0.25\ 0.3\ 0.4\ 0.5\ 0.6\ 0.7\ 0.8\ 1.0\ 1.2\ 1.5\ 1.8\ 2.2\ 2.7\ 3.3\ 4.0]$. The resulting output was passed to the Maximum Entropy implementation along with an initial guess for the input function, take as the mean of the output.

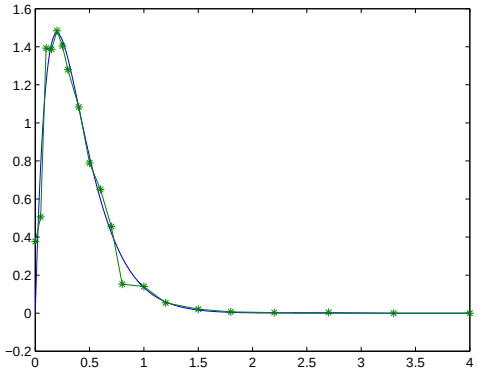
$$\begin{aligned} \dot{x} &= Ax + Bf(t) \quad x(0) = 0 \\ y &= Cx \end{aligned} \tag{E.2}$$

$$A = \begin{bmatrix} -6 & -5 \\ 1 & 0 \end{bmatrix} B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad C = \begin{bmatrix} 2 & 6 \end{bmatrix}$$

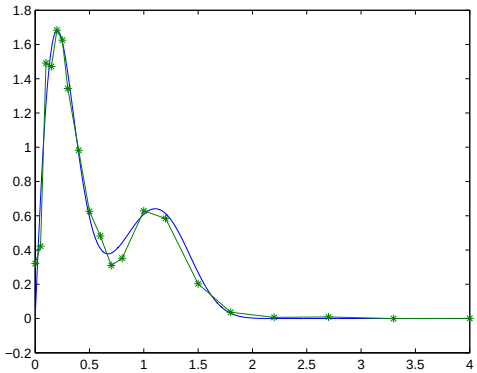
$$
\begin{aligned}
g(t) &= e^{-t} + e^{-5t} \\
y(t) &= \int_0^t g(t-s) f(s)\, ds
\end{aligned}
\tag{E.3}
$$



(a) input $f_1(t)$ 15% noise



(b) input $f_3(t)$ 15% noise



(c) input $f_4(t)$ 15% noise

Figure E.1: *Maximum Entropy recovered test input functions (eq. E.1); (solid line) the real input function, (*-line) recovered input, * indicates the input sample points as defined above.*

As can be seen in Figure E.1 the deconvolved input is succesfully recovered for each of the four input signals.

Appendix F

Implementation of Maximum Entropy Regularisation

The following code is a Matlab function implementing the Maximum Entropy Regularisation using the *fmincon* function to minimise the inverse of the entropy of the input function constrainted by the chi-squared metric.

```
function [xout,chi2out,messout,grad,hessian] ...
= maxentExample(initial_guess,...
                sampled_output,...
                t_sample,...
                t_sample_in,...
                PERC_NOISE,...
                dbg,...
                x0)
% [xout,chi2out,messout,grad,hessian] =
% maxentNonZeroInit_ode(initial_guess,sampled_output,t_sample,t_sample_in,
%                       PERC_NOISE,dbg,x0,t_on)
%
% Example function implementing max ent regularization.  The function opus
% should be changed to implement the underlying model.  In this case a
% switched system is implemented as a simpled ODE. Please note this is not
% a robust production worth function, it is here as a guide to using
% fminson and maxent regularisation.
%
```

```
% Inputs: (not all inputs are required)
%
% initial_guess – vector of initial guess, same dimension as t_sample_in
% sampled_output – observations
% t_sample       – time points of obsevation
% t_sample_in    – number of nodes in output function.
% PERC_NOISE     – Estimates on noise measurement, assumed scalar.
% dbg            – boolean flag (0 – output plots, 1 – don't).
% x0             – initial conditions for ODE
%
% Ouputs:
%
% xout           – Deconvolved input
% chi2out        – chiSquared measure (how close to output)
% messout        – Output message (if has an error check it out)
% grad           – gradient of Lagrangian at solution xout.
% hessian        – hessian of Lagrangian at solutout xout.
%
% Written by
%   J Hattersley,
%   Systems Modelling and Simulation Group
%   School of Engineering
%   University of Warwick,
%   December, 2008
%

%=========================================================================%
% Stop divide by zero issues
%=========================================================================%
ZEROISH = 1E-5;

ns = 100;

%=========================================================================%
% Setup for Chi^2 stuff.
%=========================================================================%
sigma = sampled_output*PERC_NOISE;
sigma(find(sigma == 0)) = ZEROISH;
```

```
Caim = length(sampled_output) + 2*sqrt(length(sampled_output));

%=========================================================================%
% Setup Optimization
%=========================================================================%
OPTIONS_INNER=optimset('LargeScale','on');
OPTIONS_INNER = optimset(OPTIONS_INNER,'Display','iter');
OPTIONS_INNER = optimset(OPTIONS_INNER,'Diagnostics','off');
if dbg > 0;
   OPTIONS_INNER=optimset(OPTIONS_INNER,'OutputFcn',@printEm);
end
OPTIONS_INNER = optimset(OPTIONS_INNER,'MaxIter',100);

% As it's normalised...
lb = ones(1,length(initial_guess))*ZEROISH;
lb(1) = max(initial_guess);
ub = ones(1,length(initial_guess))*max(initial_guess);

x = initial_guess;

[x,fval,exitflag,output,lambda,grad,hessian] = fmincon(@objFunc,x,[],...
                                                       [],[],[],lb,ub,...
                                                       @myconfun,...
                                                       OPTIONS_INNER);


chi2out = chiSquared(x);
xout = x;
messout = output;

   %=============================================================
   % debug - just output some plots for diagnostics
   %=============================================================
    function stop = printEm(x,optimvalues,state,userdata,varargin)
                stop = false;
        switch state
            case 'init'
                figure;
                % hold on
```

```
        case 'iter'
            [t_sim,F] = opus(x);
            output_sim_C1 = F(:,1);
            output_sim_C2 = F(:,2);

            subplot(2,1,1);
               plot(t_sample,sampled_output,'ro',...
                    t_sim,output_sim_C1,'-',...
                    t_sim,output_sim_C2,'--','Linewidth',2.0);
               ylabel('IgG Plasma g/L');
               xlabel('time (days)');
               %legend('Real','Simulated C1','Simulated C2');
               title('System output');
               axis tight;

            subplot(2,1,2);
               plot(t_sample_in,x,'-ro',...
                    t_sample_in,initial_guess,...
                    t_sample_in,ub,':','Linewidth',2.0);
               ylabel('IgG Synthesis g/day');
               xlabel('time (days)');
               % legend('Recovered','Initial');
               title('Recovered Input');
               axis tight;

               % Need this pause or you won't see the plots.
               pause(0.1);

        case 'done'
            % hold off
        otherwise
    end
 end %end of printEm

%================================================
% Objective and Constraints
%================================================
 % For nonlinear constraints
```

```
function [cin,ceq] = myconfun(x)
    % Catch all in case fmincon gets stuff with negative values.
    % This should never be the case, if this is called fix it.
    if ((isreal(x)==0) || (length(find(x<0))>0))
            cin = 1E100;
    else
        % In equality contraint
        C = chiSquared(x);
        cin = C - Caim;
        % Equality constraint - for Max Entropy
        % all inputs should only sum up to 1!
    end
    ceq = [];
end %end of myconfun

function C = chiSquared(x)
    [t1,output_sim1] = opus(x);
    output_sim1 = output_sim1(:,1);
    C = sum(((((sampled_output - output_sim1')./sigma).^2)));
end %end chiSquared

function [obj] = objFunc(x)
        % Smooth attempts to remove peaks/troughs from the signal -
        % make it more natural.
        m = zeros(1,length(x));
        m(1) = 0.5*(x(2)+x(3));
        for i=2:length(x)-1
           m(i) = 0.5*(x(i-1) + x(i+1));
        end
        m(end) =  0.5*(x(end-2)+x(end));

        smoothed_x = x./m;
        logx = log(smoothed_x);
        % Remember entropy = -sum(x log(x))
        % but we're after maximum, so *-1
        obj = sum(x.*logx);
end %objFunc
```

```
    function [T,F] = opus(x)
        [T,F] = dialODE(x,t_sample,t_sample_in,x0);
    end

end % End of Max Ent
```