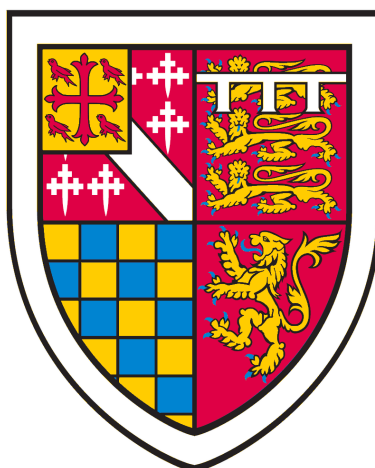VINCENT RUSSELL KNIGHT-SCHRIJVER

TOWARDS SYSTEMS PHARMACOLOGY MODELS OF DRUGGABLE TARGETS AND DISEASE MECHANISMS

UNIVERSITY OF CAMBRIDGE

TOWARDS SYSTEMS PHARMACOLOGY MODELS OF DRUGGABLE TARGETS AND DISEASE MECHANISMS

VINCENT RUSSELL KNIGHT-SCHRIJVER

This dissertation is submitted for the degree of Doctor of Philosophy

Babraham Institute
St Edmund's College
University of Cambridge

March 2018

Dedicated to Merlin "*magic paws*" Knight-Schrijver.

$2015 - 2016$

His purr-review and contributions to early versions of this work
have long since been edited, but not forgotten.

DECLARATION

I hereby declare this dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. This dissertation contains fewer than 60,000 words excluding bibliography, figures, appendices, footnotes, tables and equations.

Cambridge, March 2018

Vincent Russell
Knight-Schrijver

TOWARDS SYSTEMS PHARMACOLOGY MODELS OF DRUGGABLE TARGETS AND DISEASE MECHANISMS

The development of essential medicines is being slowed by a lack of efficiency in drug development as ninety per cent of drugs fail at some stage during clinical evaluation.

This attrition in drug development is seen not because of a reduction in pharmaceutical research expenditure nor is it caused by a declining understanding of biology, if anything, these are both increasing. Instead, drugs are failing because we are unable to effectively predict how they will work before they are given to patients. This is due to limitations of the current methods used to evaluate a drug's toxicity and efficacy prior to its development. Quite simply, these methods do not account for the full complexity of biology in humans.

Systems pharmacology models are a likely candidate for increasing the efficiency of drug discovery as they seek to comprehensively model the fundamental biology of disease mechanisms in a quantitative manner. They are computational models, designed and hailed as a strategy for making well-informed and cost effective decisions on drug viability and target druggability and therefore attempt to reduce this time-consuming and costly attrition.

Using text mining and text classification I present a growing landscape of systems pharmacology models in literature growing from humble roots because of step-wise increases in our understanding of biology. Furthermore, I develop a case for the capability of systems pharmacology models in making predictions by constructing a model of interleukin-6 signalling for rheumatoid arthritis. This model shows that druggable target selection is not necessarily an intuitive task as it results in an emergent but unanswered hypothesis for safety concerns in a monoclonal antibody. Finally, I show that predictive classification models can also be used to explore gene expression data in a novel work flow by attempting to predict patient response classes to an influenza vaccine.

PUBLICATIONS

During my studentship I have been named as an author on several published articles:

Chelliah, V. et al. (2015). 'BioModels: ten-year anniversary'. In: *Nucleic Acids Res.* 43.Database issue, pp. D542–548. DOI: 10.1093/nar/gku1181.

Knight-Schrijver, V.R., V. Chelliah, L. Cucurull-Sanchez and N. Le Novère (2016). 'The promises of quantitative systems pharmacology modelling for drug development'. In: *Computational and Structural Biotechnology Journal* 14, pp. 363–370. DOI: 10.1016/j.csbj.2016.09.002.

Waltemath, D. et al. (2016). 'Toward Community Standards and Software for Whole-Cell Modeling'. In: *IEEE Trans Biomed Eng* 63.10, pp. 2007–2014. DOI: 10.1109/TBME.2016.2560762.

"So remember to look up at the stars and not down at your feet. Try to make sense of what you see and wonder about what makes the universe exist. Be curious, and however difficult life may seem, there is always something you can do, and succeed at. It matters that you don't just give up."

— Stephen Hawking, (8 January 1942 –14 March 2018)

ACKNOWLEDGMENTS

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LISTINGS

ACRONYMS

7TD1   murine hybridoma cell line

AD   Alzheimer's disease

ADME absorption, distribution, metabolism and elimination

AIC   Akaike information criterion

ANOVA analysis of variance

Cal33   tongue cell carcinoma cell line

CD   Crohn's disease

CD3+   cluster of differentiation 3 positive

CD4+   cluster of differentiation 4 positive

C/EBPβ CCAAT/enhancer binding protein β

CNS   central nervous system

COS-7   monkey kidney fibroblast-like cell line 7

CRP   C-reactive protein

EC50   effective concentration 50 (% of maximum effect)

FDA   U.S Food and Drug Administration

gp130   glycoprotein 130

HA   haemagglutinin

HAI   haemagglutination inhibition

HepG2   human hepatocellular carcinoma cell line

$t_{1/2} - \beta$ elimination half-life

HuH7   human hepatocyte-derived cellular carcinoma cell line

IgG   immunoglobulin-G

IL-17   interleukin 17

IL-23   interleukin 23

IL-5   interleukin 5

IL-6   interleukin-6

IL-6Rα IL-6 receptor-alpha

ISF   interstitial fluid

IV   intravenous

JAK   janus kinase

K_D dissociation equilibrium constant

KNN   K-nearest neighbour

LSODA   livermore solver for ordinary differential equations

mAb   monoclonal antibody

MACE   major adverse cardiovascular events

MAMO   mathematical modelling ontology

MAP   *maximum a posteriori*

MAPK   mitogen-activated protein kinase

MeSH®   medical subject headings

mIL-6Rα membrane-bound IL-6 receptor-alpha

MIRIAM   minimal information requested in the annotation of biochemical models

MM   Michaelis-Menten

MoA   mechanism of action

MTX   methotrexate

NBC   naïve Bayes classifier

mNBC   multinomial naïve bayes classifier

NLP   natural language processing

ODE   ordinary differential equation

PBMC   peripheral blood mononuclear cell

PBPK   physiologically-based PK

PCA   principal component analysis

PD   pharmacodynamics

PDF   probability density function

PI3K   phosphatidylinositol-4,5-bisphosphate 3-kinase

PK   pharmacokinetics

PKPD   pharmacokinetic-pharmacodynamic

pSTAT3   phosphorylated STAT3

Q2W   fortnightly administration

Q4W   monthly administration

QSP   quantitative systems pharmacology

QSPM   QSP modelling

RA   rheumatoid arthritis

RF   random forest

RME   receptor-mediated endocytosis

R&D   research & development

SBGN   systems biology graphical notation

SBML   systems biology markup language

SC   subcutaneous

SF   synovial fluid

sgp130   soluble glycoprotein 130

sIL-6Rα soluble IL-6 receptor-alpha

SOCS   suppressor of cytokine signalling

SRK   sirukumab

STAT   signal transducer and activator of transcription

STAT3   signal transducer and activator of transcription 3

SVM   support vector machine

TCZ   tocilizumab

tet+   haemagglutinin tetramer-positive

TF-IDF   term frequency-inverse document frequency

TM   text mining

TMDD   target-mediated drug disposition

TN   true negative

TNF-α tumour necrosis factor-alpha

TPB   true positive biological models

TPP   true positive pharmacological models

U266   human multiple myeloma cell line

UMLS   Unified Medical Language System

WHO   World Health Organization

INTRODUCTION

1.1 THESIS OUTLINE

This thesis is presented in four main sections, each with a discussion. This is rounded by a conclusion which ties together the discussions in each chapter. Chapter one provides a general introduction to systems pharmacology models and introduces the concepts of drug discovery, target selection and modelling methods that have defined the field as it stands. A brief more specific introduction is given in each chapter where the subject deviates from the main matter.

The second chapter illustrates the results of published text mining research I performed to define the landscape of quantitative systems pharmacology modelling. This chapter also documents the effort to further this research by developing a supervised text classification model to categorise and retrieve systems pharmacology models for annotation, documentation and analysis.

Chapter three is an applied case study using quantitative systems pharmacology to aid target selection in the disease context of rheumatoid arthritis. This covers the success of such an attempt, emergent implications of current medication as well as the ease and accessibility of reusing purpose-built models for further modelling programs.

The fourth chapter illustrates an alternative use for the classification model in gene expression data to predict patient response categories before administering therapy.

The final chapter concludes the thesis with a discussion of findings as well as an outlook on future prospects for the research presented within the thesis.
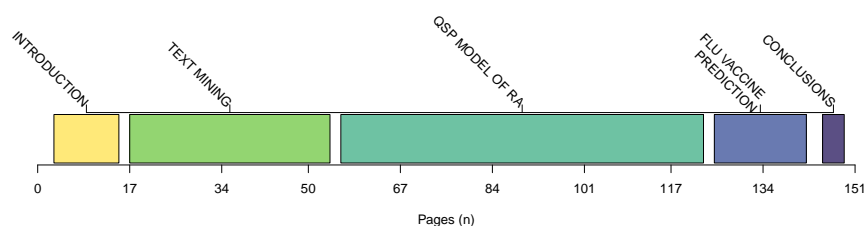
Figure 1.1: Outline of the thesis excluding front matter and bibliography.

1.2   CONTEXT AND MOTIVATION

Biomedical research continues to produce effective medicines for humanity which undoubtedly contributes to world-wide increases in life expectancy and quality of life. These medicines are typically generated through drug development and the refinement of basic biological research into a suitable means of preventative, curative or palliative intervention. This transition from laboratory to dispensary is a lengthy process. Currently, the average development time is ten years for most drugs which would be perfectly acceptable if a successful outcome was certain. However, drugs have been failing during this process due to a variety of reasons. For example, close to 90 per cent of drugs fail during their clinical phases (Hay et al., 2014). Often called attrition, failure at this stage is highly problematic because it removes time from researching otherwise successful medicines and it costs more than if the failure was detected or recognised earlier.

In context, drug development is becoming synonymous with high-cost and low yield research. To overcome the challenges faced in the domain, many methods have been used to varying degrees of success. A successful field with extensive applications in predicting drug action and informing the development process is mathematical modelling and simulation. For example, the use of mathematical models in the last 30 years may have been momentous in reducing drug failure resulting from the main causes of late-stage attrition. Despite this, tackling these causes has shifted the leading driver of attrition to other equally contributory factors (Kola and Landis, 2004; Waring et al., 2015). Another milestone in modelling and simulation may therefore harness humanity's current understanding of the disease mechanisms, drugs and druggable targets to protect against the current sources of attrition in drug development.

One solution stems from the holistic understanding of biology in applying a non-reductionist point of view. This is an effort to interpret dynamic systems as fully as possible called systems biology. In using systems-level concepts to understand how diseases and their mechanisms can be reversed by therapeutics, we can derive the term systems pharmacology as is discussed in this research. Systems pharmacology models aim to mathematically describe the detailed mechanisms that drive pathology which are now detectable through advancements in biomedical research. In doing so, the effects of drugs within a system may be predicted, giving pharmaceutical researchers a framework for detecting drug failure earlier. The concept of systems pharmacology models may be further applied to predicting suitable strategies in exploratory and proactive use, accelerating drug discovery and the speed of generating new medicines.

This research aims to put forward tangible concepts within systems pharmacology modelling, revealing a direction for future focus within the field. Herein I discuss the history and current literature on drug development and modelling in pharmacology as well as the mature and fledgling aspects of model development and the broader context of predictive modelling in therapeutics.

The focus is upon two central aspects of systems pharmacology modelling. Firstly, how can these models explore and understand disease mechanisms? Secondly, how they can be used to decide between druggable targets? The research addresses the key motivation of the emerging field by answering following question: what are systems pharmacology models, where did they come from and how can we harness them to reduce the causes of failure in the drug development and discovery and boost biomedical and pharmaceutical productivity?

1.3 DRUG DEVELOPMENT AND DISCOVERY

Drug development and discovery is the overall process by which new therapeutics are generated and approved. It can be divided into two themes, the first is one of discovery designed to produce an effective drug through basic research and refinement. The second theme is one of validation which makes use of proof-of-concept clinical trials to demonstrate the drug's worth. Significant challenges become apparent in the latter phase which can be attributed to the former discovery phase.

1.3.1 *The challenges in drug development*

A modern-day definition of drug development was neatly encapsulated by Sheiner and Steimer (2000) as

> "the information-gathering activities that begin when a lead compound is first introduced into man and that end when the accumulated information is summarized and presented to a regulatory agency for a market-access decision."

The main challenge that faces drug development today is an exponential rise in the cost of producing drugs (DiMasi, November 18, 2014; Knight-Schrijver et al., 2016). The cost of developing drugs is rising partly because more work is now required to identify, optimise and market a drug successfully but mostly because a large fraction of drugs fail during clinical phases; for each successful drug, nine fail at a stage where the costs are astronomically high. For patients, this manifests itself as a lack of effective medication for significant diseases such as cancer or Alzheimer's disease (AD) and continuing

adverse reactions in current treatment strategies. AD is an extreme example of drug attrition as nearly 100 per cent of AD therapies fail during the clinical stages (Cummings et al., 2014). Assuming that clinical trials are properly designed and rigorously carried out, we can reason that the result is wholly reliant on the input (the drug(s)). Therefore, in tackling failure during developmental stages, more needs to be understood about how drugs are generated during the initial discovery period.

1.3.2 *Rationalising drug discovery*

Drug discovery began in earnest at the beginning of the twentieth century. Early work of discovering drugs was the remit of chemists, and substances with biological effect were structurally altered to maximise their activity. Understanding the mechanism of action (MoA) for these drugs were a life time away for scientists who developed them, let alone the population receiving them. One example of this is aspirin, which was synthesised and marketed properly in 1899 by Bayer Inc after thousands of years of observation (Levesque and Lafont, 2000). The actual MoA for aspirin was only revealed seventy two years later by Vane (1971) (the magnitude of which earned him the Nobel Prize for Physiology or Medicine in 1982). These *Then* years of the medicinal chemist (Lombardino and Lowe, 2004) were blessed with a series of lucky finds (Ban, 2006) and pharmaceutical research adopted what can be described as a *drug-down* approach to the initial stages of drug development. Conceptually, the strategy was to observe that a substance has a biological effect, propose a therapeutic use and demonstrate its clinical effectiveness in *in vitro* or *in vivo* studies. For example, pioneering research in chemistry produced cortisone which was given patients for relieving rheumatoid arthritis despite the unknown mechanism behind the efficacy (Hench et al., 1949). A string of other important discoveries (Ban, 2006) are also attributed to have occurred after initially observing clinical effects with much influence upon the therapeutic landscape still today.

By probing our biological black-boxes with chemical inputs, a *drug-down* approach has allowed researchers to develop a picture of the underlying physiology and gradually form more targeted approaches in treating diseases. Now that this picture is beginning to be painted more clearly, a *biology-up* approach is gaining momentum. Advancements in technology such as the initial decoding of the human genome (Lander et al., 2001) allow for a previously unattainable understanding of the fundamental disease mechanisms contained within our bodies. More information is revealed about the normal biological networks and the disruption of them that characterises disease. This valuable information can be readily applied to selecting single or multiple points of impact within these networks to produce the desired

outcome. Intending to use the breadth of biological knowledge suggests the importance of systems-level thinking in drug discovery. A *biology-up* view is the ideal basis for modern rational or target-based drug discovery pipelines.

Preclinical rational drug discovery can be loosely boiled down to two main concepts: 1) Select a good target; and 2) Select a good drug. The first point is the comprehensive dissection of the disease mechanisms as well as the evaluation of target druggability within the disease. This is the typical starting position for a *biology-up* paradigm. The second step addresses the selection and optimisation of a drug which can affect this target (lead discovery) through to validation of use in a clinical setting. Selecting a good target and a good drug have essentially the same list of boxes to tick.

1.3.3 *A good target is a druggable target*

By target, we mean any molecule within the body associated with a biological process, typically proteins or DNA.

Apart from a few serendipitous discoveries (Ban, 2006), most therapeutic drugs have clear targets: either those they were selected against or those which were uncovered during discovery stages.

In rational drug discovery, the ability to select between targets in pre-clinical stages is essential in developing effective medicines. For example, consider two proteins A and B, which positively regulate a disease. The inhibition of either may produce markedly different effects upon the overall disease activity. Where attenuating protein A may show high efficacy, its blockade in force may result in overly toxic effects. Alternatively, the perturbation of protein B might result in negligible clinical effect. A good target is often a compromise between the demonstrable efficacy and toxicity. One description of the relative worth of a target is druggability which is often defined as how easy a target is to bind to (Bakan et al., 2012) and moreover refers to ease that a target can be modulated. I would argue that, given the example of proteins A and B, it is equally important to define druggability as a function of the target's position as an essential piece of disease mechanisms as well as the target's safety profile.

Druggability - A target's ability to be readily modulated and a dichotomy of its position as a keystone in driving disease mechanisms and its relevance to healthy function.

LOCALISATION: A good target should be easily located and bound to by drugs.

EFFECT: A good target must be a key player in the disease, offering maximal potential for pathological perturbation.

SPECIFICITY: Modulation of a good target should be relatively benign towards healthy function.

As poor clinical efficacy and poor safety of a drug are the major contributors to attrition in clinical phases (Harrison, 2016; Kola and Landis, 2004), we can consider that the success of a novel drug discovery program depends largely on initial target selection. This de-

pendence was shown by introducing a more stringent set of target entry guidelines at AstraZeneca in 2003 as implementing new criteria may have been responsible for a four-fold decrease in program failure in the following lead optimisation phase (Jackson, 2014). This confirms that selecting between targets is a large component of optimal drug discovery and suggests that evidence-based predictive measures can reduce the larger expenditure in further clinical stages. Therefore, measuring druggability is one of the key factors in drug discovery's success and thus the development of methods to evaluate druggability are of great importance.

1.3.4 *A good drug is a targetable drug*

A drug should be able to modulate the selected targets in the intended manner with ease. This means that, at biological concentrations, the drug should be able to localise efficiently with, be specific for, and be able to functionally alter its targets for a therapeutic response. This is generally either to increase or reduce the target's current functionality. The failure of a drug to maximise these three major traits results in a loss of efficacy or a gain in drug-dependent toxicity.

LOCALISATION: A good drug needs to be delivered efficiently to the desired tissue or system for drug-target interaction.

EFFECT: A good drug should alter a target's function effectively after localisation to produce the desired result.

SPECIFICITY: A good drug should not interact with endogenous or exogenous components other than the intended target(s).

Maximising the target localisation of a drug on a macro-scale is largely driven by optimising its molecular properties and administration route. For example, inhibiting a receptor in the central nervous system (CNS) would typically be difficult using a large polar molecule when delivered via the circulatory system (Mikitsh and Chacko, 2014; Scherrmann, 2002). However, a more direct administration into the CNS could help make this drug viable (Cohen-Pfeffer et al., 2017).

Optimising the effect of the drug-target interaction is regulated by the affinity of the drug to the target as well as the intrinsic efficacy of the drug. For an effect focusing upon competitive inhibition, a higher affinity results in a greater potency. Alternatively, where a functional change is sought, the potency of a drug is a combination of both the drug's affinity as well as the ability of the drug to modify the target conformation for a gain or loss of function. This is the case of both allosteric and orthosteric ligands.

Poor specificity of the drug is the potential for binding to non-target components. This results in off-target drug-dependent toxicity which can be divided into several categories including bioactivation,

idiosyncratic responses, and immunogenicity or immune hypersensitivity (Guengerich, 2011). Drug-dependent toxicity depends largely upon the formulation of the compound and the route of administration and increasing the specificity of drugs may reduce the chances of drug-dependent toxicity. Lastly, another challenge that may hinder a good drug is the properties of its metabolites.

This is to emphasise that developing medicines is not necessarily about driving towards a strong and stable compound nor is it about finding the finest of the best of targets. It is about identifying drug–target relationships demonstrating optimal efficacy and toxicity profiles which is not so straight forward. We have to consider that even after evaluating the best drug for the best target within the network, there may be many unknown results caused by perturbations of disease mechanisms. After all, it is a very large network of interacting components within a black box.

1.4 MODELLING DISEASE MECHANISMS AND DRUGGABLE TARGETS

A significant fraction of knowledge in biological research is derived from the use of *in vitro* or *in vivo* models. Animals have been used to study biology for at least the last two millennia (Ericsson et al., 2013). More recently, cultured cells and *in vitro* organoid models of biological systems have been used to understand the mechanistic details in biology (Jackson and Lu, 2016). Now, there are a plethora of *in vitro* disease models (Benam et al., 2015) which can be used to understand how small-scale systems may react to potential target perturbations. For instance, this can be used to differentiate how distinct organs may react differently to a drug. Similarly *in vivo* models can demonstrate how a target can be simultaneously essential for the maintenance of a disease model as well as safe to modulate through drug interaction using knockout models or ribonucleic acid interference (Lee, 2014).

However, *in vitro* and *in vivo* models have their drawbacks. Assessing the druggability of a target involves the context of physiology. Therefore, the use of *in vitro* systems to measure the kinetics and interaction of drugs within the wider system of human physiology is flawed. Moreover, the physiology of animal *in vivo* models may be drastically different to humans bringing great challenges in translating their outcomes into humans. Examples like the animal model of leprosy in armadillos speak volumes about the selective and sometimes anthropomorphic measures applied to secure a working representation of human disease (Truman et al., 2014). Successful outcomes in efficacy and (lack of) toxicity in animals has occasionally not been translated into humans which contributes to the picture of attrition today (Garner, 2014; Shanks et al., 2009). Even in accounting for this inter-organismal disparity, the knockout of potential targets to

demonstrate their role in animals may be an unrealistic depiction of real-world inhibition in humans where therapeutic windows are acquired and target occupancy is rarely near 100 per cent (Sams-Dodd, 2005). Fortunately, treating humans with novel and potentially fatal compounds is deemed unethical in most situations today, which limits the number of options in choosing an ideal disease model. Finally, while both *in vitro* and *in vivo* versions of disease are instrumental in gathering data and examining intricate mechanisms, to combinatorially examine a library of compounds against either of these models is not only highly time consuming, but also expensive. To this end, mathematical and computational models have been used which attempt to reproduce the important aspects of drug activity in humans by simulating and predicting drug effects while minimising cost and time.

1.4.1 *Mathematical Modelling in Drug Discovery*

Almost one century of mathematical modelling has been applied within the field of pharmacology. During this time, the core concepts and methods have been refined and the field has iteratively built upon itself, resulting in more advanced techniques and meaningful predictive methods to encompass the growing knowledge of biology.

1.4.1.1 *Definitions of modelling*

A hefty portion of mathematics in pharmacological modelling is based upon ordinary differential equations (ODEs), which are almost always numerically solved with respect to time. This is largely due to the number of measurements taken with reference to the time domain as well as the value of rendering model simulations as being readily interpretable in the context of distribution, elimination, disease progression and the majority of other tangible observations associated with biology. For example, a system of n pharmacological variables, represented as a vector X where $X = x_1, x_2, ..., x_n$, can be modelled by a series of ODEs. The whole system can be modelled by equation 1.1 where the time-dependent change in each nth variable of X (x_n) is governed by the rate function for each variable f_{x_n}, describing the molecular interactions between the variables in X. Other mathematical methods include agent-based models or stochastic models which have both been instrumental in a number of studies, especially in more recent years (Cosgrove et al., 2015; Traynard et al., 2017).

$$\frac{dx_n}{dt} = f_{x_n}(X, t) \tag{1.1}$$

1.4.1.2 *Pharmacokinetic models*

For want of a better definition, pharmacokinetics (PK) is the study of *the effect that the body has upon the drug* and was only coined in 1953 (Wagner, 1981) despite its study decades earlier. Today's models of pharmacokinetics can be ascribed to works in the 1930s which may have triggered the beginning of mathematical modelling in pharmacology. In this decade, equations were first derived for the absorption, distribution and elimination of drugs using simple compartment models (Wagner, 1981). A few examples of influential models are a series of ODEs by Teorell (1937a,b) which illustrated the concept of simulating drug PK in multiple tissue compartments following subcutaneous distribution and were some of the founding examples of PK modelling. A simplified diagram of such a PK model can be seen in Figure 1.2A. PK modelling has expanded rapidly since Teorell's work and is now a standard element in drug development pipelines. PK modelling today is used to predict the absorption, distribution, metabolism and elimination (ADME) properties of drugs and its use is streamlined by the availability of dedicated PK modelling software like SAAM and NONLIN which began to emerge towards the end of the 1960s.
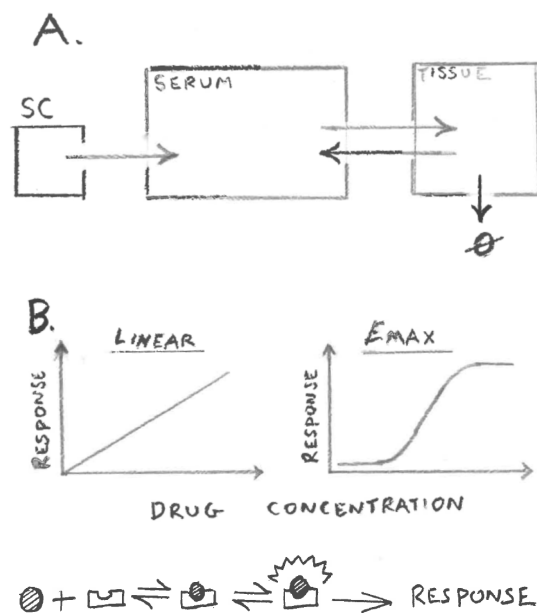
Figure 1.2: **A**, a hypothetical compartmental PK model depicting subcutaneous (SC) administration of a drug and its distribution into two compartments with elimination in the tissue only as given by arrows. **B**, linear or E_{max} response models and an example of a biochemical reaction governing the response only after agonist (circle) binding to the receptor (inverted arch) and resulting complex activation. This reaction scheme results in a sigmoidal response for certain parameter sets.

1.4.1.3 *Pharmacodynamic models*

As the complement of PK, understanding pharmacodynamics (PD) is to study *the effect that the drug has upon the body*. The data for exploring these drug-response relationships stems from *in vitro* studies and is usually in the form of a measured response to a fixed concentration of drug. Initial regression-type models observing a linear response to drug concentrations were built upon by considering that the systems have maximal and minimal responses in log-space (Holford and Sheiner, 1982). This evolved such that dynamic models of pharmacological response used today are often forms of the sigmoidal E_{max} model derived from much earlier work by Hill (1910) (Gesztelyi et al., 2012; Holford and Sheiner, 1982). Further but much later work in rationalising the link between drug concentrations and the response of the system was made by Black and Leff (1983) in forming the operational models of pharmacological agonism. The model integrates the E_{max} model with mass action binding kinetics of drugs which explicitly accounts for disparate potencies seen between compounds when acting upon the same receptor or system. The differences between these views on PD can be see in Figure 1.2B.

1.4.1.4 *Pharmacodynamic pharmacokinetic models*

Pharmacokinetic-pharmacodynamic PKPD models bring two of the essential pieces of the pharmacological puzzle together. On the one hand we saw that drugs produce a response in cultured tissues and experiments, which could be modelled according to mathematical relationships (PD) (Figure 1.2B). On the other hand, it was possible to simulate the concentration of the drug over time in multiple compartments based on experimental clinical and animal studies (PK) (Figure 1.2A). It was only a matter of time until the two were combined, using the output from PK simulations to determine the input concentration for PD response model. Pharmacokinetic-pharmacodynamic (PKPD) models first began to emerge in the 1960s and allowed pharmacologists to simulate time-dependent concentration–response scenarios using simulated PK output as the input for the PD models (Csajka and Verotta, 2006; Holford and Sheiner, 1982). Modern day PKPD modelling is used in late-stage prediction and validation of dose optimisation but has also been successfully transferred to the earlier stages of drug discovery (Agoram et al., 2007). A detailed account of the modern methods of PKPD modelling can be found in Sheiner and Steimer (2000).

1.4.1.5 *Physiologically-based pharmacokinetics models*

The use of compartments in PK modelling is simplified to theoretical volumes of distribution which limits the interpretation of the model (Figure 1.2A for example). In seeking to address this, physiologically-

based PK (PBPK) models account explicitly for the majority of physiological compartments within an organism. This aids translatability between organisms where data are derived from *in vivo* studies as well the use of the models as platforms which are not tailored to a specific compound.

1.4.1.6 *Population pharmacokinetics*

Differences in drug effects between patient populations are evident in the quantity of patients receiving pharmaceuticals in today's society. Population-PK modelling is designed to model these differences using statistical distributions and thus form population-level predictions of drug ADME as well as PD and toxicity when combined with PD models. For example, modelling patient populations is useful for understanding the effects of drugs in at-risk patient groups by assessing drug metabolism and genetic polymorphisms (Yoo et al., 2012).

1.4.2 *Challenging the success of pharmacometrics*

The use of PK, PD and PKPD models in quantifying dose-response relationships and population variance is now termed pharmacometrics (Ette and Williams, 2007; Standing, 2016). In short, pharmacometrics is a powerful approach to predicting drug kinetics, response and toxicity within individuals and populations. It is currently used by a large fraction of industry with generally good results at both preclinical as well as clinical stages (Schuck et al., 2015). Furthermore, the use of pharmacometric approaches have been seen to be pivotal in a large fraction of successful new drug applications seen by the U.S Food and Drug Administration (FDA) between 2000 and 2008 (Bhattaram et al., 2007; Bhattaram et al., 2005; Lee et al., 2011).

PK models can simulate the tissue concentrations of drugs, they can be used to estimate occupancy and to a certain extent, off-target toxicity to arrive at optimal dose predictions. However, to derive meaning from these simulations, the volumes and compartments used in the model should to be translatable into human physiology; PBPK models address this caveat. Additionally, drug PK is now well understood as the variables involved are largely known and measurable (for example, drug concentrations or metabolites). Perhaps as a result from this knowledge, the widespread use of PK models might have contributed to the attenuation of clinical PK-mediated attrition seen between 1990 and 2000 (Kola and Landis, 2004).

However, the historical array of modelling approaches needs to be adapted to address challenges faced by researchers today. Contrasting with the putative success of PK modelling, the major causes of attrition in drug development now are efficacy and toxicity (Waring et al., 2015), which are both placed firmly within the scope of understanding the disease mechanism and selecting the best druggable

target. Typical PD models such as the operational model (Black and Leff, 1983) are black-box systems. This means that they can be used to quantify the magnitude of a response even though the machinery driving the output may be poorly understood. However, as further intricacies of biology and disease systems are uncovered, PD models are forecast to progress towards a more mechanistic representation of the disease mechanisms (Csajka and Verotta, 2006). This is the remit of systems pharmacology models.

1.5   NETWORKS IN BIOLOGY

To tackle the failure of drugs resulting from toxicity or poor efficacy, more needs to be modelled within the black-box. The solution is to re-produce as much of the known biology as possible to understand the disease mechanism. Instead of simplification and reduction, the goal is (reasonable) complexity and (limited) holism. One method to man-age and infer useful information from complex biological systems is to construct networks.

Networks are a system of nodes, connected to each other by edges. Nodes represent the entities within the network and edges describe the relationship between these nodes. Tangible large-scale networks may be electrical grids (Pagani and Aiello, 2013) or road systems (Balijepalli and Oppong, 2014). However, when biological networks are discussed, nodes are usually in the form of proteins, chemicals and genes (Alm, 2003) and edges in biology often portray a physical interaction. You may think of these as drug targets and drug interactions respectively.

A drug with polypharmacology modulates multiple targets although not necessarily by design.

Polypharmacy is the concurrent use of multiple drugs to modulate multiple targets for a synergistic effect.

One example of using networks in drug discovery is seen in ef-forts to strategically develop compounds with deliberate polyphar-macology (Csermely et al., 2005; Hopkins, 2008). Polypharmacology and polypharmacy have the potential to counter pathological redund-ancy, maximise efficacy and lower toxicity by focusing on smaller but orchestrated perturbations of multiple network targets (Kitano, 2007; Korcsmáros et al., 2007; Morrow et al., 2010).

1.6   SYSTEMS BIOLOGY AND SYSTEMS PHARMACOLOGY

By progressing towards a view that networks are dynamic and not static it becomes feasible to model strictly time-dependent phenom-ena such as oscillations relating to chronotherapy or the cell cycle in cancer (Dulong et al., 2015; Jackson et al., 2017). Studying dynamical systems or networks in biology is referred to as systems biology.

1.6.1 *Systems biology*

The definition of systems biology has eluded an overwhelming subset of researchers, prompting the question *What is Systems Biology?*[1] by Breitling (2010).

The term systems biology is applied to research that studies large scale *omics* data using networks and computational modelling. A simple enough description of systems biology can be amalgamated from a few sources: Systems biology aims to be able to fully reconstruct a biological system by understanding the structure and dynamic nature of its wiring, including knowledge of substructure compositions necessary for regulation and control of the system (Breitling, 2010; Kitano, 2002). This is much like being able to fully understand how a radio operates through the complete knowledge of its circuitry (Lazebnik, 2002)[2].

"ask two systems biologists for the definition of their discipline, and you will get three answers"
— R. Breitling, (2010)

In practice systems biology is largely concerned with constructing models which integrate this mechanistic knowledge into dynamic networks, yielding a system ready for quantitative analysis. Validation of these models is the observation that the reconstructed system behaves, as a whole, like the real biology with a given set of inputs. By exploring these models through individual input perturbations, the output is the emergent and predictive response of a whole system. Mathematical models have been used for decades to reconstruct biological systems and their wiring to predict novel outputs. A famous example are the models by Hodgkin and Huxley (1952) were used to examine the system of the giant squid axon and are widely recognised as early important work towards systems biology. However, applying this approach to larger, more complex systems requires not only a greater understanding of the circuitry or components, but also the ability to manage and compute their dynamic interactions. In that sense, it is easy to see how this approach is only becoming realised only now. It is a synergistic culmination of computational power for storage and calculation, combined with the increasing resolution of experimental technologies which builds upon an already vast mountain of humankind's knowledge.

Many curated examples of systems biology models can be explored and simulated at leisure from BioModels Database [3].

The allure of systems biology in the context of drug discovery becomes obvious. Perturbations of an accurate computational counterpart to real-life biological systems could be analogous to drug activity, an understanding of which may ease the selection of drug-targets; repairing the radio becomes trivial when you have the blueprints. Fur-

---

1 With responses such as: a), "to me, all biology is systems biology";
   b), "it's just one of those hypes that you need to follow to get grant money";
   c), "this holistic approach is the future of biology, we all will have to do it"
   (Breitling, 2010).
2 A recommended and entertaining read for those unfamiliar with the analogy.
3 Access to BioModels can be found at: www.ebi.ac.uk/biomodels

thermore, successfully creating *in silico* biological systems reduces the need for real-world experiments which are both costly and time consuming.

1.6.2 *Systems pharmacology*

The field of systems pharmacology has emerged recently and aims to transition the more phenomenological *in silico* modelling methods in pharmacology such as PKPD modelling towards the mechanistic views embedded in systems biology. Hence, systems pharmacology concerns itself with the quantitative analysis of biological networks and can be defined using systems biology. However, the key difference between the fields is that systems pharmacology is solely focused upon examining these network perturbations to explore, validate and predict drug action.

A comprehensive working definition was provided by Sorger et al. (2011) in the National Institutes of Health white paper which remains fully descriptive of the subject.

"Quantitative and Systems Pharmacology" is focused on "identifying and validating drug targets, understanding existing therapeutics and discovering new ones". Furthermore, it is discussed that "(QSP) aims to develop formal mathematical and computational models that incorporate data at several temporal and spatial scales" (Sorger et al., 2011). Systems pharmacology is also typically described as the interface between systems biology and pharmacodynamics (Graaf and Benson, 2011; Mager and Kimko, 2016).

However, where the definition in the 2011 white paper (Sorger et al., 2011) refers to "Quantitative and Systems Pharmacology" I make one non-trivial amendment; this definition refers now only to systems pharmacology and not necessarily quantitative systems pharmacology (QSP). Instead, I would argue that QSP is synonymous with systems pharmacology modelling but not systems pharmacology. This distinction can be explained.

Observation of the past seven years has shown that the use of the term QSP is starting to noticeably diverge from "systems pharmacology". Where the term systems pharmacology continues to be applied in parallel with systems biology it refers to other forms of network analyses as well as dynamical computational modelling. QSP however has become solely synonymous with the pharmacology equivalent of systems biology models. In other words QSP (systems pharmacology models) is a subset of systems pharmacology. Studies which use "QSP" approaches are almost all in this category ((Geerts et al., 2018; Kosinsky et al., 2018; Leil and Bertz, 2014; Pichardo-Almarza and Diaz-Zuccarini, 2017b) to name a few), in contrast with "systems pharma-

cology" (for example: Berger and Iyengar (2009), Jalali et al. (2018) and Zhao and Iyengar (2012)).

It is important to make this distinction as a number of articles use the definitions interchangeably with potential confusion for readers. For example, Musante et al. (2016) use a definition of systems pharmacology, given by Graaf and Benson (2011), to describe QSP. If this were the definition of systems pharmacology then research describing other forms of network analysis are defined incorrectly (such as the study by Jalali et al. (2018)). Another example, applied to logic models is seen in research by Traynard et al. (2017) entitled *Logic Modeling in Quantitative Systems Pharmacology*. In this article, QSP is introduced briefly, but only systems pharmacology appears in the discussion. In this case the subject matter is of course both QSP and systems pharmacology. However, not all systems pharmacology research is QSP.

Are QSP and systems pharmacology the same?

The distinction may be implicit; the first text book on systems pharmacology occasionally uses QSP instead of systems pharmacology to distinguish between dynamic modelling and other forms of network analyses (Mager and Kimko, 2016). Furthermore, in their book, systems pharmacology was only ever defined by alluding to "Quantitative and Systems Pharmacology". Even in my previous work in 2016 "QSP" was used to separate modelling from the overall concept of systems pharmacology without explicit discussion (Knight-Schrijver et al., 2016).

Perhaps there is an unspoken rule in place...

This is not describing a deep-rooted problem or intentional misunderstanding on any part. The use of the term QSP is interesting because it may directly stem from merging of two separate terms of "quantitative pharmacology" and "systems pharmacology" (Allerheiligen, 2010). Furthermore, the phrase "quantitative and systems pharmacology" was used many times in the white paper and yet, "quantitative systems pharmacology" was not mentioned once in the 47 pages despite its use of the acronym *QSP* (Sorger et al., 2011). However, the term did appear in an executive summary of the first QSP workshop in 2008 (Lauffenburger, 2008).

Furthermore it is apparent that a whole sub-field of systems pharmacology slips under the radar, which explicitly discusses traditional chinese medicine using network analyses[4]. Recent proposals for the working definition include the use of QSP as an integrative framework for context, as opposed to more of the same modelling (Androulakis, 2016). However, the unspoken disparity creates an ambiguity that makes it difficult for the reader to discern whether a proposed frame-

---

4 A query of "systems pharmacology"[Title] AND ("herbal"[Title] OR "chinese"[Title] OR "TCM"[Title]) in Pubmed returned 14 hits before 2017. Zero of these are mentioned in *Systems Pharmacology and Pharmacodynamics* (Mager and Kimko, 2016)

work is a computational model of dynamic systems or whether the framework includes computational models as but one of its tools. A final point of interest is that quantitative systems biology has failed to become a substantial term in scientific literature. How did quantitative come to prepend systems pharmacology and not biology for describing the same subset of concepts?

Pedantics aside, the debate I raise is that the definition is unclear; there is an appreciable level of ambiguity in the subject matter discussed in this thesis. Furthermore, I would hazard a guess that more than a few pharmacologists including I have paused for thought over which of the two terms to use in context. My proposal is clear; that QSP is synonymous with systems pharmacology modelling and remains a subset of systems pharmacology which is itself the overall concept described within the NIH white paper (Sorger et al., 2011). After all, it was a working definition.

TOWARDS A LANDSCAPE OF SYSTEMS PHARMACOLOGY MODELS.

2.1 INTRODUCTION

To define the landscape of systems pharmacology models it is necessary to examine the range of published articles describing QSP and systems pharmacology approaches and assess the spread of focus and areas of biological interest. As a field, systems pharmacology has been slowly growing since conception. The first mention of *systems pharmacology* in PubMed was in 2004 and the number of papers which explicitly contain systems pharmacology in clearly visible text fields has risen since (Knight-Schrijver et al., 2016) (Figure 2.1). These pub-
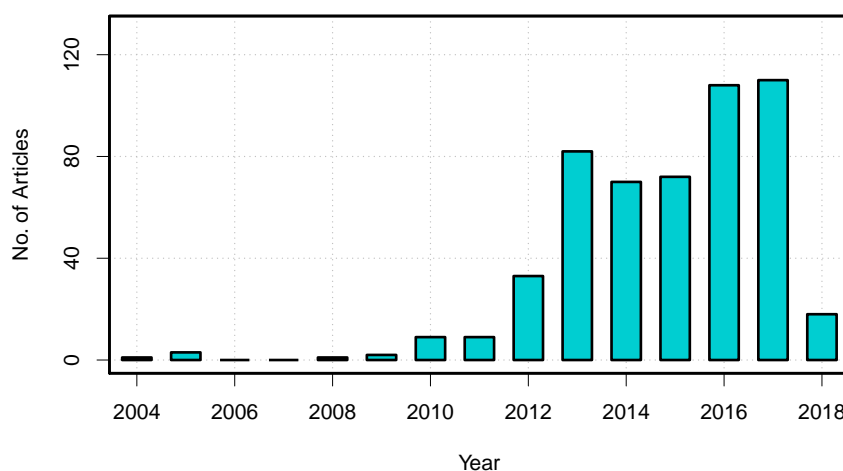
Figure 2.1: Systems Pharmacology in PubMed. The unique published occurrences of *systems pharmacology* in PubMed from 2004[1].

lished accounts of systems pharmacology models describe a whole plethora of disease mechanisms and pharmacological targets. However, it is difficult to properly examine where these models focus or to infer how effective these models are in addressing clinically relevant problems. One reason is that there is a degree of confusion as to how to interpret systems pharmacology or QSP which hinders a proper examination of literature within the modelling landscape. Similar to the rise in popularity of systems biology (Hübner et al., 2011), the definition of systems pharmacology may become obfuscated to those working within the field as the term has been used over a range of *in vitro*,

---

[1] Retrieved with a query from PubMed in February 2018: "systems pharmacology"[Title/Abstract].

statistical, and network analyses when a non-reductionist perspective is applied. Although these are approaches often seen in rational drug discovery, when systems pharmacology models are discussed, we are instead referring to the computational models akin to those in systems biology: mathematical networks which are developed which attempt to mimic real biological systems. Although this distinction is now becoming defined by the term quantitative systems pharmacology (QSP), the actual definition of systems pharmacology and QSP may still be poorly understood by a large proportion of the pharmaceutical industry (Nijsen et al., 2018). Further confounds to identifying relevant articles for analysis manifest themselves as reviews of the field or methods for the application of QSP. While they are certainly of interest to the field as a whole, these are generally not considered primary model descriptions.

In this project, text mining (TM) approaches were used to classify, extract and analyse text information in order to define the modelling landscape of QSP. In doing so, the range of focus upon disease is discussed alongside the variety of biological processes and pharmaceutical targets.

2.1.1 *Text mining*

The general explosion in data over recent years, fuelled by more efficient data generation and storage technologies, drives the need for powerful methods of data processing and extraction (Fan et al., 2014). For the analysis of large-scale text data, TM is used which specifically focuses on the parsing, processing and analysis of large corpora of texts. Typical applications of TM involve generating structured data from unstructured textual datasets to analyse trends and patterns. This is used to create population-level insights from readily accessible literature sources. For instance, the wealth of scientific literature in biology provides ample resource for studying and generating new hypotheses (Bekhuis, 2006; Swanson, 1990). TM has seen a number of uses in the biological sciences and the pharmaceutical industry. The application of TM to sift through clinical trial reports may be able to effectively reduce the workload of forming systematic reviews (O'Mara-Eves et al., 2015); at least 75 reports are being generated every day on average which necessitates a semi-automated approach of report handling (Bastian et al., 2010). TM has also been used to generate valid novel hypotheses over numerous biological interactions (Swanson, 1990). As part of rational drug-discovery, knowledge generation through TM can assist in the context of network construction (Saric et al., 2005), aiding target selection and drug re-purposing within a network pharmacology paradigm (Defteros et al., 2011; Hopkins, 2008; Tari and Patel, 2014). For example, novel drug indications can

Corpora is the plural form of corpus, a collection of documents that are usually themed.

be suggested from available literature databases such as Medline. In one case, a clinical trial for treating chronic hepatitis C with thalidomide was published a few years after a TM study suggested its potential (Milazzo et al., 2006; Weeber, 2003). However in this example, more recent evidence suggested a possible reactivation of the hepatitis c virus in patients receiving concurrent thalidomide as therapy for multiple myeloma (Mahale et al., 2015).

The need for TM here in building a picture of systems pharmacology models is evident. We can retrieve models for analysis using a simple search over an indexed database such as PubMed. In doing so however, several factors need to be considered. Firstly, the size of the data set is large. PubMed is one of the largest medical literature resources with over 28 million indexed articles; there is a vast swathe of literature to sift through to exhaustively capture the landscape of QSP and systems pharmacology. Secondly, large-scale datasets are noisy; even a simple search query using only the key defining phrase *systems pharmacology* or *model* returns many false-positives which may influence conclusions. For instance, a review article or a methodological paper might bias the result or indeed any other text that mentions a model. Finally, the ability to retrieve all relevant texts is limited using simple queries as relevant research and models undoubtedly exist which were published prior to or independent of the emergence of QSP. These studies (in content, concept and methods) would inherently be classed as QSP but are missing key words used in the current climate of QSP or systems pharmacology. Therefore, it is necessary to employ TM methods to define the evolved landscape of systems pharmacology models independent of the coined name, maximising the recall as well as the precision.

2.1.1.1 *Natural language processing*

Retrieving information from text is the basis for most text mining applications and begins by translating the text into useful data. Natural Language Processing NLP is used to split text into components and in doing so, the document is rendered accessible to interpretation by statistical methods, machine learning approaches and other data analysis techniques.

TEXT PARSING AND TOKENISATION   The conversion of structured or unstructured text into useful data often begins by parsing the text. With a few exceptions, words in written English can usually be segmentated by whitespace " " and sentences can be parsed using a full stop "." as a delimiter (the evolution from example 2.1 to example 2.2). Segmentation or tokenisation produces an ordered list of character strings ready for interpretation which can consist of single (unigrams) or multiple words (n-grams). When the order or distance

between tokens within each document is not considered important for the question, a *bag of words* model is used where the position of each token is independent of one another. Inference over this model assumes that the mere co-occurrence of words within the document provides enough evidence to draw satisfactory conclusions and can often be used for text classification tasks.

$$QSP \ increases \ drug \ development \ efficiency. \tag{2.1}$$

$$\underbrace{\overbrace{[QSP]}^{\text{word token}} [\ ][increases][\ ][drug][\ ][development][\ ][efficiency]}_{\text{sentence}}[.] \tag{2.2}$$

SEMANTICS AND LINGUISTIC ANNOTATION   In addition to tokenisation, semantic concepts can be applied to ordered sets of words so as to minimise misinterpretations resulting from polysemy or synonymy. Annotation of domain-specific text is often carried out using standard ontologies or vocabulary databases such as the Unified Medical Language System (UMLS) Metathesaurus and medical subject headings (MeSH®) (Demner-Fushman et al., 2010; Lipscomb, 2000). In a simple example (2.3), the predicate *increases* can be interpreted as one synonym for the concept of *positive regulation*. Likewise, the word *drug* is synonymous for several other words which could fall under the concept of *pharmaceutical* or other suitable definition. A simplistic example of this, is the interpretation of 2.3 as entity-relationship-entity, implying positive interaction between *QSP* and *drug development*.

$$\underset{\text{named entity}}{[QSP]} \quad \overset{\text{positive regulation}}{[increases]} \quad \overset{\text{named entity}}{(\underset{\substack{\text{compound} \\ \text{drug} \\ \text{medicine}}}{[drug]} [development])} \quad \underset{\text{positive attribute}}{[efficiency]} \tag{2.3}$$

In retaining token positions, the generation of such entity-relationship-entity "tuples" by annotating tokens with semantics is useful for higher resolution analysis. Extraction of directed relationships between known and unknown entities can be key in network construction tasks like phosphorylation interaction networks (Tudor et al., 2015), gene regulatory networks (Song and Chen, 2009) or drug interaction networks (Iyer et al., 2014). Using deeper levels of linguistic annotation and sentence parsing, natural language processing (NLP) is also applied to study more complex relationships like biological pathways with

For avid readers, an interactive demonstration of syntactic sentence parsing and parsing trees can be seen with Enju².

multiple entities. Examples can be seen in efforts such as EventMine or the Big Mechanism programme which aims to extract large mechanistic networks from the wealth of literature (Cohen, 2015; MIWA et al., 2010). Extended discussion on state-of-the-art semantic annotation is non-trivial and pushes outside the scope of research. Readers are advised to see dedicated reviews, tools and example work flows for further information (Gonzalez et al., 2015; Jensen et al., 2006; Sernadela and Oliveira, 2017). In the project here, I only consider the annotation of entity synonyms.

2.1.2 *Document classification*

Text classification has a variety of real world applications such as email filtering, plagiarism or malware detection (Alsmadi and Alhami, 2015; Oprisa et al., 2013). Document classification involves calculating the similarity of novel documents to a set of predefined categories generated in training. This means that classifiers use data extracted from texts such as tokens, structure or other content generated through NLP. Classifiers can be supervised or unsupervised. A supervised classifier is given document class labels as well as the input data and are trained by building associations between the data and the class. Examples of supervised classifiers for text categorisation include the naïve Bayes classifier (NBC), K-nearest neighbour (KNN) or support vector machine (SVM) methods (Joachims, 1998; McCallum and Nigam, 1998; Trstenjak et al., 2014). Performance between these methods has been evaluated in many studies and the accuracy is comparable (Yang and Liu, 1999; Yu, 2008).

An unsupervised classifier on the other hand uses only the input text data to generate output classes from the observations. Because no class labels are given to be trained against, unsupervised algorithms are designed to quantify differences between groups of observations, separating groups where large differences occur. This results in the clustering of similar observations. Unsupervised classifiers are generally clustering algorithms such as K-Means Clustering (Singh et al., 2011). One example of K-means clustering will randomly assign k observation values as class means. Class clusters are then created by categorising the remaining observations according to the closest mean. The class means are then recalculated by computing the class centroids and the exercise is repeated until convergence. Because they are unbiased, unsupervised classification methods can also be used in feature selection for supervised methods and have been shown to build highly decisive feature subsets for text classification tasks (Zhou et al., 2014). While these methods typically use a bag of words model, more recent approaches successfully use ordered representations of sentence and paragraphs such distributed memory models (Baker et

al., 2016; Le and Mikolov, 2014). Document classification in this project is performed using the NBC. The NBC is easy to implement within a work flow of feature selection and works "surprisingly well" under the general assumptions of term independence. Furthermore, the NBC is relatively easy on computation and copes well with handling highly-dimensional data.

2.1.2.1 *Naïve bayes classifier*

The NBC traditionally comes in two flavours, the Bernoulli model and the multinomial model (McCallum and Nigam, 1998). Both classifiers utilise Bayes' theorem of conditional probability and assumptions of independence between observed events. Here the multinomial form is used which accounts for word frequency between articles and is suitable for the text length of research abstracts. Another form of the NBC is the Bernoulli model which only takes into account binary appearances of each term within a document.

BAYES' RULE IN DOCUMENT CLASSIFICATION   A typical representation of Bayes' rule is seen in equation 2.4. The probability that C occurs, given that D occurred (denoted by $P(C|D)$), is equal to the prior probability of C occurring, $P(C)$, multiplied by the posterior probability of the event D given that C has also occured, $P(D|C)$. This is divided by the probability of D ($P(D)$).

$$P(C|D) = \frac{P(D|C) \cdot P(C)}{P(D)} \tag{2.4}$$

We can interpret this with respect to document classification where class is C and the features of a document is D, our evidence. As $P(D)$ is the probability of drawing the evidence in posterior, it can be practically omitted for classification purposes as the evidence is constant in all class decisions. The probability then of C occurring given the observation of D, is proportional to product sum of prior and posterior probabilities (Equation 2.5.).

$$P(C|D) \propto P(D|C) \cdot P(C) \tag{2.5}$$

If instead, D is a vector of character strings such as words or tokens of length n, $D = (w_1, w_2, ...w_n)$, the equation can be re-written to account for the likelihood of all evidence occurring.

$$P(C|(w_1, w_2, ..., w_n)) \propto P((w_1, w_2, ..., w_n)|C) \cdot P(C) \tag{2.6}$$

MULTINOMIAL DISTRIBUTION   Suppose that you have three large bags of words and that someone has dipped a small container into

one, retrieving a portion of words at random. To which bag does the vessel of words belong? Suppose that instead of three bags, you have now a room full of bags, of which there are three types or classes. Calculating the probability that your container was dipped into either one of these bag types is the premise of the multinomial naïve bayes classifier (mNBC) in text classification. The multinomial distribution is used to describe the probability of selecting a sequence of independent events. Using the bag-of-words model as an example, a sequence of words of length n is drawn with replacement from a bag of words with the unique vocabulary $|V|$. The words $w_1, w_2, ..., w_t$ in $|V|$ occur at different frequencies within the bag as represented as fractions (p_t) of the total number of words.

$$\sum_{t=1}^{|V|} p_t = 1 \quad p_t > 0 \tag{2.7}$$

When a document is represented by a vector of w_t counts, $d = (x_1, x_2, ..., x_{|V|})$, each word can be modelled as a x_t successive draws given by the probability of w_t: p_t in $p_1, p_2, ..., p_{|V|}$. The document, as a sequence of words, is modelled by a multinomial distribution and parametrised by the number of words n and the probability of each word in the bag p_t (Equation 2.8.).

$$P(d) = \frac{n!}{\prod_{t=1}^{|V|} x_t!} \cdot \prod_{t=1}^{|V|} p_t^{x_t} \tag{2.8}$$

For each class in C then, p_t is substituted for the conditional probability of w_t given C (Equation 2.9.). For simplicity, the multinomial coefficient is discarded from equation 2.8 as the number of sequence combinations is constant between classes for a given document.

$$P(C|d) \propto P(C) \cdot \prod_{t=1}^{|V|} p(w_t|C)^{x_{it}} \tag{2.9}$$

In forming a prediction, the document class is then chosen by a *maximum a posteriori* (MAP) decision and the least improbable category is selected to classify the document. The core structure of the algorithm used in this research is given in Listing 2.1.

Listing 2.1: A multinomial Naïve Bayes Classifier[4].

```
model(C,D)
  1  V <- ExtractFeatures(D)
  2  N <- CountDocuments(D)
  3  for each class c in C
  4    N_c <- CountDocumentsInClass(D,c)
  5    priors[c] <- N_c / N
  6    text_c <- tokeniseTextOfClass(D,c)
  7    text_c <- text_c ∈ V
```

```
8      for each term t in V
9        T_ct <- countTokensOfTerm(text_c,t)
10       condprob[c,t] <- (T_ct+1)/(length(text_c)+length(|V|))
11     return(V, priors, condprob)
```

```
predict(C,V, priors, condprob,d)
1    W <- FeaturesInModel(V,d)
2    for each |class| c in C
3      score[c] <- log(priors[c])
4        for each t in W
5          score[c] <- score[c] + log(condprob[c,t])
6    return(max(score))
```

2.1.2.2 *High-dimensionality and dimension reduction*

One challenge of TM is the abundance of features formed from the terms within the corpus. The NBC performs well with high-dimension data sets but like other classifiers, an abundance of bad predictors can lead to a low accuracy and general poor efficiency. Therefore, choosing a set of highly predictive word tokens is an essential process in text classification.

Common approaches for reducing feature dimensionality are wrapper or filter methods (Kohavi and John, 1997; Sánchez-Maroño et al., 2007). Between the two, filter methods are less computationally intensive. Wrapper methods are machine learning algorithms that select subsets of features for evaluation, often used after a filter method. One example of algorithms used for wrapper-based feature selection is the random forest algorithm (Breiman, 2001; Vora and Yang, 2017). Filter methods are implemented by indexing features by their value as predictors. A common and simple filtering approach is to rank features by the term frequency within the corpus. In this case, high frequency terms may be noisy predictors, offering little discriminatory information due to their ubiquity. More complex methods of filter approaches to dimension reduction can be seen with calculating information gain, which quantifies the information that a feature provides about a class (Dasgupta et al., 2007), or methods derived from calculating chi-square statistics (Jin et al., 2015). The results of these approaches are often seen to be correlated and their accuracy is still comparable to that of simple methods such as document frequency when using large enough feature sets (Forman, 2003; Yang and Pedersen, 1997). A high frequency term may provide a great discriminatory power as its between-class distribution may be more well-defined. Based on this observation, non-parametric statistical tests

---

4 Model training and prediction algorithm psueo-code adapted from Manning et al. (2008).

such as a Kruskal-Wallis (Kruskal and Wallis, 1952) could also be used to determine the value of features as inter-class discriminants (Vora and Yang, 2017). Additionally when comparing word features, a heuristic weighting function called term frequency-inverse document frequency (TF-IDF) can be applied to terms within a corpus. This method of term weighting calculates the frequency of terms and acts to penalises those which occur frequently in documents. It is generally considered practical to weigh against words which have very little eliteness (Robertson, 2004). These filter methods consider that features are univariate and are independent which suits the assumptions in NBC models.

2.1.3 *Previous work, aims and goals*

A review by Hübner et al. (2011) in the field of biochemistry details trends and patterns in systems biology modelling. However, no previous studies attempt to capture the picture of systems pharmacology modelling to such a degree. Furthermore, the definitions of systems pharmacology are loose which makes a simple query ineffective. Therefore, in an effort to capture the essence of systems pharmacology modelling, we constructed a more complex TM query to recall relevant research abstracts from Medline dated from 1965 to 2015. The aims were to: 1) accurately retrieve a corpus of systems pharmacology models; and 2) extract entity information to examine the properties of the pharmacological modelling over the last 50 years covering diseases, chemicals, genes and other entities in order to generate a landscape of systems pharmacology models.

2.2 MATERIALS AND METHODS

This section describes the tools and software, the sources of data and the work flows used to define the QSP modelling landscape through TM.

DOCUMENT RETRIEVAL AND CLINICAL TRIALS DATA   Initial document retrieval, including the extraction of MeSH® terms, disease categories and clinical trials data was carried out using the NLP data mining program, I2E© version 4.2 (Linguamatics). Further analysis, generation of figures and classification of the retrieved literature abstracts was carried out in in R (R Core Team, 2016). The positive terms in the initial search query are seen in list 2.2. A list of negative terms such as *structural model* or *protein folding* was added to remove noise similar to the query design by Hübner et al. (2011). Of course in my research, terms such as *pharmacokinetic* or *pharmacodynamic* were not negative terms.

Figure 2.2: I2E search query. Positive terms in the I2E search query. * = word pluralisms; ^= case sensitive. These terms are an "OR" relationship.

```
Biological Modelling
  └── Computational Biology
       └── Computational Biology Models
            ├── computational* biology model*
            ├── Dynamic Model
            │    ├── dynamic* model*
            │    ├── Model* Dynamic*
            │    └── computational* biology model*
            ├── Mathematical Model
            │    ├── Mathematical Model*
            │    ├── mathematical models
            │    ├── Quantitative Model*
            │    ├── mathematical biology model*
            │    ├── Differential Equation Models
            │    │    ├── ODE* model*
            │    │    ├── differential equation*
            │    │    ├── ordinary differential equation* model*
            │    │    ├── partial differential equation* model*
            │    │    ├── Model* by* differential equations
            │    │    └── Model* by* ODE*
            │    └── Compartmental Models
            │         ├── multi-compartmental model*
            │         ├── compartmental model*
            │         ├── compartment* model*
            │         └── compartment model
            ├── Systems Models
            │    ├── systems model
            │    ├── agent-based model*
            │    ├── Systems Biology Models
            │    │    ├── computational system*
            │    │    ├── "systems biology model*"
            │    │    ├── theoretical* model*
            │    │    ├── theoretical model
            │    │    ├── model*, systems biology
            │    │    └── model* of systems biology
            │    ├── Systems Pharmacology Models
            │    │    ├── systems pharmacology model*
            │    │    ├── model* of systems pharmacology
            │    │    └── model*, systems pharmacology
            │    ├── Systems Medicine Models
            │    │    ├── systems medicine model*
            │    │    ├── model* of systems medicine
            │    │    └── model*, Systems Medicine
            │    └── Systems Physiology Models
            │         ├── systems physiology model*
            │         ├── model*, systems physiology
            │         └── model * of systems physiology
            ├── Kinetic Models
            │    ├── Kinetic* model*
            │    └── kinetic-metabolic model*
            └── Computer Simulation
                 ├── computer simulation*
                 ├── computational model*
                 └── mathematical simulation*
       ├── Modelling Programs
       │    ├── copasi
       │    ├── gepasi
       │    ├── simbiology
       │    ├── dbsolve
       │    ├── celldesigner
       │    ├── SAAM^II^
       │    ├── jdesigner
       │    ├── biopax
       │    ├── SBworkbench
       │    ├── xpp
       │    ├── xppaut
       │    ├── Matlab
       │    ├── Dizzy^
       │    └── biology toolbox
       ├── Modelling Languages
       │    ├── cellml
       │    ├── sbml
       │    ├── systems biology markup language
       │    └── MIRIAM^
       └── Computational Biology Databases
            ├── BioModel
            ├── BASIS^
            └── Biomodels Database
                 ├── path2models
                 └── biomodel ID
```

CORPORA   Abstract corpora used for classification were downloaded directly from PubMed in extensible markup language (XML). The annotated set of abstracts for model training and testing purposes were compiled before optimising the model by reading abstracts acquired from iterative basic queries. These were manually annotated into three classes. Abstracts were marked as either true-negative (TN), true-positive biology (TPB) or true-positive pharmacology (TPP) with respect to computational systems models. Additionally if abstracts were from articles describing models (TPP or TPB classes), an effort was also made to store and curate the models on the BioModels Database. The sizes of the corpora and the distribution of classes within the annotated corpus are seen in Table 2.1.

Corpus	Number of Documents
Query retrieved abstracts [1]	372,967
Annotated TN abstracts [1,2]	845
Annotated TPB abstracts [2,3]	767
Annotated TPP abstracts [2,3]	128

Sources: [1], retrieved using I2E query 8 in 2015 (without negations) from Medline (1965 - 2015); [2], manually annotated abstracts retrieved during query optimisation and dataset sampling; [3], sorted abstracts from models stored in the BioModels Database.

Table 2.1: The size of corpora used in this study.

TEXT PROCESSING   A number of different tasks form the work flow for text processing. To create data for a useful *bag-of-words* model, 6 steps were carried out:

1. Upper case characters were replaced with lowercase equivalents.

2. Punctuation characters were removed (with the exception of hyphenated word compounds).

3. Common English stop words were deleted (listing 2.2).

4. Words were stemmed using Porter's stemming algorithm with English (Porter, 1980).

5. Abstract strings were parsed by space (" ").

6. n-grams ranging between one and ten were generated over unigram sequences.

DATA NORMALISATION AND ANALYSIS   Data were normalised and weighted using TF-IDF. In this calculation of TF-IDF, the term fre-

quency TF_{td} is created for each term t per document d (Equation 2.10.).

$$TF_{td} = \frac{F_{td}}{\sum_{t' \in d}^{|t|} F_{td}} \qquad (2.10)$$

The inverse document frequency for a given token (IDF_{tD}) is then defined by taking the logarithm of the total number of documents N in the corpus D, divided by the number of documents with a non-zero count of the term t n_t (Equation 2.11.). A psuedocount is added to all terms.

$$IDF_{tD} = log\left(\frac{N}{1 + n_t}\right) \qquad (2.11)$$

Finally, TF-IDF is given as the product of both the TF_{td} and the IDF_{tD}.

STATISTICAL TESTS AND FEATURE REDUCTION   Supervised statistical analyses were carried out over the TF-IDF distributions. Non-parametric testing of one-criterion variance was carried out using the Kruskal–Wallis test (Kruskal and Wallis, 1952). This assumes that the class word distributions can be statistically modelled by an approximate χ^2 distribution with $|K| - 1$ degrees of freedom ($|K|$ being the number of unique classes). Feature reduction was carried out by filtering over the list of tokens ordered by the descending χ^2 and *post-hoc* pairwise comparisons were made using Conover's test (Conover, 1999) in cases where the Kruskal-Wallis null hypothesis was rejected.

MEASURES OF CLASSIFICATION   Standard measures of classification were used here. Recall of individual classes is observed using the standard formula (equation 2.12). The recall measures the fraction of a total class population that the classifier is able to correctly identify. The recall of a class K in a predicted set of documents D is given by.

$$Recall = \frac{K \cap D}{K} \qquad (2.12)$$

Likewise, precision is defined by the canonical equation 2.13 and details the extent to which the classifier is stringent and correctly classifies a document as class K within D, the set of documents predicted as K.

$$Precision = \frac{K \cap D}{D} \qquad (2.13)$$

Finally, the harmonic mean of both recall and precision is often used to report the effectiveness of classification (Rijsbergen, 1979) which is usually defined as the F_1 score (equation 2.14).

$$F_1 = \frac{2 \cdot Recall \cdot Precision}{(Recall + Precision)} \tag{2.14}$$

NAMED ENTITY RECOGNITION AND ANNOTATION   Extraction of annotated entity information was carried out using automatically annotated versions of abstracts from PUBtator accessed via the REStful api in R (Wei et al., 2013). The full list of annotations were mapped from lesser Online Mendelian Inheritance in Man (OMIM$^{©}$, https://omim.org/) and MeSH$^{®}$ (2018) terms onto MEDIC-slim terms (Davis et al., 2012). Disease synonym mappings were derived from the comparative toxicogenomics database (CTD) (Davis et al., 2016). Software entity extractions were carried out straight forward regular expression (regex) string matching commands. This was under the simple observation that an ordered string of words containing the term *software* would likely have the name of the software one word adjacent to it.

Listing 2.2: Common English stopwords

```
a, about, above, across, after, again, against, all, almost, alone, along,already, also, although,

always, am, among, an, and, another, any, anybody,anyone, anything, anywhere, are, area, areas, aren't,

around, as, ask, asked,asking, asks, at, away, b, back, backed, backing, backs, be, became,

because,become, becomes, been, before, began, behind, being, beings, below, best, better,between, big,

both, but, by, c, came, can, cannot, can't, case, cases, certain,certainly, clear, clearly, come,

could, couldn't, d, did, didn't, differ,different, differently, do, does, doesn't, doing, done, don't,

down, downed,downing, downs, during, e, each, early, either, end, ended, ending, ends,enough, even,

evenly, ever, every, everybody, everyone, everything, everywhere,f, face, faces, fact, facts, far,

felt, few, find, finds, first, for, four,from, full, fully, further, furthered, furthering, furthers,

g, gave, general,generally, get, gets, give, given, gives, go, going, good, goods, got, great,greater,

greatest, group, grouped, grouping, groups, h, had, hadn't, has, hasn't,have, haven't, having, he,

he'd, he'll, her, here, here's, hers, herself, he's,high, higher, highest, him, himself, his, how,

however, how's, i, i'd, if, i'll,i'm, important, in, interest, interested, interesting, interests,

into, is,isn't, it, its, it's, itself, i've, j, just, k, keep, keeps, kind, knew, know,known, knows, l,

large, largely, last, later, latest, least, less, let, lets,let's, like, likely, long, longer, longest,

m, made, make, making, man, many,may, me, member, members, men, might, more, most, mostly, mr, mrs,

much, must,mustn't, my, myself, n, necessary, need, needed, needing, needs, never, new,newer, newest,

next, no, nobody, non, noone, nor, not, nothing, now, nowhere,number, numbers, o, of, off, often, old,

older, oldest, on, once, one, only,open, opened, opening, opens, or, order, ordered, ordering, orders,

other,others, ought, our, ours, ourselves, out, over, own, p, part, parted, parting,parts, per,

perhaps, place, places, point, pointed, pointing, points, possible,present, presented, presenting,

presents, problem, problems, put, puts, q,quite, r, rather, really, right, room, rooms, s, said, same,

saw, say, says,second, seconds, see, seem, seemed, seeming, seems, sees, several, shall,shan't, she,

she'd, she'll, she's, should, shouldn't, show, showed, showing,shows, side, sides, since, small,

smaller, smallest, so, some, somebody,someone, something, somewhere, state, states, still, such, sure,

t, take,taken, than, that, that's, the, their, theirs, them, themselves, then, there,therefore,

there's, these, they, they'd, they'll, they're, they've, thing,things, think, thinks, this, those,
```

```
though, thought, thoughts, three, through,thus, to, today, together, too, took, toward, turn, turned,

turning, turns,two, u, under, until, up, upon, us, use, used, uses, v, very, w, want,wanted, wanting,

wants, was, wasn't, way, ways, we, we'd, well, we'll, wells,went, were, we're, weren't, we've, what,

what's, when, when's, where, where's,whether, which, while, who, whole, whom, who's, whose, why, why's,

will, with,within, without, won't, work, worked, working, works, would, wouldn't, x, y,year, years,

yes, yet, you, you'd, you'll, young, younger, youngest, your,you're, yours, yourself, yourselves,

you've, z
```

2.3   RESULTS

This section documents the output of an initial model retrieval query and subsequent classifier optimisation followed by the analysis of extracted information. The initial query was formed as part of earlier work throughout my research and was published in 2016 (Knight-Schrijver et al., 2016). The account of this work here illustrates the query terms used in the initial query as well as some of the noise mitigation steps taken while constructing it. The classifier was constructed after this research and remains an account of the extra efforts taken to reduce the set of articles using an effective text classification model. Finally, the results of abstract entity annotation are reported which reveals population-level content of the predicted systems pharmacology models.

2.3.1   *Corpus retrieval*

An indexed and annotated corpus of Medline abstracts was made available through I2E©, a proprietary NLP data mining software. Using this corpus, an initial search query was constructed and optimised in an iterative fashion to maximise the recall of a positive dataset, BioModels Database (Le Novère et al., 2006). Key descriptors of computational modelling within biology and pharmacology were included 2.2.

2.3.1.1   *Building a suitable query*

Key terms used to retrieve modelling articles were generated subjectively based upon the general terms expected to be discussed within modelling research. For example, "*computational biology*", "*ordinary differential equation*" and "*mathematical model*". Improvements to the query were made by examining the output, identifying both relevant and irrelevant articles, and adding the key terms within them to the next query. These incremental modifications to the query structure and content resulted in linear increases of recall as measured using the positive control data at the time (Figure 2.3). However, the rise in recall resulted in greater gains in the total number of retrieved articles. Without drastic measures of noise control, the relationship appeared to be exponential. Recall of positive articles was the only metric avail-
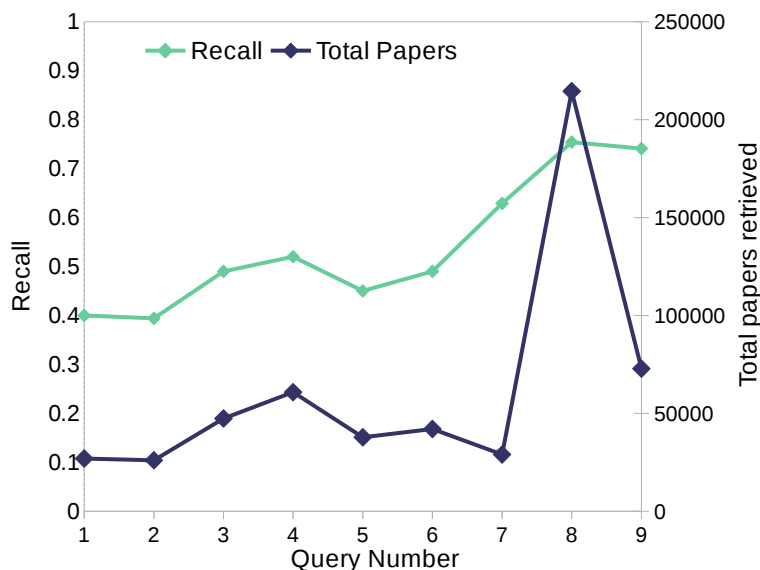
Figure 2.3: Initial query optimisation. After each query, the output was examined and obvious terms from both relevant and irrelevant abstracts were added to the query. This increased the recall of the positive control data.

able for evaluating the queries as the distribution of positive abstracts within the population of PubMed articles remained unknown. Evaluation and update of the query were carried out by assessing the output at each stage and manually identifying features which contribute to type one document retrieval errors. One effort to increase the precision and mitigate against false positive abstracts used MeSH® terms. We extracted all MeSH® terms from our annotated datasets and compared them with the MeSH® terms within our total population of retrieved abstracts (Figure 2.4). Within I2E©, the relative complement C \ B, totalling 2 454 MeSH® terms, was used as a negations list for query 9. Positive document MeSH® terms within C were shared also shared by B. This intersection of 828 MeSH® terms between the two class populations was retained in addition to the complement of B ∪ C. This ensures that potentially positive articles with both non-novel and novel MeSH® enrichment are still retrieved. Substantial noise was eliminated from the query by including this *MeSH®-fingerprint*. By actively filtering against these MeSH® terms, a three-fold reduction in the total number of retrieved articles was seen while maintaining the recall of the true positive data set. The effect is be shown in the difference between queries eight and nine in Figure 2.3. However, sampling the results of both MeSH® filtered and unfiltered retrievals suggested a bias towards BioModels-like abstracts, with a loss in the number of novel systems pharmacology models.
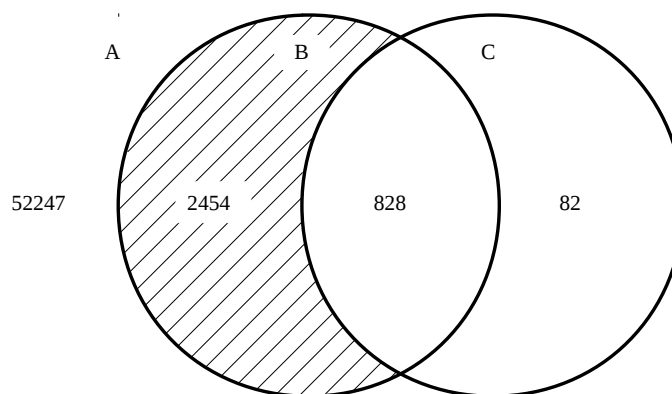
Figure 2.4: **A**, MeSH® terms not associated with either manually curated systems biology or true-negative articles; **B**, MeSH® terms associated with manually curated true-negative articles; and **C**, MeSH® terms associated with manually curated systems biology articles. The total population of MeSH® 2014 terms was 55 611 given by $(A \cup B \cup C)$.

2.3.1.2 *Comparing the focus of clinical trials and modelling*

The abstracts retrieved from query eight, clinical trial studies and the positive control of systems biology model abstracts were then passed back through I2E© to extract named disease entities. Using this approach the disease coverage of proposed systems models were compared with clinical trials and the BioModels Database as a positive control.

The findings tenuously suggested both areas of overlap and disjoint between modelling and clinical agendas covering the last 50 years (Knight-Schrijver et al., 2016) (Figure 2.5). For example, *Neoplasms* was the most enriched disease category in all datasets. However, less of an enrichment was seen in *Immune System Diseases*, *Respiratory Tract Diseases* and *Digestive System Diseases* comparing the query result to clinical trials.

The preliminary study assumed that the disease enrichment profile of the BioModels Database largely represented the focus of modelling in systems pharmacology. While this may or may not have been the case at the time, the precision was low, making any meaningful analysis difficult. It was unlikely that the retrieved abstracts were truly all published accounts of systems modelling let alone systems pharmacology. Samples suggested a maximal precision of 18 per cent for systems modelling articles but only five per cent for articles on pharmacometrics or systems pharmacology. Even using the *MeSH® fingerprint*, the query returned a substantial number of abstracts and may have biased the retrieval for BioModels-like articles. Furthermore, as the query made no serious effort to separate systems biology and
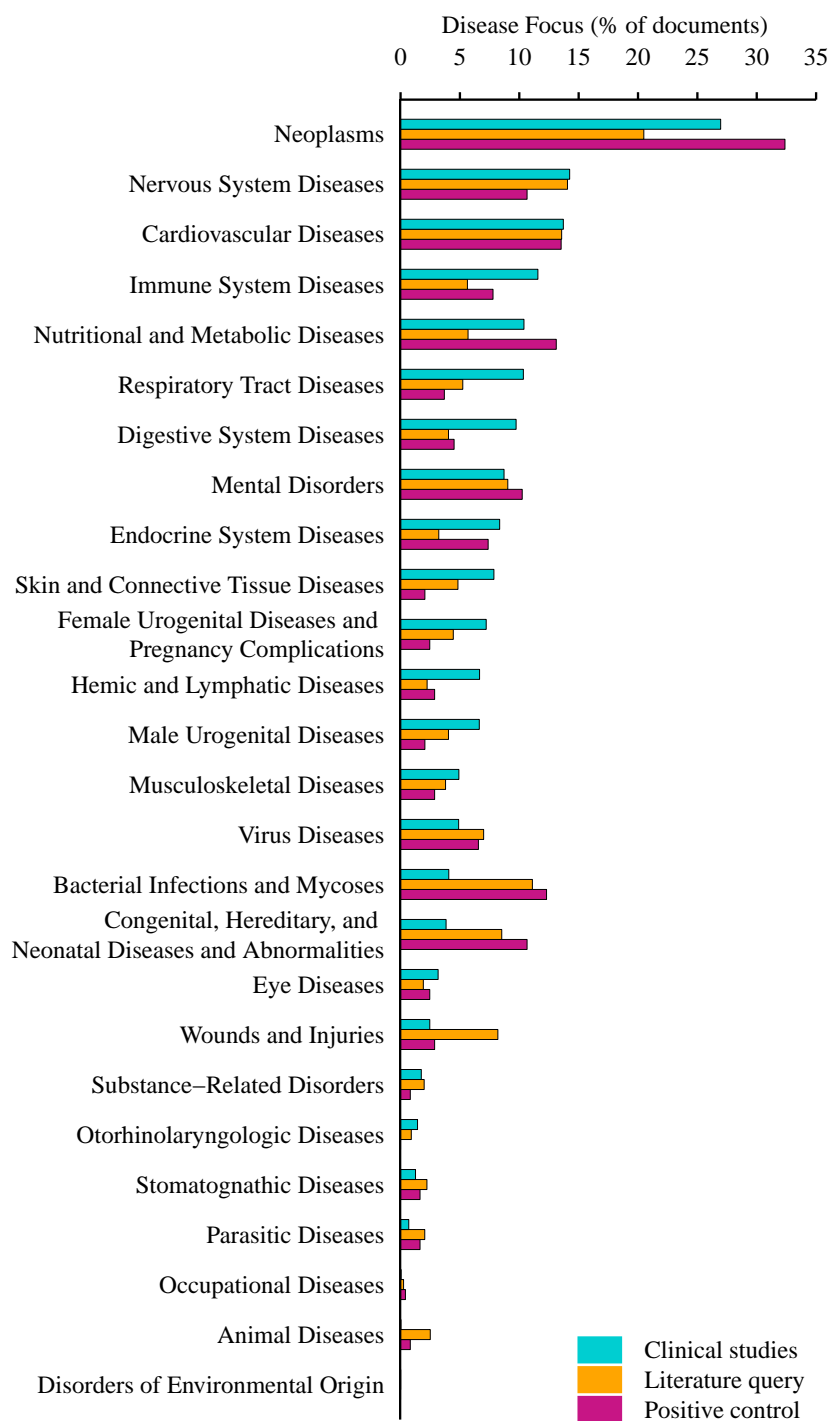
Figure 2.5: Disease category enrichment of texts in Clinical trials, BioModels Database and the I2E© query corpora covering the last 50 years. (taken from Knight-Schrijver et al. (2016)).

pharmacology models, the analysis also was not well-suited to study the landscape of systems pharmacology models alone.

For want of better precision then, the task was set about to first reduce the noise of the query results while maintaining the sensitivity in positively retrieving QSP and other systems modelling abstracts from PubMed.

2.3.2   *A naïve Bayes classifier for modelling texts*

The less-than-ideal recall and dubious precision warranted a more stringent method of document retrieval. The construction and optimisation of a mNBC for text categorisation was carried out to achieve a more accurate dataset for a more meaningful analysis. Here we consider three classes of document; true negative (TN), true positive biological models (TPB) and true positive pharmacological models (TPP). For this research we are most interested in extracting information from TPP-class abstracts as they are putative systems pharmacology models.

2.3.2.1   *Optimisation of the NBC*

The mNBC model is based upon the underlying distributions in the data and has no formal iterative learning steps. Therefore, it is best to optimise an mNBC by determining the most efficient feature subset which best discriminates between classes.

*Here, **gram** describes the number of words or tokens in a given string. I use the terms unigram, bigram and n-gram to refer to strings containing 1, 2 and n adjacent words.*

The number of words in a feature is considered important for classification. By generating n-grams, we can group multiple words together allowing the classifier to use features where words appear adjacent to each other in documents. For example, as we are classifying systems pharmacology models, the n-gram of length three, "systems pharmacology model", could be a useful predictive feature. To determine the added value of including n-grams in our feature set, a list of n-grams of lengths one to ten were generated across all training documents resulting in 1 777 337 features. Removing those that occur only once and selecting those that were significantly different between the three classes returned 3 959 n-grams of length one to six ($p < 0.05$, Kruskal–Wallis test). The performance of the mNBC was then calculated using feature sets made from different n-gram sizes in an attempt to optimise the classifier. We show that for our classification model, all feature subsets with a n-grams greater than one outperformed the unigram-only features. This can be seen as a mean increase in prediction accuracy of 2.2 per cent in test data between unigrams and bigrams ($p < 0.01$, one-way analysis of variance (ANOVA))(Figure 2.6).
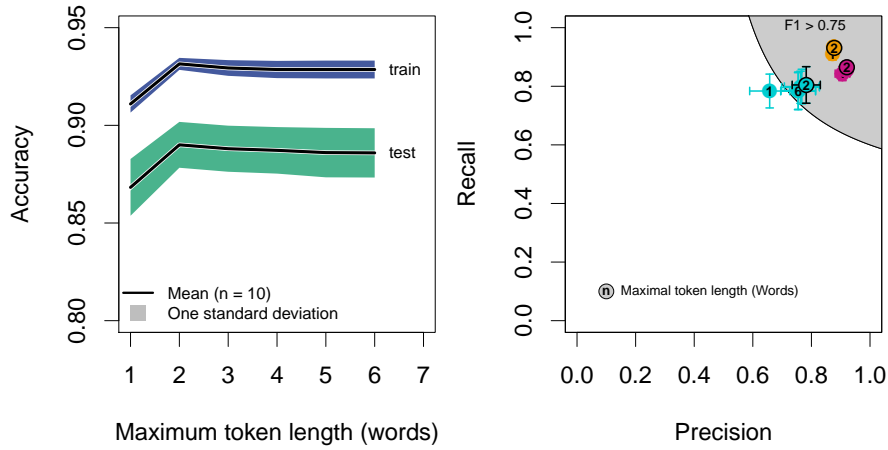
Figure 2.6: The overall accuracy of the n-gram variations shown as the fraction of correct predictions (left) or as a plot of Recall against Precision (right) of TN (magenta), TPB (orange) and TPP (cyan) predictions. For each n-gram, ten models were trained on a random sample of 60 per cent of abstracts and then used to predict the remaining 40 per cent for an unbiased measure. The abstracts were manually annotated before splitting (see Materials and Methods). Error bars are the standard deviation of these ten models.

Furthermore, significant increases in both Precision and F_1 but not Recall were seen in TPP classed abstracts when incorporating n-grams above 1. ($p < 0.001$, one-way ANOVA). The largest gain in classifier function was achieved by including bigrams as the Precision and F_1 scores of test TPP abstracts were enhanced by 12 and 8 percent respectively (Figure 2.7). Gains in recall were only modest in comparison and were only significant in TPB abstracts. The increase in Precision versus Recall was be seen in Figure 2.6 as the TPP class predictions crossed a 0.75 F_1 boundary. The highest performance was observed with a feature set containing only unigrams and bigrams. Predictive scores for TPP abstracts were lower than those for TN or TPB categories. The exception to this is the recall of training data (Figure 2.7).
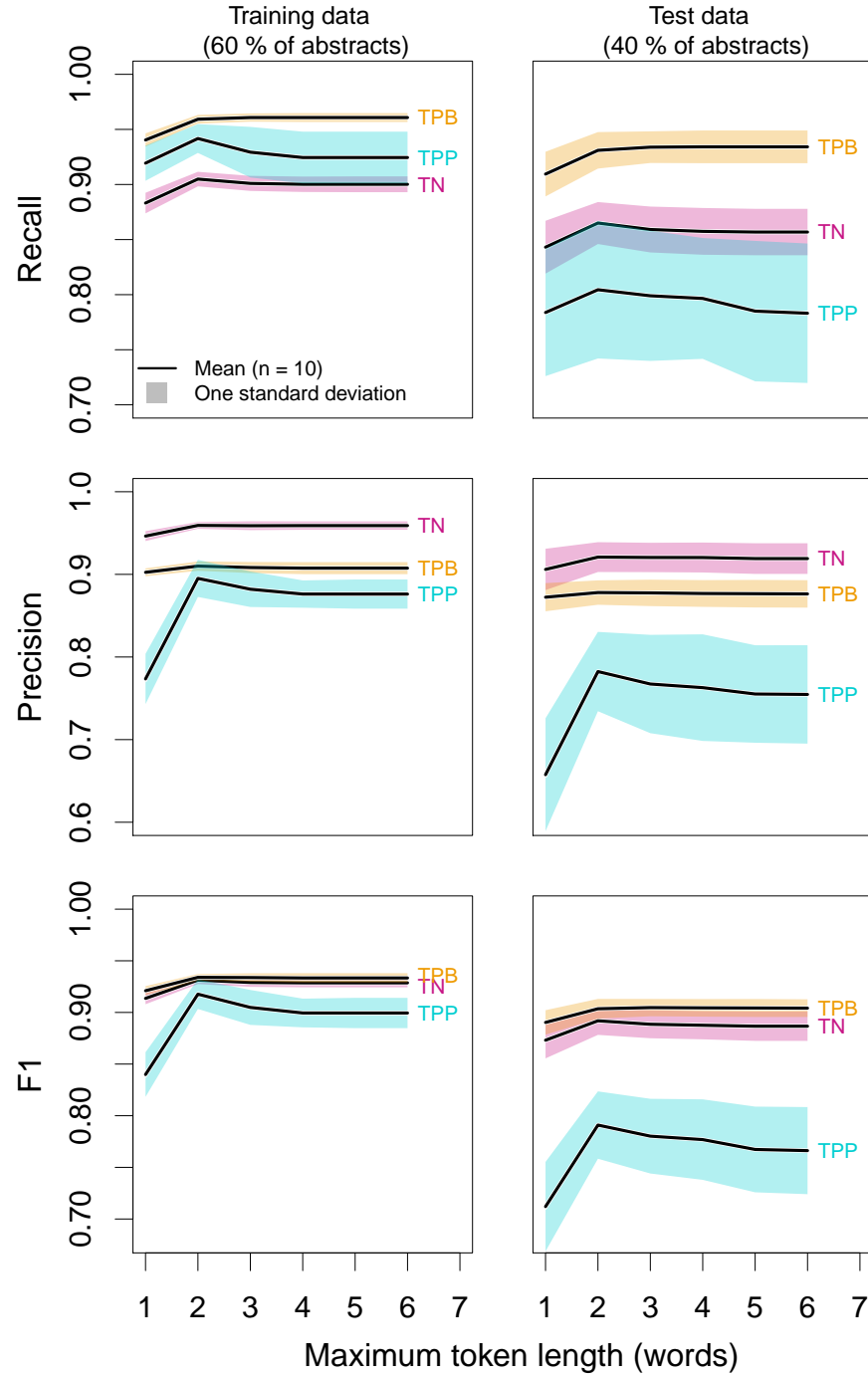
Figure 2.7: Mean Recall, Precision and F_1 scores for both training (left column) and test (right column) abstracts for n-grams of length 1 to 6. For each n-gram, ten models were trained on a random sample of 60 per cent of abstracts and then used to predict the remaining 40 per cent for an unbiased measure. The abstracts were manually annotated before splitting (see Materials and Methods). Error bars are the standard deviation of these ten models. F_1 scores are the harmonic mean between Recall and Precision (Equation 2.14).

To further define our feature set, a filter approach was used which covered the full list of unigram and bigram features. We can carry out a filtering approach by generating a feature weight metric and ordering the features by their descending score. This acts to filter out the weakest predictors by their position at the bottom of the ordered list. Firstly, features were removed that appeared in or mentioned annotated genes, diseases or chemicals this was to mitigate against bias in downstream analysis. The features were then ranked using Kruskal–Wallis χ^2 values obtained by testing between class TF-IDF distributions. The use of a non-parametric test was justified by testing for normality (Shapiro-Wilk test p values were reported between 4.7×10^{-34} and 8.8×10^{-8} across the classes). We generated optimisation curves for the classifier by descending through ranked feature space. For each set of features, three mNBC models were generated and tested over randomly sampled abstracts. The results show that a feature set consisting of the first 3 546 highest ranked features provides the best mean F_1 score (80 per cent) and precision (82 per cent) in test data with a recall of 78 per cent in TPP abstracts (Figure 2.8, Table 2.2).
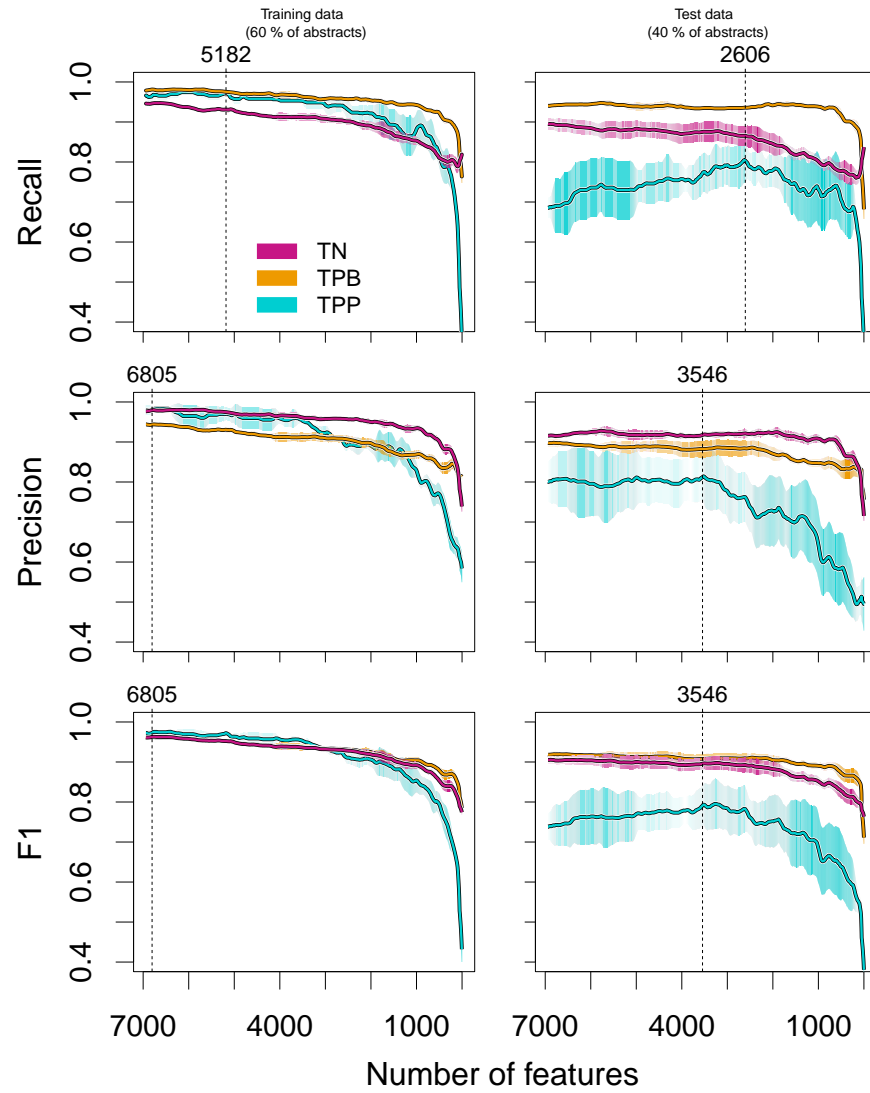
Figure 2.8: Feature reduction optimisation. Mean and standard deviation of classifier performance measures $n = 3$) across both training (left column) and test (right column) abstracts. For each feature set, the classifier was trained and tested three randomly assigned partitions of 60 per cent training and 40 per cent test abstracts. The abstracts were manually annotated before splitting (see Materials and Methods). The opacity and area are the standard deviation of these three models.

It is interesting to see that the model predicts training data well even after 6 800 features while the predictive power on test data was seen to decrease after approximately 3 500 features. This may be a result of model overfitting as more training specific and less generalised word features are used for the model. The word stem features with the most predictive power included "pharmacokinet", "system pharmacolog" "pharmacolog", "drug" "dose" and "model" which heavily influenced positive selection for TPP abstracts. Interestingly, the term "experiment" was largely associated with TPB abstracts (Figure 2.9).

Class	Recall (%)	Precision (%)	F_1 (%)
TN	87.6 ± 2.48	91.9 ± 0.660	89.7 ± 1.45
TPB	93.5 ± 0.418	88.2 ± 2.05	90.8 ± 1.25
TPP	77.9 ± 4.60	82.3 ± 5.42	80.0 ± 4.35

Values as mean ± standard deviation, $n = 3$.

Table 2.2: Naive Bayes Classifier statistics using the optimised set of 3 546 features.

The NBC classifier will use these word probabilities as seen in Figure 2.9. For example, a document with frequent occurrences of "experiment" and "pathway" would be preferentially classed as a TPB document unless it also had many counts of high-frequency TPP features such as "drug interact" or "pharmacometr".

Figure 2.9: Centroids of normalised feature frequency between classes (the top 3 546 features). Features were normalised to their maximal frequency between the classes. Significantly different TF-IDF word features are highlighted ($p < 0.01$, pairwise Conover test across all combinations).

2.3.2.2   *Classification of the query results*

The accuracy of the mNBC after optimising the features was 89.5 ±
1.44 per cent (mean ± standard deviation, $n = 3$) in predicting test
data. Using this model, the abstracts retrieved using query eight ($n =$
372 967) were classified. Only 5 432 (1.5 per cent) of these were cat-
egorised as TPP articles while 56 714 (15.2 per cent) of the abstracts
were classed as TPB. The remaining 83.3 per cent were categorised as
TN (Figure 2.10).
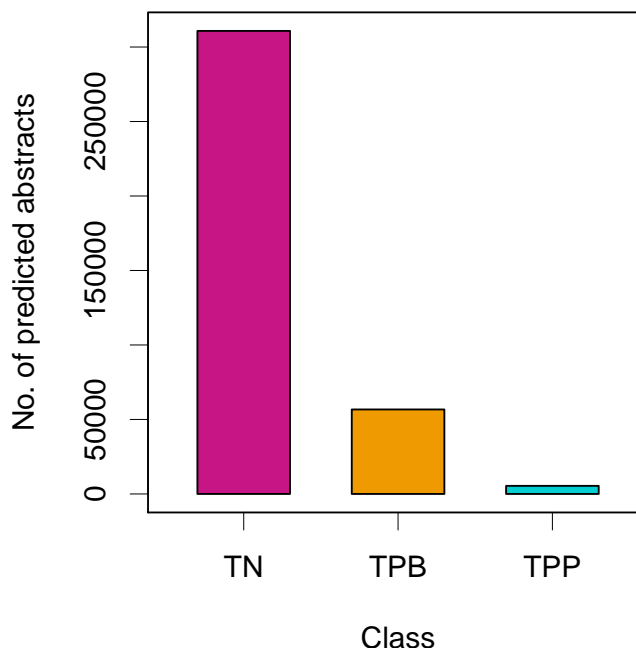
Figure 2.10: The number of predicted abstracts per class for the I2E$^{©}$ query
corpus for the period between 1965 and 2015.

2.3.3   *Annotation and entity extraction.*

We extended the methods used to examine the results of the I2E$^{©}$
query, the TPP abstracts were enriched for disease as well as gene and
chemical entities. This was carried out by passing the TPP PubMed
identifiers programmatically through PubTator (Wei et al., 2013) and
parsing the results in R. The rest of this chapter will focus upon these
annotations of TPP category abstracts under the assumption that they
represent pharmacological modelling research.

Looking at a cumulative view of the past 50 years, the top three
enriched disease categories in TPP abstracts were *Cancer*, *Nervous Sys-
tem Diseases* and *Cardiovascular Diseases* (Figure 2.11). This was similar
to the published query result (Figure 2.5)). However, in contrast with
the original query, the fraction of TPP abstracts enriched with *Immune*

System Diseases or *Digestive Systems Diseases* were two-fold higher. Additionally, the classifier appears to have removed a large fraction of documents discussing *Wounds and Injuries* which were present in the initial query. The higher precision and recall using the classifier compared with the query resulted in a closer match of systems pharmacology modelling and clinical disease agendas across this time period.
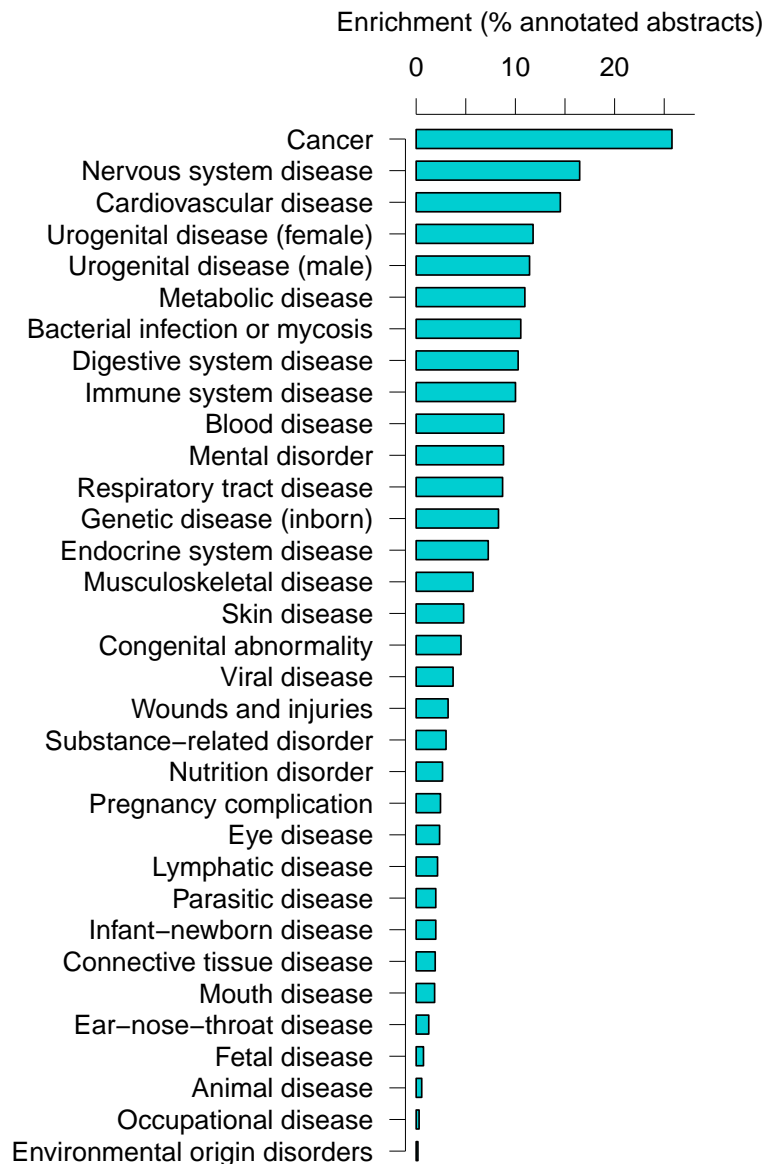
Figure 2.11: Disease enrichment of systems pharmacology. Abstract annotations were retrieved from PubTator and mapped to higher MEDIC-slim categories.

To examine trends in disease focus since 1980, the annotated abstracts were separated by date. Initially, the dominant focus was in *Urogenital Diseases* however this all but vanished by 2010 (Figure 2.12). Throughout this period TPP abstracts were increasingly annotated with *Cancer* and *Nervous system disease*. However, it it interesting to see that there was a sudden interest in the category of *Bacterial infections or mycosis* in the mid nineties which surpassed the fraction of *Cardiovascular disease* annotations in 2015.
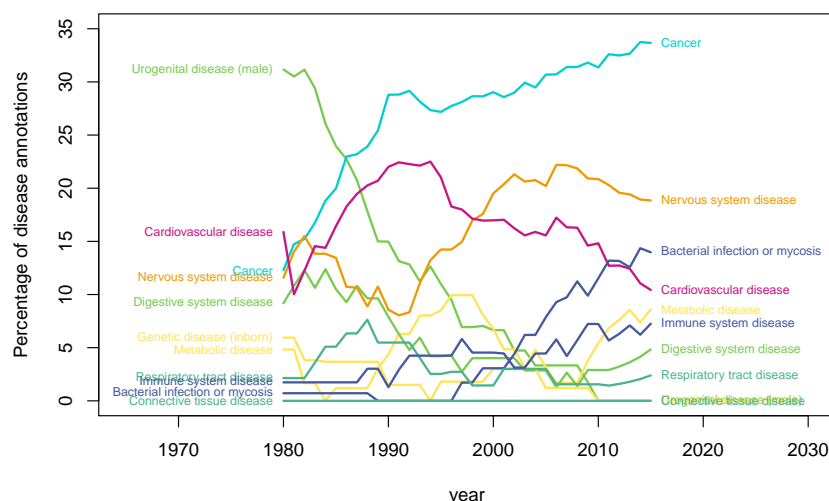
Figure 2.12: Disease enrichment of systems pharmacology over time. Abstract annotations were retrieved from PubTator and mapped to higher MEDIC-slim categories. Plotted lines are the average percentage over ten years.

The majority of predicted systems pharmacology models discuss *Homo Sapiens* within the abstract or title followed by the common rat and the house mouse (70 per cent, 21 per cent & ten per cent respectively, table 2.3). The most common gene entities in TPP abstracts code for metabolic enzymes such as CYP3A4 and important pharmacological targets such as the drug efflux transporter ABCB1, involved in regulating biodistribution and multi-drug resistance (5.6 & 3.7 per cent of TPP abstracts respectively). The chemical creatinine appeared the most frequently followed by glucose and cholesterol (5.3, 3.1 & 2.5 per cent of TPP abstracts respectively).

Table 2.3: Species, chemical and gene annotations.
[Top ten species, chemical and gene annotations (% of abstracts)]

Species (%)	Chemicals (%)	Genes (%)
Homo sapiens (69.5)	Creatinine (5.35)	CYP3A4 (5.64)
Rattus norvegicus (20.7)	Glucose (3.11)	CD59 (3.72)
Mus musculus (9.97)	Cholesterol (2.46)	ABCB1 (2.46)
Canis lupus familiaris (4.06)	Propofol (2.22)	IL1RL1 (2.28)
HIV virus 2 (2.77)	Calcium (1.59)	INS (2.22)
Oryctolagus cuniculus (2.7)	Dopamine (1.52)	CYP2D6 (2.16)
Staphylococcus aureus (1.65)	Midazolam (1.5)	EGFR (1.8)
Escherichia coli (1.61)	Morphine (1.5)	EPO (1.74)
Pseudomonas aeruginosa (1.52)	Oxygen (1.45)	CD4 (1.68)
Sus scrofa (1.34)	Cyclosporine (1.45)	VEGFA (1.5)

2.3.3.1 *Model definition and documentation*

String matches to model categories defined in the study by Hübner et al. (2011) showed that although ODEs were popular amongst annotated abstracts, regression models or general linear models were the most frequent technique discussed in TPP abstracts (88 abstracts vs 121 abstracts respectively). The favoured software in systems pharmacology modelling up until 2015 appeared to be NONMEM (Beal et al., 2009), followed by Simcyp (Jamei et al., 2009) and MATLAB® (The MathWorks, Inc. www.mathworks.com) (table 2.4). The majority of these programs are proprietary. Surprisingly, no matches for popular freely available tools such as COPASI (Hoops et al., 2006) or XPPAUT or even the format systems biology markup language (SBML) (Hucka et al., 2003), a standard language used in systems biology, were found in any TPP abstracts.

The journal which published the most pharmacological models in this time frame was *Antimicrobial Agents and Chemotherapy* (Table 2.4). The other journals were an assortment of pharmacology relevant literature with mixed clinical and theoretical research. However, at the bottom of the list is a journal dedicated to modelling and simulation in pharmacological contexts *CPT:Pharmacometrics and Systems Pharmaclogy* with an average 29 models published every year since the journal was established. The rate of publishing models in the other nine much more mature journals fared between 1.8 and 5.3 models per year.

Table 2.4: Top ten software and journal annotations.

Journals (n)	Software (n)
Antimicrob Agents Chemother (228)	NONMEM (379)
Br J Clin Pharmacol (172)	Simcyp (47)
J Clin Pharmacol (139)	MATLAB® (34)
Clin Pharmacokinet (136)	SAAM (29)
Pharm Res (110)	NONLIN (20)
J Pharm Sci (107)	Monolix (14)
Cancer Chemother Pharmacol (106)	GastroPlus™ (11)
Clin Pharmacol Ther (97)	P-PHARM (8)
J Pharmacokinet Pharmacodyn (97)	STELLA (8)
CPT Pharmacometrics Syst Pharmacol (88)	PK-Sim (6)

Out of curiosity, the top five cited articles within TPP classified abstracts were captured (descending order): (Cockcroft and Gault, 1976; Jonsson and Karlsson, 1999; Mager and Jusko, 2001; Sheiner et al., 1979; Yano et al., 2001), of which, only Mager and Jusko (2001) themselves were retrieved using the classifier.

2.3.3.2 *Model complexity and scope*

The nature and complexity of the models themselves has evolved over the last 50 years (Figure 2.13). PK models comprised nearly 100 per cent of modelling efforts until 1986 where models describing PD and gradually a mix of the two (PKPD) were seen. One moment can be seen where PBPK models gain initial popularity in 1992 followed by a four-year decline in the annotation of all model types from 1996. Interestingly, a decline between 1996 and 2000 was not observed in the total numbers of TPP classed abstracts published between these dates (not shown). From 2000 the rapid rise in all models was observed alongside the introduction of population PK models. In 2012, the emergence of a systems pharmacology model was seen as well as the the first appearance of a QSP model in 2013. Throughout this time, the majority of model types retrieved were PK models.
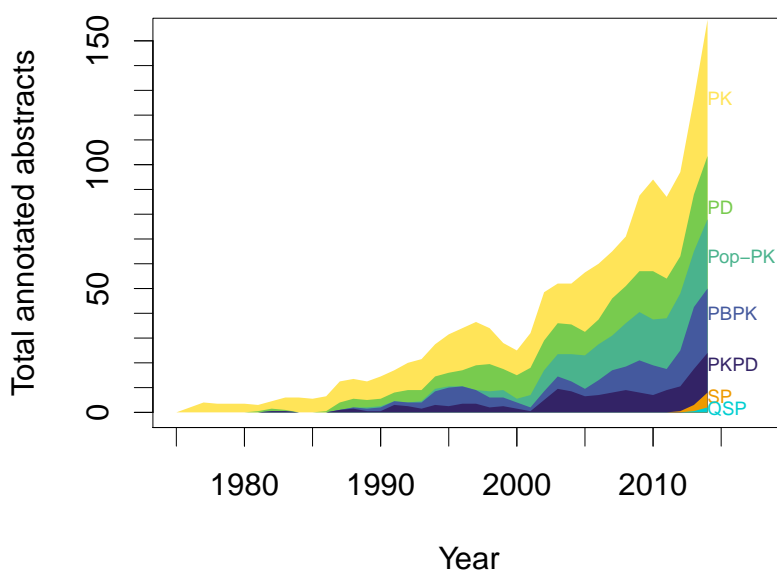
Figure 2.13: The evolution of pharmacological modelling over the last 50 years. Plotted areas illustrate the number of modelling abstracts which discuss the type of model.

2.4 DISCUSSION

The definition of systems pharmacology is sometimes poorly misunderstood even within the domain of pharmaceutical research. To resolve this I offer no additional explicit or alternative definitions of systems pharmacology as those have been provided in a number of cases (Androulakis, 2016; Cucurull-Sanchez et al., 2012; Mager and Kimko, 2016; Sorger et al., 2011). Instead, the work presented offers an emergent and tacit definition of systems pharmacology models through observation.

In profiling the current landscape of systems pharmacology modelling, a subset of literature was analysed through text mining approaches. Firstly, a corpus of systems modelling literature was retrieved and the overlap and disjoint of modelling and clinical disease agendas were examined. This revealed diseases of relative neglect where systems pharmacology can perhaps work upon in pre-clinical and clinical phases of drug development. However, this initial approach fell short of expected recall and precision measures where a large number of false positives and negatives were suspected.

Secondly, a supervised text classification of the corpus using a mNBC allowed for a more meaningful and reliable analysis of the landscape. The predicted systems pharmacology modelling abstracts retrieved a wide range of models relevant to drug discovery. Finally, retrieved abstracts were annotated, uncovering a variety software and

methods used in pharmacological modelling in the run up to the systems pharmacology paradigm as well as a glimpse at some of main biological mechanisms targeted by modellers within the last 50 years.

2.4.1 *Challenges with the initial retrieval*

A number of issues were encountered with interpreting the results of the query. Access to I2E$^©$ software was restricted and the final retrieval for analysis was run in December 2015. This means that beyond this date, published accounts of systems pharmacology models were inaccessible using the current methods.

The exact signal to noise ratio of the query was impossible to calculate without knowing the full extent of true-positive articles within the corpus. Estimated precision for systems pharmacology modelling articles was at maximum 5 per cent (Knight-Schrijver et al., 2016) which was particularly poor. The limitations of this query are described more fully in the published account (Knight-Schrijver et al., 2016). Additionally, the optimised mNBC in this study predicts that only 1.5 per cent of the corpus was TPP which actually implies that this was an overly generous estimate. The estimate for systems biology models in the query (18 per cent) however was a little closer to the mark (The mNBC predicted that 15.2 per cent of the corpus was TPB). The original query assumed a correlation between BioModels and systems pharmacology modelling as BioModels was used as a positive dataset for optimising recall. However, this and the inclusion of MeSH$^®$ terms to reduce the noise may have actually biased the query results towards BioModels-like abstracts. Hence, the MeSH$^®$ term filtering was not used in the initial analysis or further classification steps for the retrieval of systems pharmacology models.

2.4.2 *Supervised classification of text*

The aim of this study was not to develop a novel classification tool, nor was it designed to develop optimal feature selection methods; the implemented feature selection and classifier were a pragmatic approach to document retrieval. However, combining a Kruskal-Wallis test with an NBC was previously shown to give a high accuracy in text classification (Vora and Yang, 2017). As such the approach as described is a suitable supervised strategy for the accurate categorisation of research abstracts.

Implementing the mNBC in R resulted in an effective text classification with an overall accuracy of 89 per cent as well as an F_1 score of 80 per cent for the TPP class. While these scores suggested that the class predictions were of reasonable quality for further analysis it does not mean that the classified documents were without flaw;

a number of hurdles stand before the perfect classification of documents is achieved.

TRAINING DATA CURATION   Fundamental definitions of a systems pharmacology model are poor. To label documents, I attempted to apply a set of criteria for abstracts to be annotated as TPP. (Knight-Schrijver et al., 2016). A functional systems pharmacology model must contain:

1. the components that constitute the biological system of interest.

2. the temporal dynamic nature of each individual component of this system.

3. the interconnectivity and temporal dynamic interaction between these components.

4. the modulation of components and dynamic interactions by putative therapy or compound(s).

It is hoped that the spread of models contained within training data capture the terms given by participants in the survey by (Nijsen et al., 2018), including PBPK and mechanistic PKPD models as well as more comprehensive models. That said, a degree of shared methods and terminology is seen between PKPD, pharmacokinetic, pharmacodynamic and pharmacometric models which often makes them challenging to separate unanimously. With mutual exclusivity necessary for the classification, it was difficult to clearly label some abstracts within either the TPP or TN classes as abstracts were occasionally ambiguous. For instance, where pharmacokinetics are studied in a model, how many biological variables does it take to constitute as a system for a systems pharmacology approach? Likewise, distinguishing a model between TPP and TPB articles is partially subjective in assessing the model system's contribution to understanding drug action. For example, small-system models studying receptor perturbation may be classed as TPP or TPB depending on the author's aims. With a self-annotated abstracts this may have lead to certain biases within the results.

Chapter three of Systems Pharmacology and Pharmacodynamics illustrates the variety of models which could be described as QSP (Mager and Kimko, 2016).

These issues could be lessened by employing a group of annotators from pharmacological modelling sub-disciplines which could provide a more realistic labelling of training documents. Additionally, future work could provide a definitive set of annotated texts for intra-pharmacological modelling classification should the need arise.

Another mNBC consideration is that the training set was not balanced. Because the distribution of TPP, TPB and TN classes is unknown within the total population, the *prior* probabilities were configured manually to 0.1, 0.5 and 0.94 respectively. It was assumed that any evidence would be strong enough to distinguish between classes on

its own merit and a conservative *prior* value was used to retain TPP precision.

NOISE AND FALSE POSITIVES   The field of systems pharmacology modelling is young and only 18 of the abstracts classified here explicitly mention systems pharmacology or QSP in text. Therefore, despite using the classifier, an appreciable level of uncertainty still shrouds the precision of the result. However, models which apply systems methods are likely to exist which do not make use of the term "systems pharmacology model". These were expected to be retrieved by this classifier.

Some false positives can be inferred through their effect upon the results. The number of abstracts annotated with species other than *homo sapiens* (*rattus norvegicus* for instance) suggests a large presence of false-positive *in vivo* PKPD models with minimal or no modelling. Annotation of text with morphine exemplifies a false-positive TPP article and no morphine annotated abstracts contained a systems pharmacology model; most were rodent models of kinetics and the occasional multi-compartment PK model. Even so, one PBPK model (Willmann et al., 2009) was found. Another example in the disease category of *Cancer*, sampled false-positives may be seen with clinical studies in patients or *in vivo* animal models where PK parameters were predicted (Kamath et al., 2011; Würthwein et al., 2013). To combat this, the expansion of TN training samples may be necessary by feeding such models back into the training set.

Given the small size of the field at the time of research, one issue in this study could be that collecting systems pharmacology modelling abstracts for training may have exhausted the positive abstracts in literature. If this were the case, the retrieval exercise and classification would only serve to dilute the current (manually retrieved) data. While a total depletion of positive abstracts is not likely, this can only be wholly disregarded through a thorough examination of the classifier's output.

2.4.3   *Named entity recognition and information extraction*

Annotation of texts using PubTator was shown to have a high degree of accuracy with the Disease, Gene and Species F_1 scores above 80 per cent (Chemical entity extraction was seen at 53 per cent) (Wei et al., 2013). With this in mind, only a handful of entity extraction mistakes were expected and on the whole this was the case. However, with real implications to the domain of systems pharmacology modelling (or indeed other modelling fields), some instances of false entity tagging may drastically change the result of such an analysis.

CHALLENGES IN AUTOMATIC ANNOTATION   In diseases a notable false positive was seen where PD was recognised as *Parkinson's disease*. Furthermore the erroneous recognition CD59 and IL1R1 tagged gene entities influence the results in this analysis. These were caused by misidentification of units min^{-1} as MIN1 (an alternative name for the gene CD59) and IL1R1 as a pseudonym for descriptions of half-lives; most if not all retrieved IL1R1 tagging events were due to the appearance of *t1/2 alpha* or *t1/2 beta* in text.

Mismatches like this are not made abundantly clear in studies on annotation issues in pharmacology (Herrero-Zazo et al., 2013), and personal communications with text mining experts have suggested that such domain-specific language needs to be incorporated for greater reliability in name entity recognition tasks. Additionally the inclusion of specific ontologies such as the units ontology (UO) (Gkoutos et al., 2012) within name entity recognition tasks may mitigate against some false positives such as units of time.

Mathematical terms were difficult to extract from text and the use of an organised database for mathematical entities could prove useful. The recognition of mathematical terms was using the mathematical modelling ontology (MAMO) was attempted. However synonyms for preferred terms did not appear to be exhaustive for the regular expression methods used here (for example, although the term *ODE model* was present, the acronym *ODE* was not. However, the result of mathematics annotation using MAMO was largely identical to the current result as *statistical models* were the top category within TPP abstracts.

The use of full texts could be considered to improve annotation as well as classification as key entities are often reported within the main body of research and not always in the abstract. Furthermore, the methods section is bound to hold a wealth of domain-specific data which could aid classification. This may increase the performance of both classification and annotation tasks. However this relies upon the availability of text for access and distribution. PubMed Central® offers a range of structured full text data which could be used.

2.4.4 *The landscape of systems pharmacology models*

We can use the annotations of predicted TPP abstracts to build a picture of which disease mechanisms and targets are being explored in systems pharmacology as well as how they are explored and documented.

2.4.4.1 *Disease mechanisms and druggable targets*

A PICTURE OF DISEASE MECHANISMS IN CANCER.   Perhaps unsurprisingly, the appearance of cancer-related terms dominates the results as cancer is becoming a major portion of medical literature

(Reyes-Aldasoro, 2017) (Figure 2.12). As well as observing that *cancer* is the most common disease annotation in TPP abstracts, the focus of the genes confirms that systems pharmacology models may be focused upon mechanisms related to oncology.

ABCB1 (also multi-drug resistance protein or p-glycoprotein), as well as being involved in PK, has a role in drug resistance in *Cancer* (Vaidyanathan et al., 2016). As it is one of the most common genes in TPP abstracts, it suggests that a number of models explore drug resistance in cancer using pharmacological models. Twenty seven per cent of documents which mention ABCB1 were also annotated with cancer ((Zhang et al., 2011) and (Kolesar et al., 2011) for example).

A popular target in the landscape of systems pharmacology is EGFR and *cancer* appears 48 per cent of EGFR abstracts, for example Foo et al. (2012), Schoeberl et al. (2009) and Sharma et al. (2013). Sharma et al. (2013) carried out a mixed experimental and computational study using *in vivo* and PBPK models which illustrates how there can be an overlap between mathematical modelling and *in vivo* models. Although the prediction was correct in this example, these studies are a challenge for the mNBC to interpret, which results in false positives as well as false-negatives.

Additionally, Erythropoetin (EPO) is largely associated with both *Cancer* and *Blood diseases* in systems pharmacology abstracts. EPO administration is used to reverse anaemia caused by chemotherapy for the treatment cancer but it can also be detrimental to the effect of anti-cancer therapeutics (Hardee, 2006). Understanding the mechanistic basis of using EPO for therapy is important which places EPO in the top 10 list here seen in a number of models (Pérez-Ruixo et al., 2009; Singh et al., 2015; Woo et al., 2007). The appearance of vascular endothelial growth factor A (VEGFA) is also strongly associated with cancer in abstracts (Hansson et al., 2013).

Random sampling of the abstracts annotated for *Cancer* provides a number of mathematical models and PKPD studies for consideration (Kamei et al., 2010; Pop et al., 1996; Tian et al., 2008; Tornøe et al., 2007; Yates et al., 2015). However, a few false positives also become apparent (Kamath et al., 2011; Würthwein et al., 2013).

DISEASE MECHANISMS IN NERVOUS SYSTEM DISEASES   One of the most common chemical entities found in TPP abstracts was dopamine which was mostly associated with the *Nervous system diseases* and *Mental disorders*. This suggests that systems pharmacology is occupied with exploring the mechanisms of numerous dopamine system dysfunctions (Parkinson's Disease and schizophrenia being the most prominent). For example, a QSP model and a PKPD model were retrieved with both annotations of *Nervous system diseases* and dopamine (Geerts et al., 2015; Reddy et al., 2012). An analysis of systems biology and pharmacology models in *Nervous system diseases* in the

context of Alzheimer's disease has previously shown the abundance of systems models in this disease category (Lloret-Villas et al., 2017).

CARDIOVASCULAR DISEASES AND DIABETES   Extraction of cholesterol was most often seen with *Cardiovascular disease* and *Metabolic disease* annotations in TPP abstracts. The overwhelming majority of these models were regression models using clinical data. However, one example of a systems model was found within this subset (Lu et al., 2014). Although *Cardiovascular diseases* is the third most enriched category within TPP abstracts, the entities of genes or chemicals do not clearly define a disease mechanism or target within this branch of conditions.

Another target highlighted is INS (Insulin) with a high rate of occurrence in both *Metabolic Disease* and *Endocrine system disease* annotated abstracts. Insulin is the cornerstone of diabetes and its part in the disease mechanism of diabetes has been well-studied through systems biology modelling (Ajmera et al., 2013). Diabetes is a large slice of the global disease burden (Tabish, 2007) and it is expected that it will retain a high position in the landscape.

PHARMACOKINETICS ARE HISTORICALLY THE MAIN TOPIC   A large number of genes identified in the landscape show a tendency to be associated with PK modelling and an over-representation of empirical or population PK modelling was suggested due to several observations. Firstly, evidence implies that the most common genes (CYP3A4, ABCB1, CYP2D6) in TPP abstracts influence the absorption, distribution, metabolism, excretion and toxicity (ADMET) attributes of drugs (Bosch et al., 2006; Wolking et al., 2015; Zanger and Schwab, 2013). As such they are of significant interest to population PK studies. Additionally, the most observed chemical is creatinine which is often used in measuring kidney function in PK studies (Mangoni and Jackson, 2003). This is also implied through the most cited paper in these classified abstracts (Cockcroft and Gault, 1976). Furthermore, the majority of models observed in Figure 2.13 are PK models across all years. The observation that early models were highly enriched for *Urogenital disease* terms is further evidence of this. It is clear that historically, mechanisms regulating the PK of drugs have been a mainstay focus of computational modelling in pharmacology. However, these models may be published at a lower rate in today's research climate.

ANAESTHESIOLOGY AND SURGERY   A relatively large fraction of chemical annotations suggest a large focus on anesthesiology and surgery. Propofol is used as an anaesthetic and its kinetics are therefore highly important to its activity in patients. PBPK modelling of propo-

fol is seen in work by Edginton et al. (2006). Similarly, midazolam is simulated in PBPK models for anaesthesia (Gaohua et al., 2012). However, morphine's appearance here exemplifies a false-positive TPP article and no morphine annotated abstracts contained a systems pharmacology model; most were rodent models of kinetics and the occasional multi-compartment PK model. One such can be described as a PBPK model (Willmann et al., 2009). Cyclosporine annotations are largely derived from post-surgery models including a few PBPK models (Jonge et al., 2005; Wilhelm et al., 2012). Although anaesthesia is an intertwined field with PKPD modelling (Gambús and Trocóniz, 2014), no strong examples of systems pharmacology modelling approaches are seen.

OTHER NOTABLE HITS   Cluster of differentiation 4 (CD4) hits were largely associated with in-text references to CD4 counts used as variables in statistical predictive models (Revell et al., 2012). However, at least one example of using the CD4 expression in a mechanistic model is seen (Page et al., 2015). The major disease categories recognised in CD4 containing abstracts were *Viral Diseases* and *Immune System Diseases*.

2.4.4.2 *Documentation, methods and the evolving landscape*

By and large, the physiological scale of the models retrieved was above the cellular level, focusing more upon tissue distribution and systemic response of drugs and interactions rather than the intricacies of intracellular signal transduction and events. A significant contribution to this observation is the motivation at the time for PK and PD model development within the pharmaceutical industry. This was a time to explore dose–effect relationships and predict optimal dose regimens for patients (Holford and Sheiner, 1981). This has been shifting to incorporate higher resolution networks in more recent years.

Only one unique journal published a pharmacological model in the sixties, 63 journals in the seventies and 206 in the eighties. Since the new millennium, 913 journals published pharmacology models.

DOCUMENTATION OF MODELS   The number and variability of models published each year is certainly increasing. More and more journals with a pharmacological focus are publishing systems pharmacology models and dedicated journals start to present the variety of modelling in the field. Despite their youth, these dedicated journals such as CPT:PSP will likely become centre-pieces of the landscape as the rate they published systems pharmacology models was far greater than mixed-methods papers.

METHODS OF MODELLING   ODEs are the most popular mathematical technique used for simulation in systems pharmacology. However, other methods such as agent-based or logic models are occasionally seen in systems pharmacology models (Cosgrove et al., 2015; Traynard et al., 2017). The results suggest that PK and population PK

models were the most common type of model throughout the last 50 years. The large spike in PK models seen here adds evidence to the role that PK modelling played in reducing PK-mediated attrition between 1990 and 2000 (Knight-Schrijver et al., 2016; Kola and Landis, 2004).

The tools used to construct models during the examined time-frame also agree that PK models were the most common type of pharmacology model. The most frequent software listed are designed for PK modelling (Table 2.4). One criticism of the field is the use of proprietary software which may hinder reproducibility of models and the access to modelling environments for would-be systems pharmacologists. However, the variety of software is changing to accommodate for the structural complexities and data required by today's modelling which brings a selection of alternative tools.

THE EVOLUTION OF MODELLING INTO SYSTEMS PHARMACOLOGY. Models are only as complex as they need to be. For example, the rise of retrieved population-PK models coincides with the initial sequencing of the human genome (Lander et al., 2001) (figure 2.13). Sequencing of the human genome completed the identification of cytochrome p450 polymorphisms (Ingelman-Sundberg, 2004). Although population-PK models were used before this time, it is clear that growth in the field was suddenly triggered. Examples like this show that as information becomes available, models seek to utilise it in a meaningful way.

An analogy using evolution can be used to describe this as the progressing complexity of mathematical models in pharmacology is seen to be collectively curious, rational and methodical as more is revealed about the nature of biology (figure 2.13). In minimising the objective function of the difference between a simulated world and the real world, models evolve to accommodate for our increased understanding of complex dynamic systems. First generation models observe that the body affects drug movement when measured and explore the drug PK. Future generations of models understand that the opposite also occurs; drugs alter the body and a response is seen. Mathematical models of the PD begin to simulate this effect. Collective theories about the input-output relationships of drug and body are formed, models simulate the dose-response and two sets of observations are combined in PK and PD models. Evidence sheds light upon the differences within populations and population models begin to emerge. Finally, the current iteration of models using PBPK, QSP and systems-biology tries to understand the higher resolution dynamic networks that we see in biology which together form the state-of-the-art generation of modelling in systems pharmacology. A neat example of this modelling evolution is a review within the context of atherosclerosis by Pichardo-Almarza and Diaz-Zuccarini (2017a).

2.4.5   *Concluding remarks and future direction*

The current landscape of systems pharmacology models stands on the shoulders of a previous foundation of pharmacological modelling. This foundation, generated by the development of early PK and PD models, harbours a scientific effort to understand dose-response relationships that are invaluable in drug discovery. Now, current models aim towards a building dynamical networks from the fundamental understanding of disease mechanisms and targets and their focus is similar to that seen in clinic.

To expand upon the study, future work should benchmark the classification approach on domain-independent datasets. Furthermore, work should be undertaken to extend the dates predicted to include the most current models as the field will have changed in the last two years with the sudden interest in systems pharmacology and QSP. Additionally, predictions should be run across all of PubMed as elements of bias from the initial query cannot be ruled out. Finally, I had planned to reassess the focus of clinical trials alongside models using the same pipeline of annotation described here. However, the older set of MeSH® terms and the work flow through I2E© could not be updated. The retrieval of clinical trial data fields may be carried out through alternative means.

3

PREDICTING DRUGGABLE TARGETS:
DEVELOPING A QUANTITATIVE SYSTEMS
PHARMACOLOGY MODEL.

3.1 INTRODUCTION

This chapter presents an opportunity taken to build a model in a disease area of relatively low enrichment in systems pharmacology models. The value of this research is three-fold. Firstly, rheumatoid arthritis is an area of clinical focus within the ageing population and aiming to treat the disease effectively is a major desire in today's society. Secondly, the number of biologics and therapeutic monoclonal antibodies (mAbs) in drug development is increasing; constructing QSP models of mAbs contributes to the knowledge surrounding their use. their use. Thirdly, I describe the process of adapting a previous model from its components which highlights some of the challenges in mathematical modelling and QSP surrounding model documentation, standards and validation.

3.1.1 *Rheumatoid arthritis and interleukin-6*

THE IMPACT OF RHEUMATOID ARTHRITIS   Rheumatoid arthritis is typically defined as an inflammatory disease of the joints. Driven and maintained by a complex and varied array of autoimmune responses, primary clinical manifestations are painful swelling of synovial joint capsules and the destruction of cartilage and bone. This leads to a loss of function in patient mobility and quality of life. The world-wide prevalence of rheumatoid arthritis (RA) is 0.24- one per cent (Cross et al., 2014; Silman and Pearson, 2002) rising to two per cent in individuals over 60 years of age and higher still as age increases (Helmick et al., 2007; Rasch et al., 2003). Furthermore, in terms of contribution to global disability, RA has been ranked nearly as high as malaria (Cross et al., 2014). Therefore, in conjunction with the ageing population, the increasing and age-associated global incidence of RA (Minichiello et al., 2016), places the disease in realms of significant clinical interest.

DISEASE MECHANISMS   Much has been uncovered about the putative mechanisms implicated in RA pathology with both genetic and environmental factors contributing to the disease. Moreover, a major and well-documented component of RA pathogenesis is mediated by the regulation of cytokines (McInnes and Schett, 2007; Stahl et al.,

2010) such as tumour necrosis factor-alpha (TNF-α) and interleukins 6, 17 and 23 (IL-6, IL-17 & IL-23) ("Siebert et al., 2015). One cytokine in particular, the pleitropic interleukin-6 (IL-6), has been shown to contribute to the disease state by promoting the infiltration of immune cells into the synovial tissue and sustaining active neutrophil populations (Lally et al., 2005). Further deleterious IL-6 effects are also shown through its role in joint destruction by regulating osteoclast numbers and activity (Le Goff et al., 2010). Therefore, inhibiting the activity of IL-6 is deemed viable as a strategy for treating RA.

3.1.2 *Endogenous roles of IL-6*

IL-6 IN INFLAMMATORY DISEASES   The dysfunctional potentiation of IL-6 signalling drives pathology in a large number of inflammatory conditions including RA and Crohn's disease (CD) (Jones et al., 2011). On the whole, ubiquitous increases in the synthesis of IL-6 is seen in inflamed, injured and stressed tissues which is thought to determine the major differences between healthy and disease physiology. In patients with RA, there is a distinct increase in the synthesis of IL-6 from synoviocytes and the concentration of IL-6 in the synovial fluid (SF) which implicates synovial IL-6 production as a hallmark of the rheumatic joint (Guerne et al., 1989; Kotake et al., 1996; Rosenbaum et al., 1992). The raised SF concentration of IL-6 in RA contributes to joint destruction by increasing osteoclast and matrix metalloproteinase activity (Kotake et al., 1996; Srirangan and Choy, 2010). IL-6 has been shown to drive differentiation, strong survival and anti-apoptosis signals in T cells (Dienz and Rincon, 2009; Durant et al., 2010; Takeda et al., 1998). Furthermore, IL-6 induces T cell migration in acute inflammation in mice (McLoughlin et al., 2005), potentiating exaggerated local immune responses; a persistent T cell population is one of the mechanisms by which chronic inflammation is maintained. In hepatic tissue, high IL-6 activity promotes an acute phase response with secretion of proteins such as C-reactive protein (CRP) (Norris et al., 2014), and the systemic effects of IL-6 in RA are observed through this positive effect upon CRP secretion (Rosenbaum et al., 1992). As such, CRP is used as a systemic marker of inflammation. Additionally, CRP is a likely factor in the progression of coronary heart disease (Shrivastava et al., 2015).

SOURCES OF IL-6   Secretion of IL-6 is a cellular response to stress and occurs in a wide variety of tissues. For example, IL-6 synthesis in the liver is a relatively normal process which is seen to increase in response to trauma (Norris et al., 2014). Evidence also suggests that rapid and transient secretion of IL-6 is a major player in liver regeneration post-injury (Schmidt-Arras and Rose-John, 2016). Another important source of IL-6 lies with peripheral blood mononuclear

cells (PBMCs) (Jansky et al., 2003) and is essential to consider when accounting for an immune response and maintenance of autoimmune diseases. A further major source of IL-6 is reported to be adipose tissue which may contribute to insulin resistance (Lagathu et al., 2003).

3.1.3 *IL-6 targets and signalling*

Evidence suggests that IL-6 has two populations of cell-surface binding sites, low and high-affinity, presumably caused by the sequential binding of IL-6 to receptor subunits. It is likely that IL-6 associates with its primary binding target membrane-bound IL-6 receptor-alpha (mIL-6Rα) followed by its signal transduction subunit, glycoprotein 130 (gp130), forming a heterotrimer (Boulanger, 2003; Taga et al., 1989). Finally, a dimerisation step forms the putative hexameric receptor complex necessary for downstream signalling (Boulanger, 2003). It is argued that the low-affinity binding is the association of IL-6 to mIL-6Rα only and that the high-affinity site is the binding of mIL-6Rα to IL-6 and the subsequent fast binding of the mIL-6Rα:IL-6 complex to gp130 (Schroers, 2005). The observed low-affinity binding has a dissociation equilibrium constant (K_D) of 500 pM whilst the high-affinity interaction has a K_D of 15 pM in human hepatocellular carcinoma cell line (HepG2) cells (Baumann et al., 1988). Attempts to reproduce IL-6 signalling with known components of the canonical receptor transfected into monkey kidney fibroblast-like cell line 7 (COS-7) cells resulted in lower affinity interactions with a K_D between 0.3 and three nM (Dittrich et al., 1994; Gearing et al., 1992; Heinrich et al., 1998). We also see that binding of IL-6 to a soluble form of the receptor soluble IL-6 receptor-alpha (sIL-6Rα) occurs at the high picomolar range with a K_D of 500 pM (Weiergraber et al., 1995). This is similar to the interaction with the membrane-bound receptor. In the same study using HepG2 cells, the binding of IL-6 to the complete array of cell-surface receptor components resulted in a K_D of 50 pM, which was comparable to previous investigations describing high-affinity binding sites (Sonne et al., 1990; Weiergraber et al., 1995; Zohlnhöfer et al., 1992). The K_D for the binding reaction between IL-6 and its receptors *in vivo* is in the range of 20-1000 pM. After receptor binding, the plieotropic nature of the *in vivo* response is driven by the relative tissue expression of the cytokine's receptor components. The binding and transduction of IL-6 signals occurs through two selective pathways, classical and trans-signalling pathways.

CLASSICAL SIGNALLING   The first pathway is coined as classical IL-6 signalling and is mediated through the membrane-bound forms of the receptor and signal transduction element of the receptor (mIL-6Rα and gp130 respectively). Cell-surface expression of the receptor components depends on the organ, tissue or cell type. Hepatocytes con-

stitutively express mIL-6Rα at high concentrations on the cell-surface which results in a high concentration of the receptor in the liver. Other notable organs with high messenger RNA expression include the small intestine and the ovaries (Uhlen et al., 2015). Lymphocytes are one non-hepatic cell type which typically have a high expression of mIL-6Rα (Uhlen et al., 2015). In contrast, most other organs have a low cell-surface expression of mIL-6Rα. This suggests that classical IL-6 signalling is relatively cell-selective.

TRANS-SIGNALLING   The second IL-6 response pathway called trans-signalling is through a soluble form of IL-6 receptor-alpha (IL-6Rα), (sIL-6Rα), generated through the shedding of mIL-6Rα by matrix metalloproteinases as well as synthesis through alternative splice variants (Briso et al., 2008; Jones et al., 1998; Lust et al., 1992; Matthews et al., 2003; Schumacher et al., 2015). Shedding is the process by which cell-surface receptors are dispersed into the surrounding medium. Immune cells are largely responsible for sIL-6Rα production and active T cells may preferentially produce sIL-6Rα. Furthermore, cluster of differentiation 4 positive (CD4+) T cells have also been shown to downregulate mIL-6Rα (Briso et al., 2008; Jones et al., 2010). It is interesting to note that sIL-6Rα sourced from splice variants is seen to decrease with age whilst the total concentration in serum remains unchanged. This suggests that receptor shedding is partially age-dependent (Jones, 2001). Elevated concentrations of sIL-6Rα are seen in RA patient's synovial fluid in comparison with serum fluid which may be driven by presence by synovial leucocytes (Desgeorges et al., 1997; Jones, 2001). The mechanism by which this pathway elicits an effect is through the binding of sIL-6Rα:IL-6 to membrane gp130. This results in the formation of an active cell-surface receptor complex necessary for IL-6 response in tissues that otherwise do not inherently express mIL-6Rα. Hence, gp130's ubiquitous cell-surface expression in tissues (Uhlen et al., 2015) enables IL-6 to bind and elicit a response at a large number of sites via the soluble form of receptor. The availability of two pathways results in a tissue selection mechanism and the dynamic interplay between IL-6 binding to either classical mIL-6Rα or sIL-6Rα could be key to IL-6 pathology. Evidence in literature suggests that through sIL-6Rα, IL-6 activity may be pro-inflammatory whereas the interaction of IL-6 with mIL-6Rα may be anti-inflammatory (Garbers et al., 2015; Rose-John, 2012; Scheller et al., 2011). For example trans-signalling via sIL-6Rα stimulates T cell migration through chemokine secretion in multiple cell lines (Hunter and Jones, 2015; McLoughlin et al., 2005), one of the drivers in maintained T cell populations in RA.

OTHER SIGNALLING COMPONENTS   A soluble form of gp130 can be found in patients (Diamant et al., 1997; Narazaki et al., 1993).

With similarity to sIL-6Rα, soluble glycoprotein 130 (sgp130) is sourced through alternative splice variants and membrane shedding (Mullberg et al., 1993). Association is seen between sIL-6Rα and sgp130 forming a soluble receptor complex. However, as sgp130 contains no transmembrane region for signal transduction, the interaction of sIL-6Rα with sgp130 has been shown to inhibit IL-6 signalling (Garbers et al., 2011; Tanaka et al., 2000). Evidence suggests that this could be a negative feedback mechanism (Tanaka et al., 2000).

RECEPTOR ACTIVATION   Evidence for the step-wise formation of the active hexameric receptor was shown to 3.65 Å (Boulanger, 2003). The hexamer, comprised of two IL-6, mIL-6Rα and gp130 heterotrimers results in an active receptor with a 2:2:2 component composition. Signal transduction is driven by the phosphorylation of intracellular gp130 domains. However, the hexameric structure has been disputed as evidence also indicates receptor activity in a tetrameric form with component stoichiometry of 1:1:2 (IL-6, mIL-6Rα and gp130 respectively) (Grötzinger et al., 1999). Further controversy is found in discovering that mIL-6Rα can pre-form a dimer both in the cell-surface membrane and in solution (Schuster et al., 2003). Additionally, pre-formed dimerisation may also occur in membrane-anchored gp130 (Tenhumberg et al., 2006) which aids the tetrameric receptor hypothesis as a pre-formed gp130 dimer may bind to one IL-6:mIL-6Rα complex to activate. Tenhumberg et al. (2006) also demonstrate that the ligand binding to this pre-formed gp130 dimer is essential for the tyrosine residue phosphorylation and signalling confirming that all three components are required for receptor activation. Ultimately, both active forms of the receptor may exist and could result in different response behaviours at varying concentrations of IL-6 (Scheller et al., 2011).

INTRACELLULAR SIGNALLING AND STAT3   Alongside gp130 phosphorylation, janus kinases (JAKs) one and two and then signal transducer and activator of transcription (STAT) proteins, primarily one and three, are phosphorylated (Hunter and Jones, 2015). signal transducer and activator of transcription 3 (STAT3) translocation to the nucleus results in transcription of and upregulation of various pro-inflammatory genes (Bild, 2002; Cimica et al., 2011; Ushijima et al., 2005). In RA, phosphorylated STAT3 (pSTAT3) is constitutively elevated in peripheral and synovial fluid T cells at levels significantly greater than healthy controls (Gao et al., 2014; Isomaki et al., 2014). Greater levels of pSTAT3 are also seen in cluster of differentiation 3 positive (CD3+) cells (Anderson et al., 2015). The immediate sensitivity of STAT3 to IL-6 in some RA patient cells may be reduced and the fraction of pSTAT3-positive T cells is negatively correlated with plasma IL-6 (Isomaki et al., 2014). In the synovium, pSTAT3 activity promotes survival of the dysfunctional synovial fibroblasts (Krause et al., 2002).

An alternative intracellular signalling pathway is seen through SHP-2 with links to both mitogen-activated protein kinase (MAPK) and phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K) pathways and promotes cell survival and proliferation (Chen et al., 1999). Downstream from this pathway is another transcription factor, CCAAT/enhancer binding protein beta (C/EBPβ), which is also a player in genetic regulation through IL-6 (Cantwell et al., 1998; Heinrich et al., 2003). In depth descriptions of the transduction processes are discussed by Kamimura et al. (2003) and systems biology models mapping these components of IL-6 intracellular signalling attempt to unravel the connections (Moya et al., 2011; Qi et al., 2013; Singh et al., 2006). More simply, a scheme of reversible reactions has been used to describe STAT3 phosphorylation in a published model (Sadreev et al., 2014).

TERMINATION OF SIGNALS   The heterodimer of IL-6:mIL-6Rα internalises with a half-life of two hours (Fujimoto et al., 2015) whilst the heterotrimeric complex of IL-6:mIL-6Rα:gp130 internalises rapidly with an approximate half-life of 15 minutes (Dittrich et al., 1996; Nesbitt and Fuller, 1992; Zohlnhöfer et al., 1992). Following internalisation, both gp130 and mIL-6Rα are thought to be preferentially degraded instead of recycled (Heinrich et al., 1998; Nesbitt and Fuller, 1992; Zohlnhöfer et al., 1992). Receptor depletion occurs rapidly and is recovered slowly after eight hours, arguing against recycling pathways (Zohlnhöfer et al., 1992). The overwhelming contribution of *de novo* protein synthesis to the restoration of cell-surface receptors was confirmed by cycloheximide (Zohlnhöfer et al., 1992). Further evidence against recycling is that the pH of endosomes in HepG2 cells did not favour dissocation of the iodinated ([^{125}I]) recombinant human IL-6 used in the study suggesting that mIL-6Rα remained bound to the cytokine for degradation. The same fate is suggested for gp130 where it is also preferentially degraded as opposed to recycled (Wang and Fuller, 1994). The signal transduction protein, gp130, is also capable of internalisation through a ligand, dimer, and activation-independent manner (Thiel et al., 1998). Using rat hepatocytes, internalised radiolabelled IL-6 was shown to be eliminated rapidly, further implying a strong degradation versus recycling pathway (Nesbitt and Fuller, 1992). Known contributors to signal termination downstream of the receptor, however, are intracellular phosphatases such as SHP-2 or the suppressor of cytokine signalling (SOCS) proteins (Kim et al., 1998; Lehmann, 2002; Starr et al., 1997). Signal attenuation through these pathways occurs independently of the receptor conformation or receptor internalisation (Thiel et al., 2000). With trans-signalling, sIL-6Rα may also be internalised through interaction with gp130 via clathrin-coated pits (Graeve et al., 1996). This has been further validated where IL-6:sIL-6Rα receptor complex was seen to clear more rapidly than unbound sIL-6Rα in rats (Weiergraber et al., 1995). Therefore trans-

signalling may also drive a target-mediated downregulation of receptor components.

CRP AS A BIOMARKER AND FEEDBACK MECHANISM   Clinical observation of IL-6 signalling and pSTAT3 activity lies in the biomarker CRP. CRP is produced in an IL-6-dependent manner from hepatocytes (although non-hepatic sources of CRP have been observed) (Kuta, 1986; Ramji et al., 1993; Zhang et al., 1996). This accounts for serum CRP concentrations found in patients with inflammatory conditions including CD and RA (Shadick, 2006; Srirangan and Choy, 2010; Vermeire, 2006). In RA, the serum concentration of CRP is used as a surrogate biomarker for disease activity; it is measured during clinical trials for anti-IL-6 compounds. CRP has been shown to cause mIL-6Rα shedding (Jones et al., 1999).

3.1.4 *Cytokine-targeting therapeutics in RA*

CURRENT GOLD-STANDARDS   In the pharmacological treatment of RA, the current gold standard is the disease-modifying antirheumatic drug methotrexate (MTX) (Shinde et al., 2014). As the first port of call, the mechanism of action of MTX is attributed to its immunosuppressive and anti-proliferative effects (Cutolo et al., 2000; Wessels et al., 2008) and is seen to be effective in reducing disease progression and relieving the debilitating symptoms of RA. However, MTX monotherapy has a poor effect in a large subset of RA patients with remission rates in patients between 16 per cent and 46 per cent (Breedveld et al., 2006; Grigor et al., 2004). Furthermore, while MTX as monotherapy is effective in slowing or halting disease progression, it may be relatively ineffective in reversing the damage in moderate to severe RA. This is in contrast to its combination with a TNF-α antagonist infliximab, a mAb, which is also seen to reverse the damage (Rau, 2010).

MONOCLONAL ANTIBODIES   The recent development of biologics targeting cytokine signalling is further evidence that cytokines present themselves as significant and readily-targeted components of RA pathology. For example the mAbs infliximab, adalimumab (anti-TNF-α), and sarilumab or tocilizumab (TCZ) (anti-IL-6 receptor) are all effective in clinic as monotherapy or in combination with MTX (Fleischmann et al., 2017; Keystone et al., 2014; Ogata et al., 2014; Shinde et al., 2014). Moreover, therapeutic polypharmacy of MTX and biologics such as infliximab or adalimumab may result in much higher rates of remission (Breedveld et al., 2006). The specific blockade of IL-6Rα function using TCZ or sarilumab confirms that inhibiting cytokines may be an effective alternative to MTX treatment and may even be comparable with MTX-based combination therapies after one year of treatment (Flipo et al., 2017).

TOCILIZUMAB AND SIRUKUMAB   We have discussed that inhibiting IL-6 signalling by blocking IL-6Rα appears to be a valid treatment for RA. However, if competitive inhibition of the receptor is effective, an equally important target could be the ligand. The development of the anti-IL-6 antibody sirukumab (SRK) (Xu et al., 2011), currently in clinical phases, is driven by the promise of an alternative target. However, it is not fully understood what relative merits or caveats exist between inhibiting either target of this binary interaction. Thus, when inhibiting IL-6 for the treatment of RA, how do the two antibodies compare? As mAbs are highly selective and their mechanism of action is generally that of true antagonism, with respect then to target druggability, which component of IL-6 signalling is better as an option for treating RA?

Given the similar strategy of IL-6 blockade, which drug offers greater therapeutic potential?

3.1.5 *Comparing drug targets*

Several approaches can be utilised to compare the druggability of targets like IL-6 or IL-6Rα. Perhaps the most obvious method would be to carry out *in vitro* studies using the biological system in question. For example, assays using receptor expressing cell-lines or tissues could indicate differences in target kinetics or redundancy in cellular responses. This lacks the full array of inter-tissue reactions present in organisms. A more holistic method would be to compare the targets in an *in vivo* setting using animal models. Comparison of target knockout or drug administration in disease models would give directly comparable results for targets. However, animal models may lack the translation of efficacy or toxicity into humans, both essential drivers of drug attrition (Garner, 2014; Shanks et al., 2009). The logical and best comparison of two targets for human intervention would be to use the target population itself. Head-to-head clinical trials are the best interpretation of relative drug-target combinations. However, head-to-head trials of drugs in development is not a common or trivial practice. Decision-making methods exist which focus upon the comparison of drugs without these comparative trials and yet, these approaches still require clinical data of the individual compounds (Kim et al., 2013). As such they are limited by the availability of clinical trials or human data and require significant investment for their decision-making power. Alternatively, QSP can utilise and interpret pre-clinical data through systems modelling and simulation to compare targets in the earlier stages of drug-discovery.
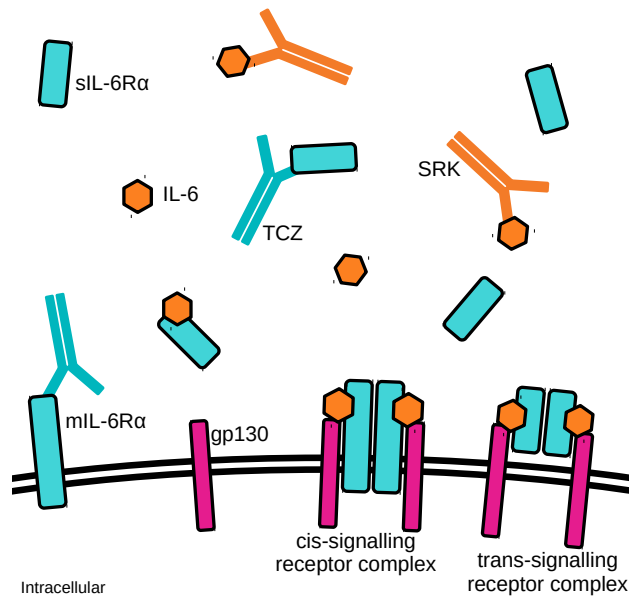
(a) TCZ and SRK mechanisms. The biological targets of TCZ and SRK are the receptors (sIL-6Rα and mIL-6Rα) and IL-6 respectively.
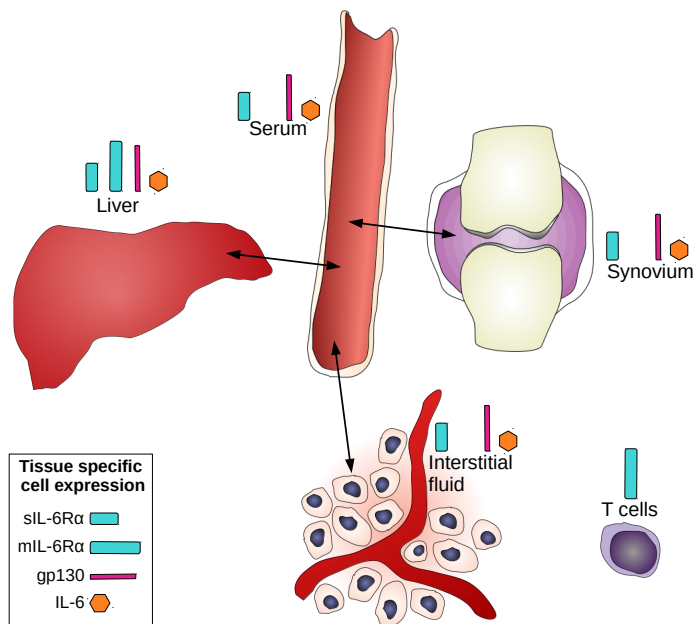
(b) Physiological compartments for IL-6 signalling in RA and the assumed tissue-specific expression of essential components.

Figure 3.1: (a). Competitive binding to both targets attenuates the dimerisation and binding of gp130, reducing the pool of active receptors but which target is better? (b), The expression of receptor components differs across tissues and cell types but T cell levels in the compartments regulate the expression of mIL-6Rα where the tissue-specific cell types do not express mIL-6Rα at a high level.

3.1.6   *Systems modelling of IL-6 signalling*

Studying RA therapeutics within IL-6 signalling by systems biology or pharmacology modelling is not entirely new. Previous modelling efforts have constructed networks for the IL-6 intracellular signalling cascades (Moya et al., 2011; Qi et al., 2013; Singh et al., 2006). Together these models comprehensively explore the downstream signalling of IL-6 in isolation and in unison with other cytokines. At a lower resolution, RA has been examined by modelling the opposing effects of pro and anti-inflammatory cytokines and their net effect upon ageing and in therapy (Baker et al., 2012). Furthermore, specific PK and systems pharmacology models have been developed which examine TCZ in its effectiveness and adverse activity in treating RA (Frey et al., 2010; Gibiansky and Frey, 2011). A deterministic QSP model of IL-6 signalling and drug target comparison was carried out by Dwivedi et al. (2014) where a QSP model was developed to explore the activity of IL-6 signalling in CD. The model was used to assess the relative merits and caveats of targeting either the sIL-6Rα, mIL-6Rα or IL-6.

For the sake of brevity, the CD model by Dwivedi et al. (2014) shall be frequently referred to as the original model.

3.1.7   *Research scope, aims and goals*

To avoid re-inventing the wheel, the task was set to adapt and re-purpose an existing model on IL-6 in CD to direct questions of target selection and drug comparisons in RA. Presented as a network of deterministic ODEs, the QSP model presented herein sought to simulate the biology of IL-6 signalling in RA to the level of detail required to adequately describe mAb-mediated perturbation of extracellular IL-6 signalling. The biological scope of the model was therefore restricted to the mechanistic cell-surface and extracellular, tissue-level and whole-body interactions of IL-6-signalling as well as a phenomenological downstream response towards IL-6. The process of repurposing questions the value and ease of reusing existing model networks for QSP approaches and provides a useful model for predictive comparison of current and in-development mAbs. Furthermore, the model aims to highlight areas of uncertainty in the knowledge of the disease mechanisms where the resulting simulations fail to fit the experimental studies or clinical data. Additionally, recent news highlights that there were safety concerns in clinical trials with the use of SRK in RA patients and it has since been withdrawn from development in treating this disease (Taylor, 2017a,b). The model may be able to explore potential reasons. Moreover, a model can still be applicable to alternative uses of sirukimab in depression or other anti-IL-6 mAbs such as siltuximab used in Castleman's disease (Rhee et al., 2010; Sun et al., 2017).

The primary goals are to:

- outline the ease of model re-usability with respect to structure and parametrisation.

- compare drugs and drug targets in rheumatoid arthritis, discussing merits and caveats of both PK and PK in each instance.

- evaluate optimal dose-regimens for SRK through simulation.

- explore potential causes of SRK adverse reactions.

Ultimately, The presented research sought to demonstrate that a QSP modelling (QSPM) approach can be used to predict and compare several aspects of two targets or compounds without the full research & development (R&D) expenditure of large-scale clinical trials.

3.1.8 *Modelling approach and development*

The approach taken to developing this model was largely to re-structure and re-parametrise an existing model. However, with differences in pathology, tissue of interest, and clinical data altering a model of CD into one of RA warranted a full re-construction.

I approached this modelling task in a modular fashion so as to implement these differences gradually and methodically. Firstly, I generated a model that describes the PK of mAbs across multiple compartments of relevance to treating and monitoring the severity of RA. Secondly and after the PK model was deemed suitable, the PK compartments were populated with mechanistic biological reactions between IL-6 signalling components in the form of a receptor turnover model. This used literature derived data to govern the rates and component concentrations. These reactions formed the PD component of the PKPD model. Lastly, after combining these individually developed modules, the model was re-parametrised with RA clinical data instead of CD data. This altered the *original model*'s parameter values using equivalent data that were not available for a CD model thus allowing for a more meaningful description of RA IL-6 signalling to answer questions specific to RA.

3.2 MATERIALS AND METHODS

3.2.1 *Modelling tools*

3.2.1.1 *Model source and model definition*

The model was derived from a published model on CD by Dwivedi et al. (2014) and was stored in the BioModels Database, a repository for systems biology models (Le Novère et al., 2006). This involved curation of the published supplementary materials file by checking

the consistency through reproducing published simulations and annotating model components according to minimal information requested in the annotation of biochemical models (MIRIAM) guidelines (Le Novère et al., 2005). The model was defined in SBML, a model description format which allows for model exchange and aims for standardisation (Hucka et al., 2003). The curated CD model can be accessed from the BioModels Database (www.ebi.ac.uk/biomodels) using the unique identifiers *[BIOMD0000000534] - [BIOMD0000000537]*. The model definition language, antimony (Smith et al., 2009), was used to write the SBML for the RA model and graphical representations of the RA model such as those seen in Figure 3.10 were compliant with systems biology graphical notation (SBGN) standards (Le Novère et al., 2009), designed using open graphics software LibreOffice Draw.

3.2.1.2   *Simulation and parameter estimation*

Simulations including parameter scans, estimations and sensitivity analyses were carried out in COPASI version 4.22 (Hoops et al., 2006) using the robust deterministic livermore solver for ordinary differential equations (LSODA) algorithm for stiff and non-stiff ODEs (Hindmarsh, 1983; Petzold, 1983).

Parameter estimation tasks were carried out using the particle swarm algorithm (Kennedy and Eberhart, 1995) available in COPASI. The objective function is a weighted sum of squares between the experimental data and the simulated values as defined as follows in the COPASI user manual (Equation 3.1).

$$E(P) = \sum_{i,j} w_j \cdot (x_{i,j} - x_{i,j}(P))^2 \tag{3.1}$$

Here, P is a parameter set being simulated and evaluated. The indices i and j refer to rows and columns in the datasetand $x_{i,j}$ and $y_{i,j}(P)$ are the experimental and simulated data points respectively. The weight for each data column is given by ω_j and is calculated both between and within experiments. Dataset weights are assigned based on the standard deviation of each experiment. To improve the whole-model estimation, the value of ω_j was overridden manually to accommodate for the nonlinearity of antibody elimination. This was to facilitate better fitting to a log scale when datasets differed by several orders of magnitude. Comparisons of model PK fits were carried out using Akaike information criterion (AIC), a model comparison statistic which can be applied to a multitude of mathematical models including PK models (Akaike, 1974; Yamaoka et al., 1978). The AIC penalises models for the number of estimated parameters used in estimation. Therefore it can be used to select between models differing in structure where more or fewer parameters are unknown.

3.2.1.3 *Data Statistics and Graphics*

Time-course and steady-state data were sourced from an array of published experiments and clinical trials and were extracted using Engauge Digitizer, version 4.1 (Mitchell et al., 2016). These data were cleaned and processed in the statistical programming environment R (R Core Team, 2016) prior to parameter estimation steps. Statistical calculations and simulation plots of the results were generated in R using a variety of bespoke functions.

3.2.2 *Parameter estimation and model assumptions.*

Because of the incompleteness of human knowledge, assumptions have to be made about the physical mechanisms, molecular interactions and concentrations, and reaction rates which allow the model to be constructed in the current framework. Estimation was carried out using a modular approach. This helps to identify parameter sets within each sub-model. For each module, assumptions are listed alongside evidence. General model assumptions are that:

COMPARTMENTS ARE HOMOGENOUS AND WELL-MIXED: Molecules are instantly and equally diffused throughout the compartment as they enter. Cells in this scenario are homogenised and readily accessible. This simplifies the model in using mass-action kinetics.

COMPONENT AFFINITY IS SIMPLIFIED: Association and dissociation rate parameters are a simplified ratio to give the apparent K_D seen in experimental studies.

3.2.2.1 *Monoclonal antibody pharmacokinetics and compartmentalisation*

The anti–mIL-6Rα mAb, TCZ, is a prominent biological therapeutic in treating RA. A number of clinical studies have characterised the PK and PD in both RA and healthy patients, providing a reasonable amount of data (Bao et al., 2012; Morcos et al., 2013; Ogata et al., 2014; Ohta et al., 2013; Zhang et al., 2013a,b). As such, TCZ was ideal for the fitting and validation of the PK model. The PK model was created assuming that:

ELIMINATION OF TCZ IS THROUGH PHYSIOLOGICAL CLEARANCE: A non-saturable route of mAb elimination occurs via site-independent mechanisms. The clearance of TCZ is not heavily hepatic and is largely influenced by body mass, implying a degradation in tissues typical of antibodies (Frey et al., 2010; Tabrizi et al., 2006). For example one part of non-specific immunoglobulin-G (IgG) antibody elimination may occur via the reticuloendothelial

system in phagocyte-rich tissues (Ferl et al., 2016). The parameter k_{el} is derived from the elimination half-life ($t_{1/2} - \beta$) of 23 days for IgG-type antibodies (Paul, 2008).

TCZ IS SUBJECT TO TMDD: A saturable route of TCZ elimination is the result of degradation or internalisation of the drug-target complex. This is known as target-mediated drug disposition (TMDD) and occurs mainly through non-central routes. This gives rise to a nonlinear rate of TCZ disappearance as seen in PK data. This is modelled through a Michaelis-Menten (MM) reaction rate in both the *liver* and *peripheral* compartments. The equilibrium constant K_M is defined as the K_D of the drug-receptor binding interaction which is between 0.7 and 2.54 nM (Mihara et al., 2005; *U.S Food and Drug Administration, BLA: 125276*). We assume that R_{tot_3} is the total concentration of hepatic mIL-6Rα, 0.83 nM (Baumann et al., 1988; Sonne et al., 1990; Zohlnhöfer et al., 1992). For the *peripheral* receptor concentration R_{tot_2} we make the assumption that it is at least the concentration of mIL-6Rα expressed in *peripheral* fluid T cells, 0.0347 nM (Bongioanni et al., 2000; Mascio et al., 2009; Zola and Flego, 1992). The internalisation parameter $k_{R_{int}}$ defines the first-order rate at which the receptor-antibody complex internalises and is removed from the system, 0.35 h^{-1} (Fujimoto et al., 2015). The MM reaction scheme was previously shown as an adequate approximation for TMDD and receptor-mediated endocytosis (RME) (Krippendorff et al., 2009; Yan et al., 2009).

ADMINISTRATION OF MABS IS FIRST-ORDER: Administration of the antibody was modelled assuming a first-order reaction whereby the antibody is dosed into a theoretical depot pool $[Ab_{depot}]$ which is then absorbed into the *serum* through a first order reaction. For subcutaneous administration a parameter *bioavailability* describes the bioavailability of the administered dose compared with intravenous (IV) administration. A value for Bioavailability taken from literature adequately fits the experimental data. Likewise, an absorption rate within ranges measured by clinical studies was applied here.

VOLUMES ARE PHYSIOLOGICAL: The *serum* or central compartment volume is 3.5 l and is similar to physiological blood plasma volumes. It fits with the IV administration PK data of TCZ in humans and is a reasonable central volume for IgG-class mAbs. Two tissue compartments are explicitly defined within the scope of the model, the *liver* and the *synovium*. The *synovium* volume is an estimated total of synovial joint volume within arthritic humans (Buckwalter, 2007; R.W et al., 2007; Simkin et al., 1995). The *liver* volume was derived from observations that total liver volume lies between one and two l (Heinemann et al., 1999; Kan

and Hopkins, 1979; Kwon et al., 2001). The remaining volume of distribution is attributed as miscellaneous tissues and was roughly derived from the total interstitial fluid (ISF). The final volume of 11.25 l in the *peripheral* compartment accounts for subtracting the synovial tissue fluid as well as the 30 per cent extracellular fluid fraction within the liver.

A SINGLE MACRO PARAMETER DESCRIBES DISTRIBUTIONS: The flow between the central and tissue compartments was described using a global parameter Q from which the first-order rate constants k_{12}, k_{21}, k_{13}, k_{31}, k_{14} and k_{41} were derived.

SYNOVIAL PERMEABILITY LIMITS DISTRIBUTION: The transfer rate is assumed to be equal in all tissues apart from the *synovium*. There may be impedance across vascular barriers to the *synovium*, poor vascular perfusion or an efflux process significantly larger than influx. The rate of distribution of molecules through the *synovium* is determined largely by the molecular radius of the molecule in question due to putative pore structures. Previous work to assess the effect of pore-size is seen in a two-pore model which describes two populations of pores which traverse the vascular barrier (Rippe and Haraldsson, 1994). Further work describes a three-pore model, which includes vascular fenestrae, and appears to adequately describe synovial distribution data for a large range of molecular radii (Simkin, 2014). This model describes how large pores are non-selective for radius and smaller pores select against increasingly large molecules. Therefore smaller radius molecules are capable of diffusion through both sets of pores whilst larger macromolecules are limited in diffusion rates through small pores. In context, IgG has a hydrodynamic radius of 5.3 nM (Armstrong et al., 2004) which is relatively large and has a significantly limited small-pore diffusion rate. Other evidence shows that the rate at which a macromolecule the size of albumin is cleared from the *synovium* is twice as fast as the rate at which it enters (Wallis and Simkin, 1983). This means that we consider two factors to alter synovial distribution, the overall permeability or perfusion, and the rate at which the antibody leaves the synovium. The assumption is made that the exit of TCZ occurs twice as fast as the entry and that a perfusion coefficient V_{4_p} further limits the distribution by factoring perfusion and diffusion barriers.

PARAMETER ESTIMATION STEPS   To parametrise a four-compartment model we split a two-compartment model incrementally into a four compartments in three steps by fitting to clinical TCZ data (Choy et al., 2000; Zhang et al., 2013a) (See Figure 3.2). Progressive partitioning of the *peripheral* compartment into the final four-compartment
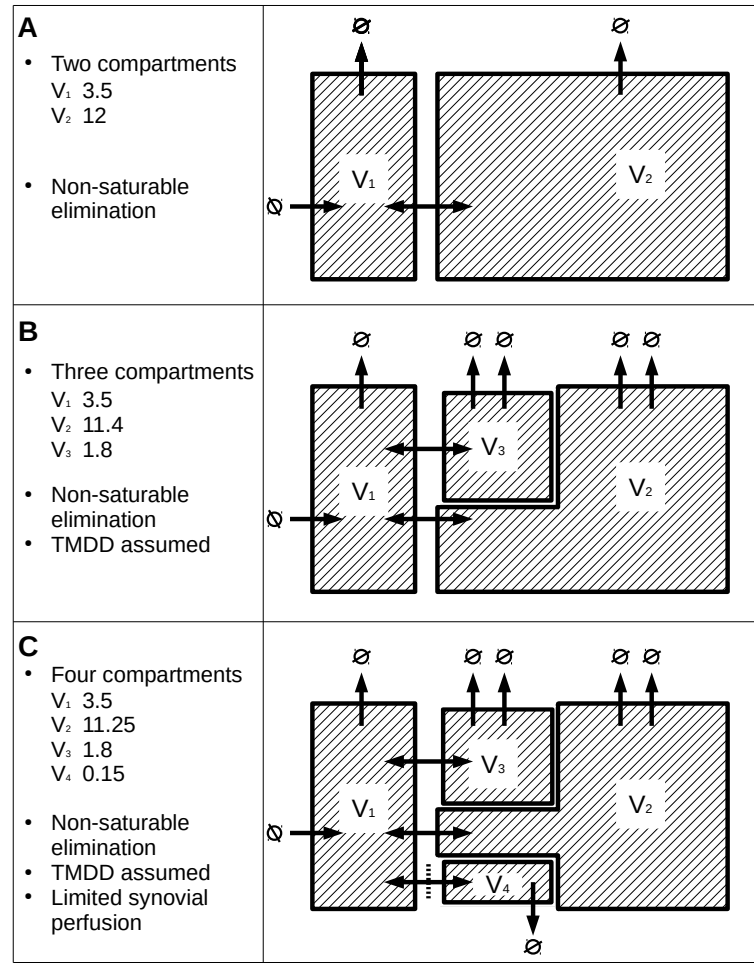
Figure 3.2: A two compartment PK model (A) was split into three (B) incorporating literature-derived values for TMDD parameters within both the liver and whole body T cell estimates. The fourth compartment (C) was added by including the synovial compartment. Estimates were made across the same three clinical data sets (Choy et al., 2000; Zhang et al., 2013a).

model was measured using AIC. The first model (A, Figure 3.2) did not fit to the *serum* PK data at later time-points, presumably due to the lack of a saturable elimination route. Additionally, no *synovium* compartment meant that the model would not fit to the given synovial data. We improved upon this by assuming the presence of a saturable TMDD elimination route in the *peripheral* compartment as well as an additional *liver* compartment using experimentally-derived concentrations of mIL-6Rα (Tables 3.1 and 3.2). While this PK model succeeded in fitting *serum* PK data for TCZ, the synovial data set was not modelled successfully. We further split the peripheral compartment and completed the four compartment structure by the addition of the *synovium*. This reduced the peripheral volume by 0.15. Estimation was carried out for a further parameter, V_{4_P} (Table 3.3).

Parameter	ID	Units	Value	Standard deviation	CV (%)
Intercompartmental clearance	Q	$l \cdot h^{-1}$	0.0595	0.0261	43.9

Akaike Information Criterion: 460.3

Table 3.1: Parameter estimates for the two-compartment PK model without nonlinear elimination.

Parameter	ID	Units	Value	Standard deviation	CV (%)
Intercompartmental clearance	Q	$l \cdot h^{-1}$	0.0301	0.0132	43.7

Akaike Information Criterion: 459.3

Table 3.2: Parameter estimates for the three-compartment PK model with nonlinear elimination.

Parameter	ID	Units	Value	Standard deviation	CV (%)
Intercompartmental clearance	Q	$l \cdot h^{-1}$	0.0356	0.00964	9.37
Synovial perfusion fraction	V_{4_p}	N/A	0.0662	0.0137	20.6

Akaike Information Criterion: 379.7

Table 3.3: Parameter estimates for the four-compartment PK model with nonlinear elimination and synovial perfusion constraints.

3.2.2.2   *Receptor dynamics*

Several receptor turnover datasets were available in two cell lines (Dittrich et al., 1996; Zohlnhöfer et al., 1992). However these are not entirely reminiscent of true hepatocytes. Furthermore, the application of a hepatic IL-6Rα signalling phenotypes to the other cells throughout the body is a large simplification. The assumptions in this module were that:

ACTIVE RECEPTOR STOICHIOMETRY IS SIMPLIFIED: The model assumed that receptor activation occurs independently of heterotrimer formation. The stoichiometry, mechanism and dimerisation status of gp130 complexes are still unclear in both IL-6-bound and unbound forms of the full receptor complex (Grötzinger et al., 1999; Scheller et al., 2011). The extent of or whether gp130 is preformed as a dimer or dimerisation occurs following receptor binding *in vivo* is yet to be confirmed (Tenhumberg et al., 2006). As such, we model an activation reaction with a 1:1:1 trimer (IL-6:IL-6Rα:gp130) reactant as simplified in *original model* Dwivedi et al., 2014. This bypasses controversy of the receptor

stoichiometry and approximates the gp130 dimerisation and phosphorylation as a single first-order reaction. The parameters for receptor activation were taken from the original model and the reaction assumes that intracellular gp130 tyrosine residues are readily phosphorylated upon assuming the active receptor conformation. A reverse reaction assumes the presence of intracellular phosphatases.

MIL-6Rα UNDERGOES FAST INTERNALISATION: The receptor, mIL-6Rα, has a relatively fast internalisation without the aid of gp130 binding as previously thought (Fujimoto et al., 2015). Therefore IL-6 can internalise through interaction with the receptor.

NO RECEPTOR RECYCLING OCCURS: IL-6 remains bound to the receptor and recovery of the cell-surface receptor population is slow after depletion (Wang and Fuller, 1994; Zohlnhöfer et al., 1992). Therefore the intracellular degradation reactions are absent from the whole model. This is because the model is only concerned with the dynamics of the receptors at the cell-surface.

PARAMETER ESTIMATION   The final parameter set for the *in vitro* receptor dynamics sub-model was estimated in two steps. Firstly, values for the zero-order gp130 synthesis parameter, first-order intracellular receptor degradation parameter, and the global low temperature modifier ($k_{gp130_{synth}}$, k_{deg} and T_4 respectively) were estimated in COPASI (Table 3.4, Figure 3.3a). First-order degradation parameters for all intracellular receptor components were assumed to be the identical. The temperature modifier T_4 was included to simulate reduced reaction rates prior to warming the cell-media. Estimation was carried out using data from a transfected HepG2 cell-line overexpressing mIL-6Rα (Zohlnhöfer et al., 1992). The initial receptor concentration was derived from the approximate number of IL-6 binding sites in the transfected HepG2 cells. Refinement of the estimation was carried out using a parameter scan to arrive at a suitable value for gp130 synthesis rates in mIL-6Rα-transfected HepG2 cells (Figure 3.3a).

Parameter	ID	Units	Value	Standard deviation	CV (%)
gp130 synthesis	$k_{gp130_{synth}}$	$nM\ h^{-1}$	2.91×10^5	5.18×10^7	1.78×10^4
Intracellular degradation	k_{deg}	h^{-1}	0.623	0.0390	6.26
Temperature modifier	T_4	N/A	0.0494	0.00700	14.1

Table 3.4: Parameter estimates for receptor dynamics in transfected HepG2 cells.
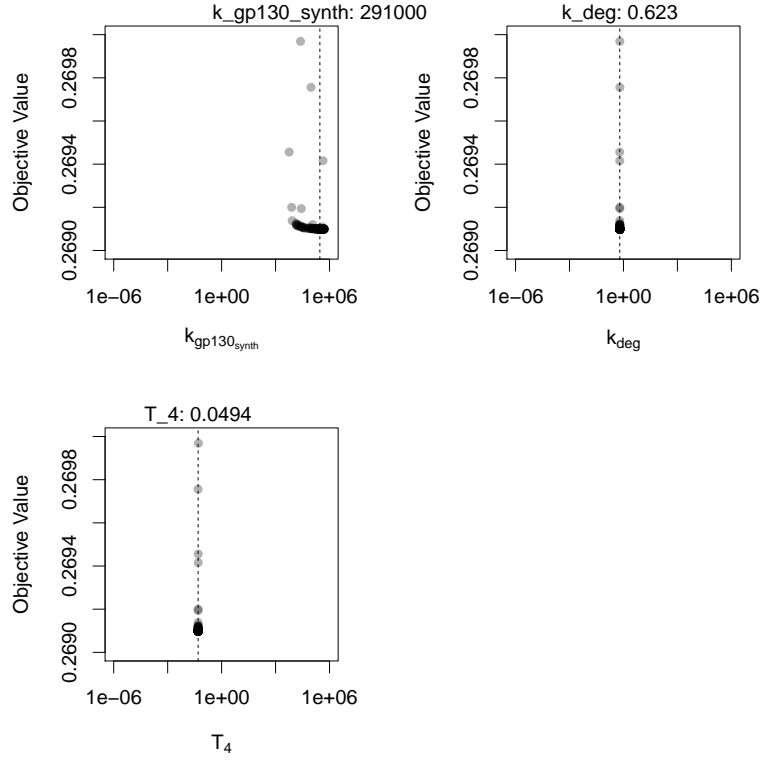
(a) Minimising the objective value by estimating $k_{gp130_{synth}}$, k_{deg} and T_4.
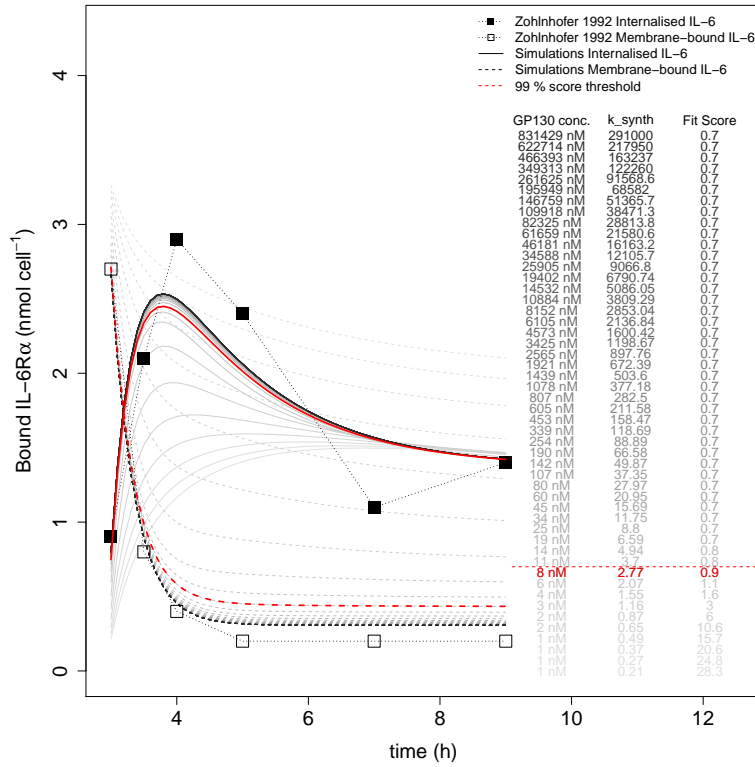
(b) A realistic biological range of gp130 expression is given by a small range of $k_{gp130_{synth}}$.

Figure 3.3: Parameter estimation of receptor dynamics with mIL-6R transfected HepG2 cells.

3.2.2.3 *Intracellular signalling and STAT3.*

The level of intracellular signalling presented in the model is only detailed enough to simulate phenomenological responses to IL-6. Thus, the model assumes that:

ONLY STAT3 IS CONSIDERED: Whilst multiple downstream pathways are perturbed in response to gp130 phosphorylation, this model assumes that the main response is captured as a function of STAT3. In simplifying the structure, the intricate intracellur pathways that regulate STAT activity are absent. For example, SOCS3 negative feedback was proposed to be ineffective in CD (Dwivedi et al., 2014; Lovato et al., 2003). Interestingly, the expression of SOCS3 may be decreased in RA T cells, providing a case for this assumption in the RA model (Ye et al., 2015). This adequately provides a clinical end-point within synovial tissue and the clinical biomarker within the hepatic tissue through the secretion of CRP.

STAT3 DIMERISATION IS OMITTED: Evidence suggests that STAT3 dimerisation may not be wholly necessary for activity (Cimica et al., 2011). For simplicity, the dimerisation step was not included in the *original model*.

AVAILABLE STAT3 IS CONSTANT: The sum of both STAT3 and pSTAT3 is assumed to be constant, that a steady-state of synthesis and degradation is present within the cell population that is not immediately perturbed by either the activity of IL-6 or other components within the model. This may not necessarily be the case in biology (Henkel et al., 2011), however, the simplification is necessary due to the lack of useful data and the level of model scope. The concentration of STAT3 was estimated from the number of molecules found in porcine heart cells for a general approximation of the concentration (Equation 3.2). This assumed that there were between 73 and 123 STAT3 molecules per cell. With respect to hepatocytes, the cell volume was 2.8 *pl*.

$$[STAT3] + [pSTAT3] = 98/((2.8 \times 10^{-12}) \cdot 6.02 \times 10^{23}) \quad (3.2)$$

Both tissue and inter-organism effects may vastly alter the concentration in real human physiology, and pragmatically speaking, 98 is a very low number of molecules per cell. Alternatively, evidence suggests that the relative expression of of STAT6 could be upwards of 10 000 molecules per cell (Raia et al., 2011). However, no quantitative contrary evidence was readily available for STAT3 and which remains a caveat in this module.

PHOSPHORYLATION IS REVERSIBLE: A scheme of reversible Michaelis–Menten STAT3 phosphorylation reactions is presented in a

previous model (Sadreev et al., 2014). The forward reaction is catalysed by the active receptor complex while the reverse is driven by an implicit presence of STAT3 phosphatases.

STAT3 PHOSPHORYLATION IS ONLY IL-6 DEPENDENT: The reaction structure does not account for IL-6-independent pSTAT3 reactions, the model only assesses the effect of IL-6. It is assumed that the other mechanisms are unaffected by perturbation in IL-6-signalling pathways and are therefore excluded from the model.

PARAMETER ESTIMATION   Reactions for the reversible phosphorylation of STAT3 were added to the receptor dynamics model. The MM reaction rate parameters were estimated simultaneously using data from multiple cell-lines (human hepatocyte-derived cellular carcinoma cell line (HuH7), tongue cell carcinoma cell line (Cal33), & murine hybridoma cell line (7TD1)). The data were digitised and converted from published studies (Casanovas et al., 2014; Gough et al., 2016; Simard et al., 2014).

Parameter	ID	Units	Value	Standard deviation	CV (%)
Catalytic rate of phosphorylation	$k_{cat}STAT_{phos}$	h^{-1}	5280	4.23×10^7	9.17×10^3
STAT3 phosphorylation MM constant	$K_M STAT_{phos}$	nM	0.0164	0.0727	412
Maximal rate of dephosphorylation	$V_m STAT_{dephos}$	$nM\ h^{-1}$	144000	9.09×10^7	1.08×10^4
pSTAT3 dephosphorylation MM constant	$K_M STAT_{dephos}$	nM	31.4	195	8.56×10^3

Table 3.5: Parameter estimates for STAT3 phosphorylation reactions. MM, Michaelis-Menten; CV, coefficient of variation. Parameter values are median values from parameter sets within 0.1 % of the lowest objective value.

3.2.2.4 *Tissue compartments, CRP and steady-state concentrations*

Connection of all modules was carried out and the model assumed its final structure. Assumptions were made to simplify this process and ease the parameter estimation and retain identifiability. The final model assumed that:

COMPONENTS WERE SOURCED AND DEGRADED IN TISSUES: Evidence suggests that tissue compartments are major contributors to the production and degradation of trans-signalling components. The rate of sIL-6Rα and sgp130 degradation is similar to other soluble cytokine receptors (Jacobs et al., 1993).

RA IS MAINTAINED BY LOCAL SECRETIONS: Synovial synthesis of signalling components in RA creates a local environment optimal for maintaining the arthritic state. High concentrations IL-6 are presumed to be responsible for the pathology of RA. Its increase in disease is assumed to be driven largely by local synovial secretions (Guerne et al., 1989; Rosenbaum et al., 1992) Furthermore, a reaction for sIL-6Rα synthesis is included in the *synovium*. One modification is made to the *original model* whereby the *serum* IL-6 synthesis reaction is removed. Here the assumption is that the circulatory system secretes an insignificant portion in systemic IL-6 concentrations relative to the RA disease state.

T CELLS DRIVE SIL-6Rα SYNTHESIS: Production of sIL-6Rα is seen from activated CD4+ mononuclear cells (Briso et al., 2008). The model assumes that serum sIL-6Rα production is proportional to the compartmental concentration of T cells.

RECEPTOR SHEDDING IS DRIVEN BY CRP: Receptor shedding of sIL-6Rα occurs in the *serum* where CRP catalyses a shedding reaction. The *in vivo* contribution of this reactions has not been determined. The assumption is made that the increase in CRP between healthy and RA state causes the increase in sIL-6Rα in RA.

CRP SYNTHESIS IS IL-6 DEPENDENT: Similar to omitting IL-6-independent STAT3 phosphorylation, the model's scope is solely that of IL-6-mediated events. In this model, CRP production is simplified so that all CRP is sourced from hepatocytes via IL-6 signalling. The IL-6-independent CRP synthesis seen in the *original model* was removed.

CRP SYNTHESIS IS SIMPLIFIED: In the presence of pSTAT3, CRP synthesis was assumed to be an enzymatic Michaelis-Menten reaction with pSTAT3 as the catalytic enzyme. This simplification merges the steps of transcription and translation.

PROTEIN TISSUE CONCENTRATIONS ARE HIGHER: The *original model* defined that steady-state protein concentrations in tissue are two-fold higher than those in *serum* at steady-state. This applies to all components except those where experimental data suggests otherwise. A coefficient of 2 was applied to *serum* to tissue rate constants to reflect this.

SYNOVIAL PERFUSION ALSO LIMITS IL-6 COMPONENTS The synovial perfusion constant $V4_p$ used in the PK model was also used here for calculating the transfer between compartments. Furthermore, the smaller molecule sizes of the majority of com-

ponents may require the removal of the 2-fold modifier seen in k_{41} as seen in the PK model.

BOTH ACTIVE RECEPTOR COMPLEXES ARE INDEPENDENT: In contrast to the original model, a reaction was implemented which separates both active forms of the receptor. The *original model* allowed for a flux from [mR_IL6_gp130] to [sR_IL6_gp130] via the receptor activation reaction which is unlikely to occur in biology at a great rate. This would only be explainable if the whole membrane-anchored receptor complex was cleaved after inactivation. Hence in this model, an additional reaction was implemented for the proper separation of cis and trans-signalling active receptor complexes.

PARAMETER ESTIMATION   In total, 15 unknown value, kinetic rate constants were estimated with published data. These included zero-order synthesis and first-order constants for molecular species as well as MM constant for protein synthesis in CRP production and the inter-compartmental flow rates for PK and PD components (Table 3.6). This was carried out in a single step with the full model in an optimisation task fitting to steady-state data as well as dynamic drug response data. The steady-state concentrations of molecular species were collected from a number of studies in RA patients (table 3.7).

See figure 3.9, section 3.3.1.4 for the figure exploring parameter space.

Parameter	ID	Units	Value	Standard deviation	CV (%)
PD distribution flow	Q_{PD}	$l\,h^{-1}$	0.255	0.0154	5.82
PK distribution flow	Q	$l\,h^{-1}$	0.0354	9.71×10^{-4}	2.72
Soluble mAb complex degradation	$k_{Ab\_sR_{deg}}$	h^{-1}	9.76×10^{-3}	5.11×10^{-4}	5.33
Synovial IL-6 synthesis	$k_{IL6\_synovium_{synth}}$	$nM\,h^{-1}$	0.226	8.38×10^{-3}	3.75
Systemic IL-6 synthesis	$k_{IL6_{synth}}$	$nM\,h^{-1}$	0.0165	4.29×10^{-4}	2.62
Liver sIL-6Rα synthesis	$k_{sR\_liver_{synth}}$	$nM\,h^{-1}$	0.365	0.0183	5.04
Peripheral sIL-6Rα synthesis	$k_{sR\_peripheral_{synth}}$	$nM\,h^{-1}$	0.361	8.84×10^{-3}	2.45
Synovial sIL-6Rα synthesis	$k_{sR\_synovium_{synth}}$	$nM\,h^{-1}$	0.291	0.0189	6.51
Liver sgp130 synthesis	$k_{sgp130\_liver_{synth}}$	$nM\,h^{-1}$	1.23	0.0776	6.23
Peripheral sgp130 synthesis	$k_{sgp130\_peripheral_{synth}}$	$nM\,h^{-1}$	1.00	0.0350	3.50
Serum sgp130 synthesis	$k_{sgp130\_serum_{synth}}$	$nM\,h^{-1}$	0.0986	0.0595	66.3
Synovial sgp130 synthesis	$k_{sgp130\_synovium_{synth}}$	$nM\,h^{-1}$	2.44×10^{-6}	2.62×10^{-3}	1.17×10^{4}
Protein synthesis rate max	$V_{m}prot_{synth}$	$nM\,h^{-1}$	1.36×10^{8}	5.45×10^{9}	8.87×10^{3}
Protein MM constant	$K_{M}prot_{synth}$	nM	1.51×10^{4}	2.72×10^{6}	8.75×10^{3}
mIL-6Rα shedding rate	$k_{R_{shedding}}$	h^{-1}	1.02×10^{-6}	6.98×10^{-5}	6.71×10^{3}

Table 3.6: Estimated parameter values for the whole model. These were derived by fitting to experimental temporal PK and PD data as well as RA steady-state concentrations of IL-6 signalling components.

	Tissue concentration (nM unless stated)						
	$IL6_1$ (pM)	$IL6_4$ (pM)	CRP_1	$sIL6Ra_1$	$sIL6Ra_4$	$sgp130_1$	$sgp130_4$
mean (SD)	6.24 (6.39)	137 (85.6)	245 (171)	1.36 (1.16)	1.50 (1.05)	4.34 (0.424)	2.78 (NA)
median	2.45	125	230	0.907	1.337	4.34	2.78
Source							
A	–	14.9	–	–	–	–	–
B	–	11.0	–	–	–	–	–
C	2.43	–	–	0.995	–	–	–
D	2.34	–	–	–	–	–	–
E	–	–	–	–	0.714	–	–
F	–	2202	–	–	0.57	–	2.78
G	–	409	–	–	2.76	–	–
H	2.11	70.3	112	3.08	1.96	–	–
I	2.46	–	432	0.554	–	–	–
J	10.8	–	–	0.818	–	4.64	–
K	–	–	–	–	–	4.04	–
L	–	–	88	–	–	–	–
M	17.3	179	348	–	–	–	–

A, Kokebie et al. (2011); B, Tsuchida et al. (2012); C, Robak et al. (1998); D, Manicourt et al. (1993); E, Nowell et al. (2003); F, Richards et al. (2006); G, Kotake et al. (1996); H, Kohno et al. (1998); I, Nishimoto et al. (2008); J, Spano et al. (2011); K, Pignatti et al. (2003); L, Zhang et al. (2013b); M, Okamoto et al. (1997).

Table 3.7: Steady-State disease concentrations of IL-6 signalling components. The subscript integer refers to compartment number ($_1$), serum; ($_4$), synovium.

3.3   RESULTS

The results focus upon several aspects of developing QSP models. Firstly, the model re-parametrisation is presented in modules, detailing the results of fits to data and the choice parameter set. Secondly, the resulting RA model's validity is shown through simulation, compared with experimental data both at equilibrium and with respect to time covering a variety of experiments. Finally, results of predictions are shown where the model is used to address the posed questions. In this last section, simulations are made to predict the action of SRK and TCZ over several dose regimens.

3.3.1   *Parametrisation*

As described in the methods, parametrisation of the model was carried out in modules: mAb PK, cell-surface receptor dynamics, and STAT3 phosphorylation. These modules were separated from the overall model by fixing their inputs and the whole model was pieced together over several estimation steps to maximise module identifiability. To determine whether the parameter values chosen were indeed suitable for use in the model or identifiable, multiple iterations of each estimation task were carried out for each estimation step

throughout the model development. Without performing a bonefide identifiability analysis, we used the distributions of lowest objective function values to declare a parameter as identifiable with our current model structure and data. For example, Figure 3.9 showed that even in parameter sets that fit our data most closely, parameters may take a very wide range of values for an equally good fit. This type of distribution suggests that either the experimental data or the model were not fully descriptive of the processes occurring in biology. Parameters with a wide distribution across best fitting parameter sets can be considered unidentifiable as they have no identified value. On the other hand, we can call parameters with well-defined minima as identifiable and we can confidently determine their value within our model.

3.3.1.1 *Monoclonal antibody pharmacokinetics*

In striving to achieve the *original model*'s reaction structure, the PK reactions were re-parametrised to fit experimental IgG-class mAb PK data. This produced a four-compartment model comprised of the central circulatory compartment *serum* and the tissue compartments of the *liver* and *synovium* as well as a *peripheral* tissue compartment. The tissues were connected via the *serum* which allowed for the transfer of soluble IL-6 signalling components between the tissues. The reaction scheme and diagram can be seen in Table 3.8. To reconstruct this PK model, a two-compartment model of physiological volumes was progressively split into further compartments, accommodating the *liver* and the *synovium* tissues necessary for the scope of the model. During this process, assumptions were made about the PK of TCZ and other IgG mAbs which increased the model complexity and the fit to IgG-mAb data. For example, it was apparent in IV data that TCZ exhibited nonlinear elimination and that this would have to be accounted for in a more complex model. The resulting PK model and parametrisation was able to fit to the kinetics of TCZ in several dose regimens and administration routes (Figure 3.4). The fit of the PK model to the experimental data for TCZ shows nonlinearity which was parametrised using mIL-6Rα concentrations in the *liver* and *peripheral* compartments calculated from published experiments (Figure 3.4A).
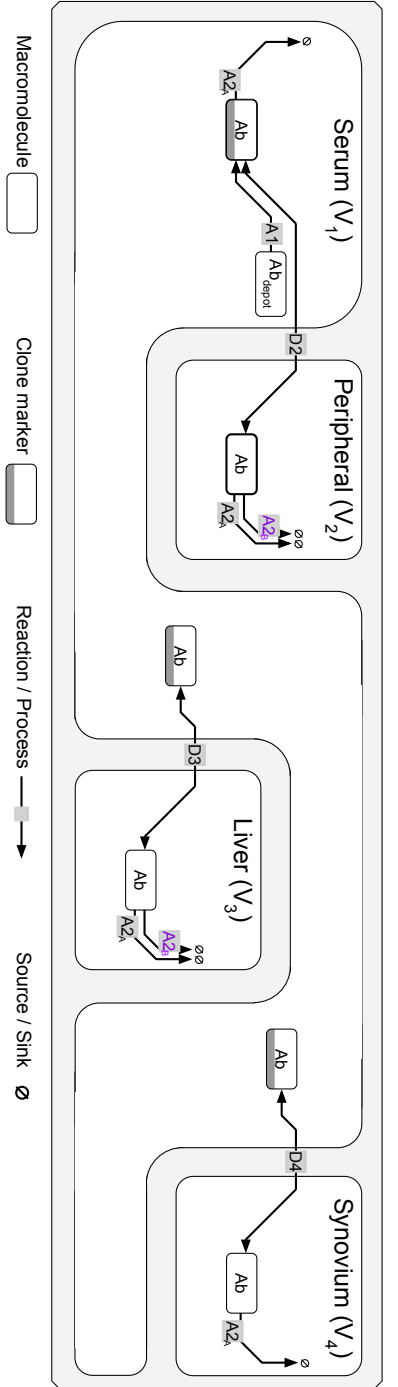
Description	ID	Reaction	Rate
Antibody absorption	serum A1	$[Ab_{depot}] \to [Ab_1]$	$k_{abs} \cdot [Ab_{depot}]$
Linear elimination	serum A2$_A$	$[Ab_1] \to \varnothing$	$k_{el} \cdot [Ab_1]$
Linear elimination	peripheral A2$_A$	$[Ab_2] \to \varnothing$	$k_{el} \cdot [Ab_2]$
Linear elimination	liver A2$_A$	$[Ab_3] \to \varnothing$	$k_{el} \cdot [Ab_3]$
Linear elimination	synovium A2$_A$	$[Ab_4] \to \varnothing$	$k_{el} \cdot [Ab_4]$
Nonlinear elimination	peripheral A2$_B$	$[Ab_2] \to \varnothing$	$\dfrac{V_{max_2} \cdot [Ab_2]}{K_M + [Ab_2]}$
Nonlinear elimination	liver A2$_B$	$[Ab_3] \to \varnothing$	$\dfrac{V_{max_3} \cdot [Ab_3]}{K_M + [Ab_3]}$
Distribution to peripheral	D2$_f$	$[Ab_1] \to [Ab_2]$	$k_{12} \cdot [Ab_1]$
Distribution from peripheral	D2$_r$	$[Ab_2] \to [Ab_1]$	$k_{21} \cdot [Ab_2]$
Distribution to liver	D3$_f$	$[Ab_1] \to [Ab_3]$	$k_{13} \cdot [Ab_1]$
Distribution from liver	D3$_r$	$[Ab_3] \to [Ab_1]$	$k_{31} \cdot [Ab_3]$
Distribution to synovium	D4$_f$	$[Ab_1] \to [Ab_4]$	$k_{14} \cdot [Ab_1]$
Distribution from synovium	D4$_r$	$[Ab_4] \to [Ab_1]$	$k_{41} \cdot [Ab_4]$

1, serum; 2, peripheral; 3, liver; 4, synovium

Table 3.8: Final four-compartment PK model of the monoclonal antibody, and the list of corresponding reactions.
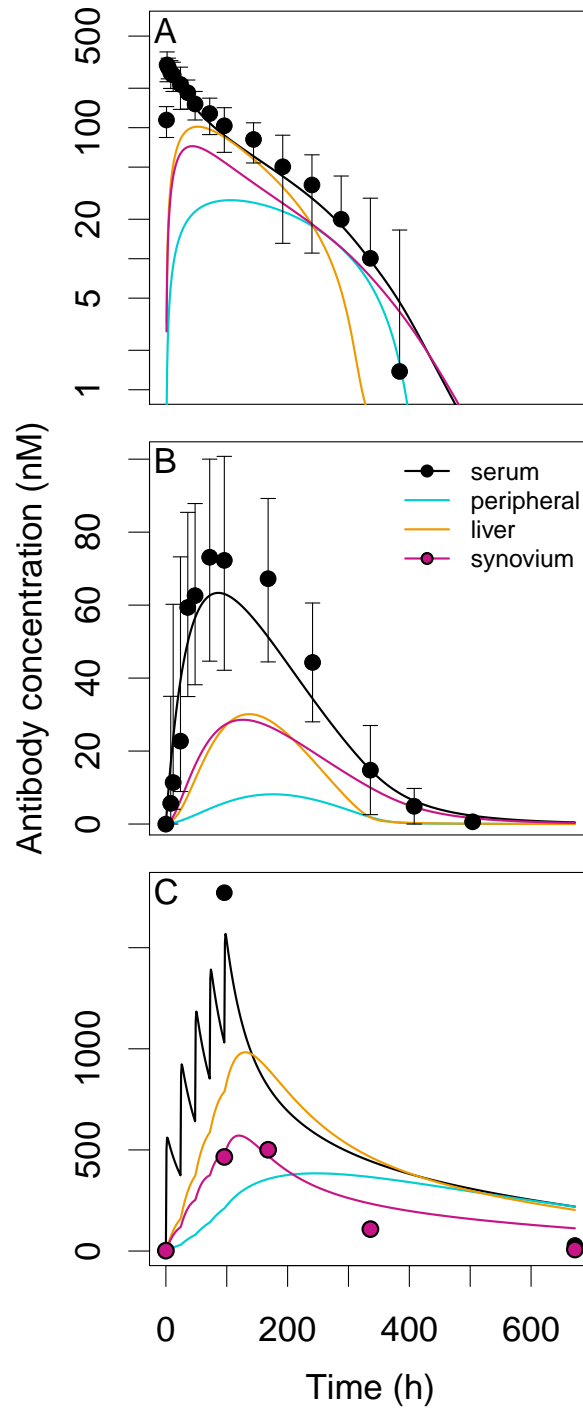
Figure 3.4: PK trajectories for the final four-compartment model. The above trajectories are experimental data points used in fitting the PK model alongside simulated events for **A**, single-dose IV administration of 162 mg TCZ; **B**, single-dose SC administration of 162 mg TCZ (Zhang et al., 2013a); and **C**, five-dose IV administration of 300 mg anti–CD4+ mAb over 120 h (Choy et al., 2000).

3.3.1.2 *Cell-Surface receptor dynamics*

The core reactions that govern IL-6 signalling are related to the regulation of cell-surface receptors and the binding of IL-6 to these receptors. It is within this network that the proposed mAbs bind and exert their perturbation. The structure of the model accounts for the dynamic turnover of both mIL-6Rα and gp130 as well as receptor activation. Receptor components are synthesised, internalised and degraded, establishing a cell-surface equilibrium. Perturbations caused by IL-6 binding and changes in temperature shift the expression equilibrium. The first step was to use parameters fully derived from experimental sources in literature. This provided an approximate fit to experimental internalisation data (Figure 3.6a). However, the simulation depicted a much lower peak of mIL-6Rα internalised mIL-6Rα than at least that seen experimentally in the HepG2 cell-culture. Subsequent estimation of gp130 synthesis rates and degradation rates of intracellular receptor components resulted in a much greater fit to the data (Figure 3.6b). Emerging from this model parametrisation was the observation that, after the removal of IL-6, the cell-surface expression of mIL-6Rα was nearly restored after 8 hours (Figure 3.6c). This was also seen in published experiments (Zohlnhöfer et al., 1992). Details of the parameter estimation figures can be seen in the materials and methods section.
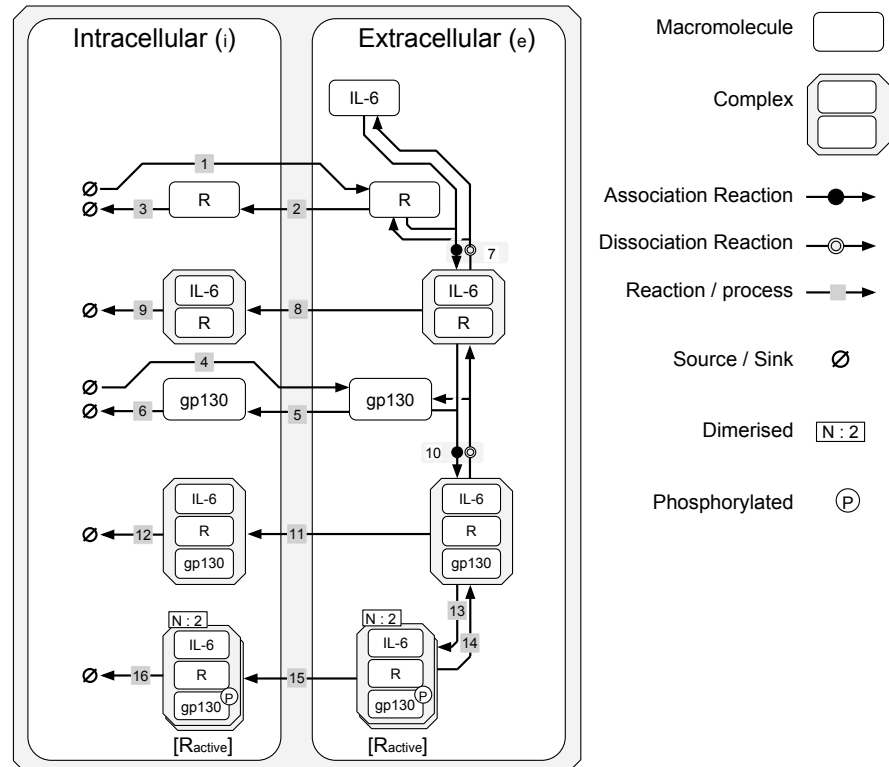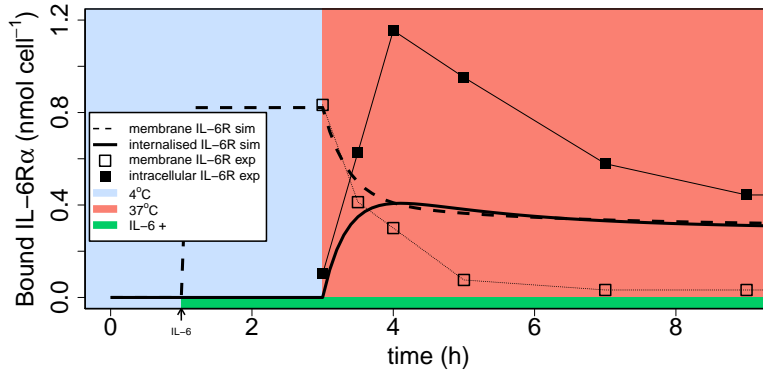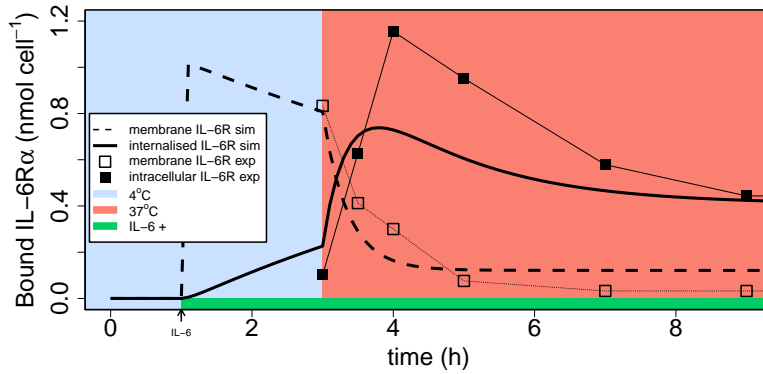

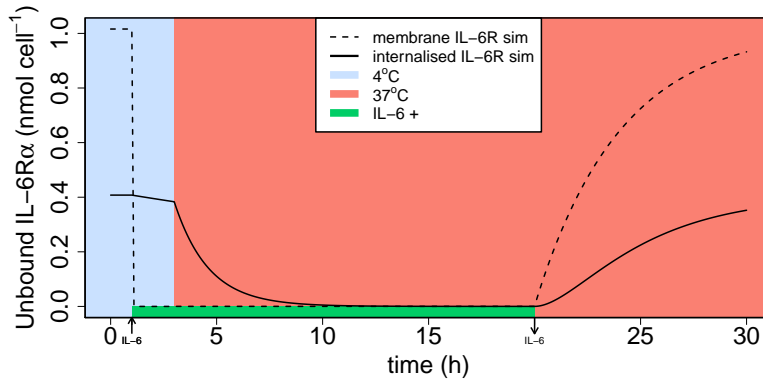Figure 3.5: Diagram of the receptor dynamics model. Reaction IDs are independent of those in the full model.

The complete receptor dynamics module was capable of simulating the dynamic turnover of IL-6 receptor components in cell culture to a reasonable semi-quantitative degree (Figure 3.6).

(a) Parameters set derived from experimental findings in literature.

(b) Estimated gp130 expression and receptor degradation parameters.

(c) Recovery of cell-surface mIL-6Rα expression after removing IL-6.

Figure 3.6: Reproducing an experiment of wild-type HepG2 cells using the receptor dynamics model (Zohlnhöfer et al., 1992). Model simulations are denoted by "sim" while experimental data is referred to as "exp". Internalisation dynamics of mIL-6Rα were observed after heating and exposure to IL-6.

Description	ID	Reaction	Rate
Receptor synthesis	1	$\emptyset \to [R_e]$	$k_{R_{synth}}$
Receptor internalisation	2	$[R_e] \to [R_i]$	$k_{R_{int}} \cdot [R_e]$
Receptor degradation	3	$[R_i] \to \emptyset$	$k_{R_{deg}} \cdot [R_i]$
Gp130 synthesis	4	$\emptyset \to [gp130_e]$	$k_{gp130_{synth}}$
Gp130 internalisation	5	$[gp130_e] \to [gp130_i]$	$k_{gp130_{int}} \cdot [gp130_e]$
Gp130 degradation	6	$[gp130_i] \to \emptyset$	$k_{gp130_{deg}} \cdot [gp130_i]$
IL-6 association	7f	$[IL6] + [R_e] \to [R\text{-}IL6_e]$	$k_{IL6_{on}} \cdot [IL6] \cdot [R_e]$
IL-6 dissociation	7r	$[R\text{-}IL6_e] \to [IL6] + [R_e]$	$k_{IL6_{off}} \cdot [R\text{-}IL6_e]$
R-IL6 internalisation	8	$[R\_IL6_e] \to [R\_IL6_i]$	$k_{R\text{-}IL6_{int}} \cdot [R\text{-}IL6_e]$
R-IL6 degradation	9	$[R\text{-}IL6_i] \to \emptyset$	$k_{R\text{-}IL6_{deg}} \cdot [R\text{-}IL6_i]$
gp130 association	10f	$[R\text{-}IL6_e] + [gp130_e] \to [R\text{-}IL6\text{-}gp130_e]$	$k_{gp130_{on}} \cdot [R\text{-}IL6_e] \cdot [gp130_e]$
gp130 dissociation	10r	$[R\text{-}IL6_e] + [gp130_e] \to [R\text{-}IL6\text{-}gp130_e]$	$k_{gp130_{off}} \cdot [R\text{-}IL6\text{-}gp130_e]$
Trimer internalisation	11	$[R\text{-}IL6\text{-}gp130_e] \to [R\text{-}IL6\text{-}gp130_i]$	$k_{R\text{-}IL6\text{-}gp130_{int}} \cdot [R\text{-}IL6\text{-}gp130_e]$
Trimer degradation	12	$[R\text{-}IL6\text{-}gp130_i] \to \emptyset$	$k_{R\text{-}IL6\text{-}gp130_{deg}} \cdot [R\text{-}IL6\text{-}gp130_i]$
Receptor activation	13	$[R\text{-}IL6\text{-}gp130_e] \to [R_{active_e}]$	$k_{R_{activation}} \cdot [R\text{-}IL6\text{-}gp130_e]$
Receptor deactivation	14	$[R_{active_e}] \to [R\text{-}IL6\text{-}gp130_e]$	$\frac{V_m R_{dephos} R_{active}}{K_M R_{dephos} + R_{active}}$
Active R internalisation	15	$[R_{active_e}] \to [R_{active_i}]$	$k_{R\text{-}IL6\text{-}gp130_{int}} \cdot [R_{active_e}]$
Active R degradation	16	$[R_{active_i}] \to \emptyset$	$k_{R\text{-}IL6\text{-}gp130_{deg}} \cdot [R_{active_i}]$

Table 3.9: The reactions present in the receptor dynamics module. This model assumed that receptor activation occurs following trimerisation and does not refer to a specific complex. Reaction IDs in this module are independent of those seen in the whole model and refer instead to the diagram seen in Figure 3.5.

Description	Name	Value	Units	Source
gp130 synthesis	$k_{gp130_{synth}}$	2.77 (0.29)	nM h⁻¹	†
basal gp130 internalisation	$k_{gp130_{int}}$	0.35	h⁻¹	(Blanchard et al., 2001; Thiel et al., 1998)
mIL-6Rα synthesis	$k_{R_{synth}}$	0.25 (0.21)	nM h⁻¹	†
basal mIL-6Rα internalisation	$k_{R_{int}}$	0.25	h⁻¹	(Fujimoto et al., 2015)
IL-6 association	$k_{IL6_{on}}$	1.0	nM⁻¹ h⁻¹	(Dittrich et al., 1994; Gearing et al., 1992; Heinrich et al., 1998; Weiergraber et al., 1995)
IL-6 dissociation	$k_{IL6_{off}}$	0.5	h⁻¹	(Dittrich et al., 1994; Gearing et al., 1992; Heinrich et al., 1998; Weiergraber et al., 1995)
IL-6 driven mIL-6Rα internalisation	$k_{R-IL6_{int}}$	0.35	h⁻¹	(Fujimoto et al., 2015)
gp130 association	$k_{gp130_{on}}$	1.0	nM⁻¹ h⁻¹	(Jostock et al., 2001; Richards et al., 2006; Schroers, 2005)
gp130 dissociation	$k_{gp130_{off}}$	0.05	h⁻¹	(Jostock et al., 2001; Richards et al., 2006; Schroers, 2005)
IL-6 driven gp130 internalisation	$k_{R-IL6-gp130_{int}}$	2.77	h⁻¹	(Dittrich et al., 1996; Nesbitt and Fuller, 1992; Zohlnhöfer et al., 1992)
mIL-6Rα degradation	$k_{R_{deg}}$	0.62 (0.69)	h⁻¹	†(Nesbitt and Fuller, 1992)
gp130 degradation	$k_{gp130_{deg}}$	0.62 (0.69)	h⁻¹	†(Nesbitt and Fuller, 1992)
trimer degradation	$k_{R-IL6-gp130_{deg}}$	0.62 (0.69)	h⁻¹	†(Nesbitt and Fuller, 1992)
Receptor activation	$k_{R_{activation}}$	155	h⁻¹	(Dwivedi et al., 2014)
Receptor dephosphorylation	$V_m R_{dephos}$	155.3	h⁻¹	(Dwivedi et al., 2014)
MM constant receptor dephosphorylation	$K_M R_{dephos}$	0.525	nM	(Dwivedi et al., 2014)

An additional parameter, T_4, was used in the model to describe the reaction rate at low temperature as a fraction of the rate at $37^o C$

Table 3.10: The parameters used in the final receptor dynamics module. † Estimated values resulting from initial parameter estimation (see Materials and Methods, section 3.2.2), calculation of mIL-6Rα binding sites in hepatocytes and manual estimation of gp130 (Figure 3.3). Values in parentheses show the parameter value derived from literature prior to estimation.

3.3.1.3 *Intracellular signalling and STAT3 phosphorylation.*

Phosphorylation of STAT3 was modelled using reversible Michaelis-Menten reactions and conveys the transduction of gp130 phosphorylation. This gives a phenomenological interpretation of a downstream response to IL-6. Estimation of the STAT3 response parameters was carried out by adding STAT3 reactions to the receptor dynamics model. The fitted model captures the sigmoidal curve of STAT3 phosphorylation seen with *in vitro* IL-6 dose–response experiments (Figure 3.7).
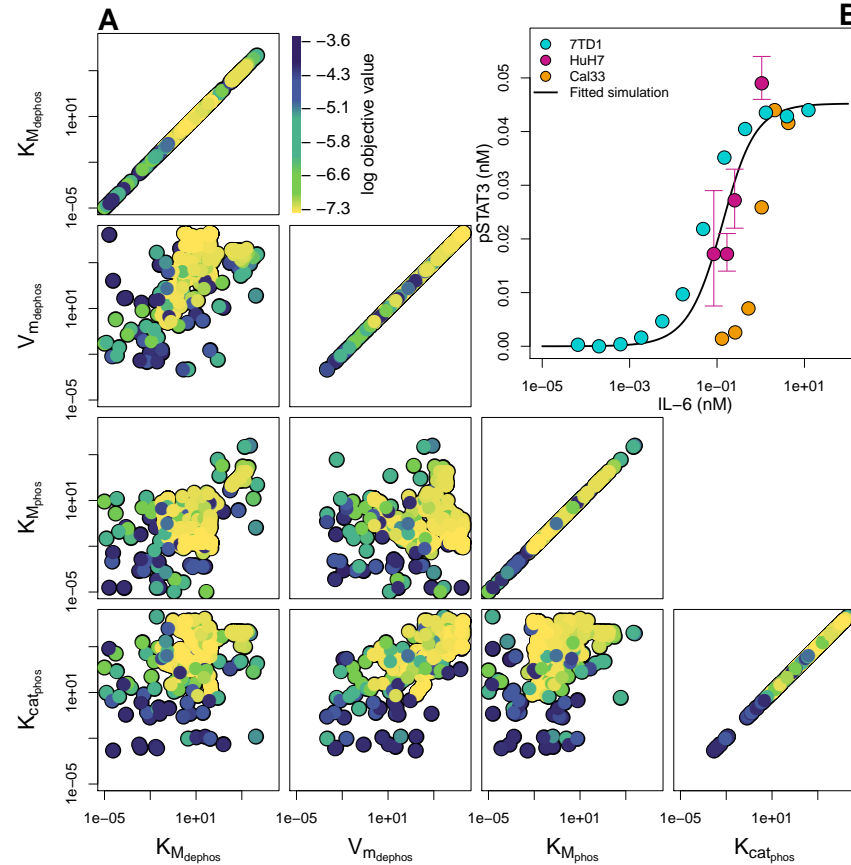

Figure 3.7: STAT3 model estimation and result. **A**, parameter correlation and convergence towards minima; **B**, steady-state dose–response of tissue STAT3 phosphorylation to a range of fixed IL-6 concentrations plotted with response data from HuH7, 7TD1 and Cal33 cells.

3.3.1.4 *Whole model fitting; CRP and steady-state concentrations*

Estimation of synthesis, inter-compartmental distribution and degradation rates of IL-6 signalling components in a final estimation step yielded a system that fits the dynamics of IL-6Rα blockade after TCZ administration, albeit semi-quantitatively (Figure 3.8A-E). Furthermore,

the steady-state concentrations resemble those derived from published studies and clinical trials with RA patients (Figure 3.8F).
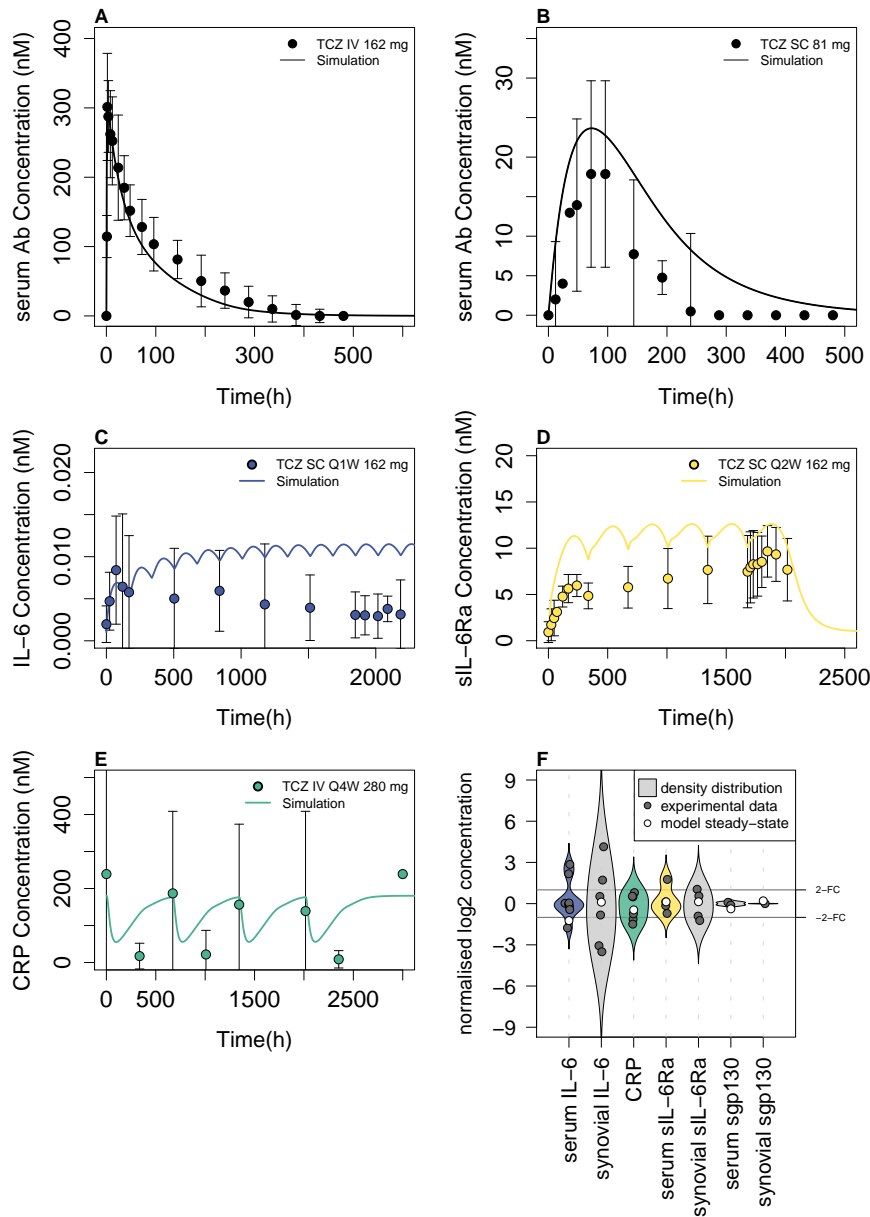

Figure 3.8: Fitted simulations of the whole model parameter estimation. Simulations were optimised to fit the experimental PK data for TCZ in **A**, single-dose IV administration (Zhang et al., 2013a); **B**, single-dose SC administration (Zhang et al., 2013a); as well as a combination of serum biomarker responses with **C**, IL-6 dissociation; **D**, sIL-6Rα elevation; and **E**, suppression of CRP (Zhang et al., 2013b). At the same time, the model was calibrated to fit **F**, the RA steady-state concentrations of IL-6 signalling components.

The estimation of parameter space for the whole model fitting including CRP and steady-state concentrations can be seen in Figure 3.9).
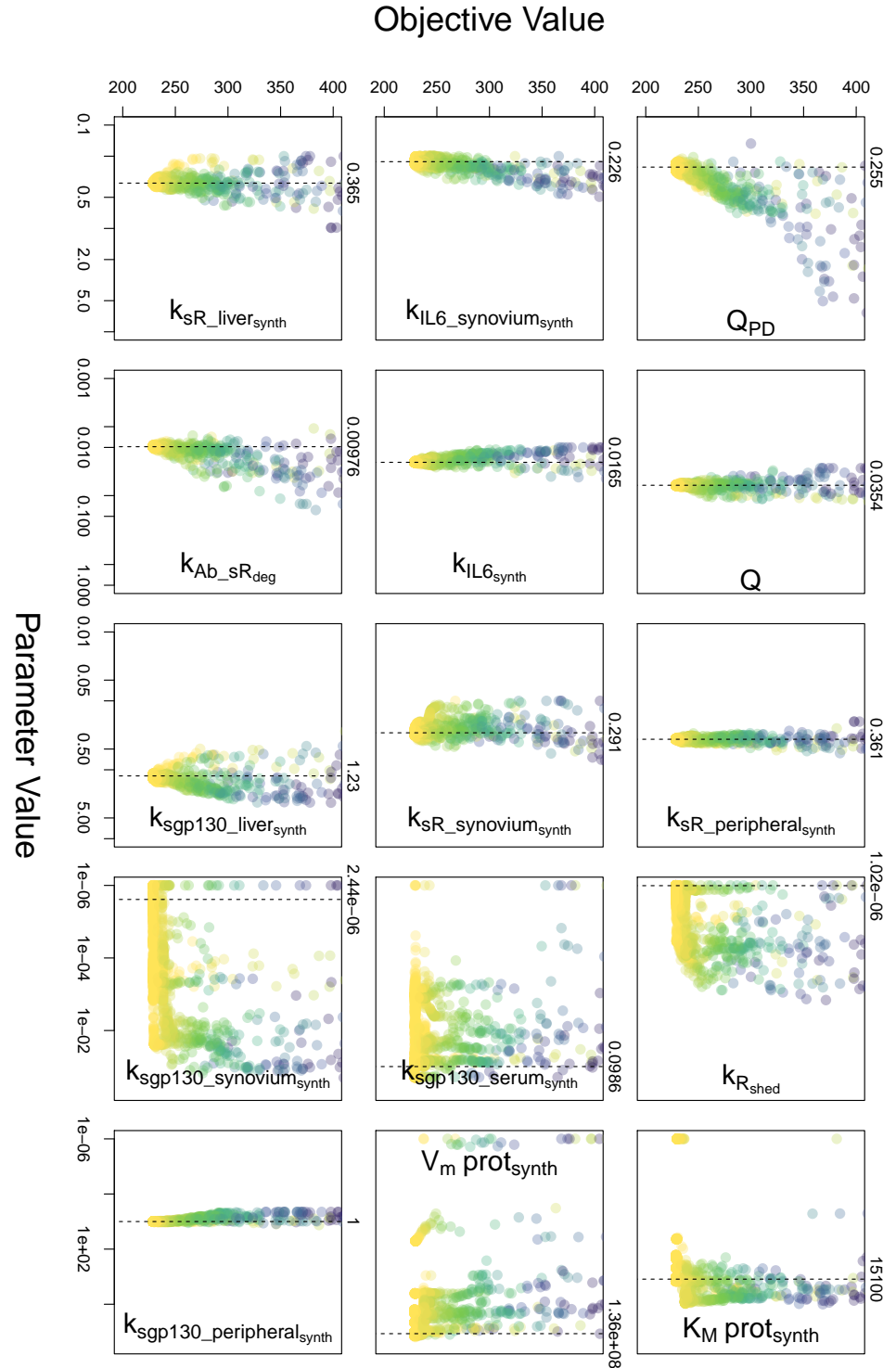

Figure 3.9: Estimation of synthesis and distribution rate constants for the whole model.

Additionally, the completed model, detailing reactions, species and compartments is represented in Figure 3.10 on the next double page with the corresponding reactions in Tables 3.11 to 3.17. The parameters can be found in Tables 3.18 and 3.19.

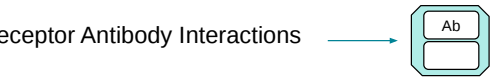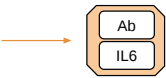Figure 3.10: IL-6 model schematic in SBGN.

Liver (V$_3$)

Synovium (V$_4$)
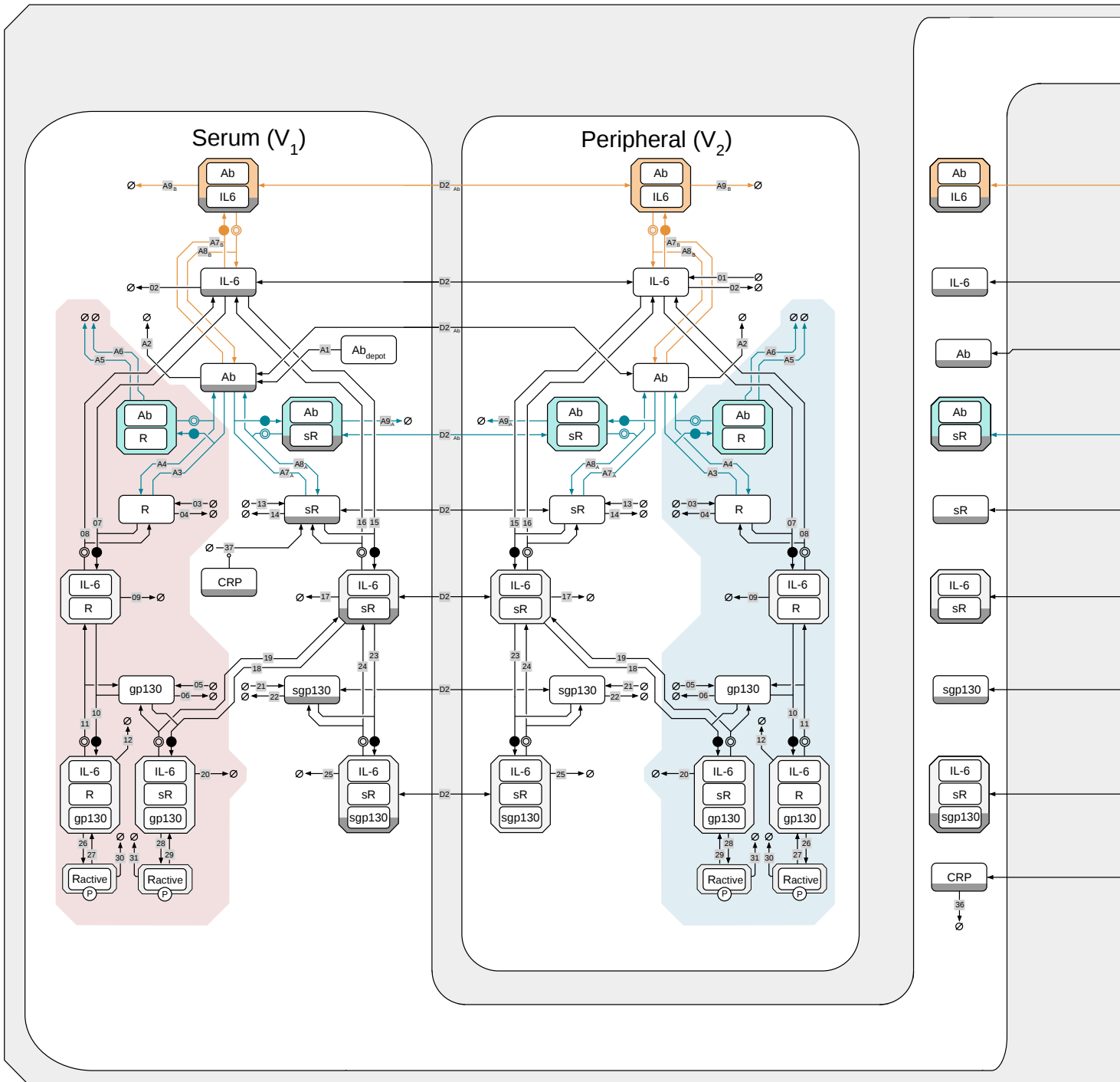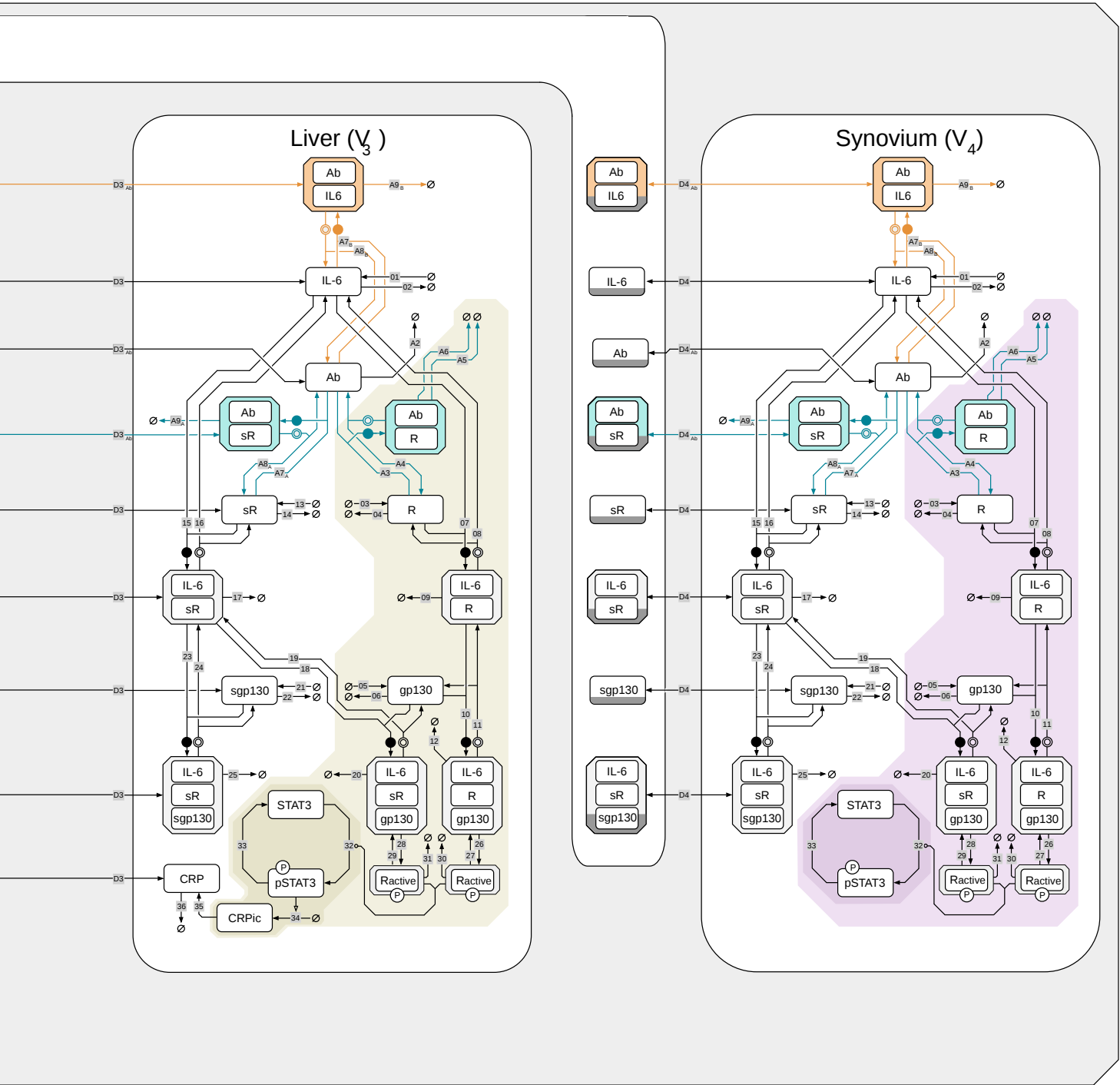
Source / Sink ∅

Phosphorylation Ⓟ

Receptor Antibody Interactions

Ligand Antibody Interactions

Description	ID	Reaction	Rate
IL-6 degradation	2	$[IL6_{serum}] \to \varnothing$	$k_{IL6_{deg}} \cdot [IL6_{serum}]$
mIL-6Rα synthesis	3	$\varnothing \to [R_{serum}]$	$k_{R_{serum\,synth}}$
mIL-6Rα internalisation	4	$[R_{serum}] \to \varnothing$	$k_{R_{int}} \cdot [R_{serum}]$
gp130 synthesis	5	$\varnothing \to [gp130_{serum}]$	$k_{gp130_{synth}}$
gp130 internalisation	6	$[gp130_{serum}] \to \varnothing$	$k_{gp130_{int}} \cdot [gp130_{serum}]$
IL-6 cis association	7	$[IL6_{serum}] + [R_{serum}] \to [R\_IL6_{serum}]$	$k_{IL6_{on}} \cdot [IL6_{serum}] \cdot [R_{serum}]$
IL-6 cis dissociation	8	$[R\_IL6_{serum}] \to [IL6_{serum}] + [R_{serum}]$	$k_{IL6_{off}} \cdot [R\_IL6_{serum}]$
IL-6 cis internalisation	9	$[R\_IL6_{serum}] \to \varnothing$	$k_{R\_IL6_{deg}} \cdot [R\_IL6_{serum}]$
gp130 cis association	10	$[R\_IL6_{serum}] + [gp130_{serum}] \to [R\_IL6\_gp130_{serum}]$	$k_{gp130_{on}} \cdot [R\_IL6_{serum}] \cdot [gp130_{serum}]$
gp130 cis dissociation	11	$[R\_IL6\_gp130_{serum}] \to [R\_IL6_{serum}] + [gp130_{serum}]$	$k_{gp130_{off}} \cdot [R\_IL6\_gp130_{serum}]$
gp130 cis internalisation	12	$[R\_IL6\_gp130_{serum}] \to \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [R\_IL6\_gp130_{serum}]$
sIL-6Rα synthesis	13	$\varnothing \to [sR_{serum}]$	$k_{sR_{serum\,synth}}$
sIL-6Rα degradation	14	$[sR_{serum}] \to \varnothing$	$k_{sR_{deg}} \cdot [sR_{serum}]$
IL-6 trans association	15	$[IL6_{serum}] + [sR_{serum}] \to [sR\_IL6_{serum}]$	$k_{IL6_{on}} \cdot [IL6_{serum}] \cdot [sR_{serum}]$
IL-6 trans dissociation	16	$[sR\_IL6_{serum}] \to [IL6_{serum}] + [sR_{serum}]$	$k_{IL6_{off}} \cdot [sR\_IL6_{serum}]$
IL-6 trans degradation	17	$[sR\_IL6_{serum}] \to \varnothing$	$k_{IL6_{deg}} \cdot [sR\_IL6_{serum}]$
gp130 trans association	18	$[sR\_IL6_{serum}] + [gp130_{serum}] \to [sR\_IL6\_gp130_{serum}]$	$k_{gp130_{on}} \cdot [sR\_IL6_{serum}] \cdot [gp130_{serum}]$
gp130 trans dissociation	19	$[sR\_IL6\_gp130_{serum}] \to [sR\_IL6_{serum}] + [gp130_{serum}]$	$k_{gp130_{off}} \cdot [sR\_IL6\_gp130_{serum}]$
gp130 trans internalisation	20	$[sR\_IL6\_gp130_{serum}] \to \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [sR\_IL6\_gp130_{serum}]$
sgp130 synthesis	21	$\varnothing \to [sgp130_{serum}]$	$k_{sgp130_{serum\,synth}}$
sgp130 degradation	22	$[sgp130_{serum}] \to \varnothing$	$k_{sgp130_{deg}} \cdot [sgp130_{serum}]$
sgp130 association	23	$[sR\_IL6_{serum}] + [sgp130_{serum}] \to [sR\_IL6\_sgp130_{serum}]$	$k_{gp130_{on}} \cdot [sR\_IL6_{serum}] \cdot [sgp130_{serum}]$
sgp130 dissociation	24	$[sR\_IL6\_sgp130_{serum}] \to [sR\_IL6_{serum}] + [sgp130_{serum}]$	$k_{gp130_{off}} \cdot [sR\_IL6\_sgp130_{serum}]$
IL-6 trans degradation	25	$[sR\_IL6\_sgp130_{serum}] \to \varnothing$	$k_{sR_{deg}} \cdot [sR\_IL6\_sgp130_{serum}]$
IL-6Rα cis activation	26	$[R\_IL6\_gp130_{serum}] \to [mR_{active_{serum}}]$	$k_{R_{activation}} \cdot [R\_IL6\_gp130_{serum}]$
IL-6Rα cis deactivation	27	$[mR_{active_{serum}}] \to [R\_IL6\_gp130_{serum}]$	$\dfrac{V_{M_R dephos} \cdot [mR_{active_{serum}}]}{K_{M_R dephos} + [mR_{active_{serum}}]}$
IL-6Rα trans activation	28	$[sR\_IL6\_gp130_{serum}] \to [sR_{active_{serum}}]$	$k_{R_{activation}} \cdot [sR\_IL6\_gp130_{serum}]$
IL-6Rα trans deactivation	29	$[sR_{active_{serum}}] \to [sR\_IL6\_gp130_{serum}]$	$\dfrac{V_{M_R dephos} \cdot [sR_{active_{serum}}]}{K_{M_R dephos} + [sR_{active_{serum}}]}$
IL-6Rα cis internalisation	30	$[mR_{active_{serum}}] \to \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [mR_{active_{serum}}]$
IL-6Rα trans internalisation	31	$[sR_{active_{serum}}] \to \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [sR_{active_{serum}}]$
CRP Degradation	36	$[CRP_{serum}] \to \varnothing$	$k_{CRP_{deg}} \cdot [CRP_{serum}]$
Receptor Shedding	37	$[CRP_{serum}] \to [CRP_{serum}] + [sR_{serum}]$	$k_{R_{shedding}} \cdot [CRP_{serum}]$

Table 3.11: Serum reactions for IL-6 signalling components.

Description	ID	Reaction	Rate
IL-6 synthesis	1	$\varnothing \rightarrow [IL6_{peripheral}]$	$k_{synth_{IL6}}$
IL-6 degradation	2	$[IL6_{peripheral}] \rightarrow \varnothing$	$k_{IL6_{deg}} \cdot [IL6_{peripheral}]$
mIL-6Rα synthesis	3	$\varnothing \rightarrow [R_{peripheral}]$	$k_{R_{peripheral\ synth}}$
mIL-6Rα internalisation	4	$[R_{peripheral}] \rightarrow \varnothing$	$k_{R_{int}} \cdot [R_{peripheral}]$
gp130 synthesis	5	$\varnothing \rightarrow [gp130_{peripheral}]$	$k_{gp130_{synth}}$
gp130 internalisation	6	$[gp130_{peripheral}] \rightarrow \varnothing$	$k_{gp130_{int}} \cdot [gp130_{peripheral}]$
IL-6 cis association	7	$[IL6_{peripheral}] + [R_{peripheral}] \rightarrow [R\_IL6_{peripheral}]$	$k_{IL6_{on}} \cdot [IL6_{peripheral}] \cdot [R_{peripheral}]$
IL-6 cis dissociation	8	$[R\_IL6_{peripheral}] \rightarrow [IL6_{peripheral}] + [R_{peripheral}]$	$k_{IL6_{off}} \cdot [R\_IL6_{peripheral}]$
IL-6 cis internalisation	9	$[R\_IL6_{peripheral}] \rightarrow \varnothing$	$k_{R\_IL6_{deg}} \cdot [R\_IL6_{peripheral}]$
gp130 cis association	10	$[R\_IL6_{peripheral}] + [gp130_{peripheral}] \rightarrow [R\_IL6\_gp130_{peripheral}]$	$k_{sgp130_{on}} \cdot [R\_IL6_{peripheral}] \cdot [gp130_{peripheral}]$
gp130 cis dissociation	11	$[R\_IL6\_gp130_{peripheral}] \rightarrow [R\_IL6_{peripheral}] + [gp130_{peripheral}]$	$k_{gp130_{off}} \cdot [R\_IL6\_gp130_{peripheral}]$
gp130 cis internalisation	12	$[R\_IL6\_gp130_{peripheral}] \rightarrow \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [R\_IL6\_gp130_{peripheral}]$
sIL-6Rα synthesis	13	$\varnothing \rightarrow [sR_{peripheral}]$	$k_{sR_{peripheral\ synth}}$
sIL-6Rα degradation	14	$[sR_{peripheral}] \rightarrow \varnothing$	$k_{sR_{deg}} \cdot [sR_{peripheral}]$
IL-6 trans association	15	$[IL6_{peripheral}] + [sR_{peripheral}] \rightarrow [sR\_IL6_{peripheral}]$	$k_{IL6_{on}} \cdot [IL6_{peripheral}] \cdot [sR_{peripheral}]$
IL-6 trans dissociation	16	$[sR\_IL6_{peripheral}] \rightarrow [IL6_{peripheral}] + [sR_{peripheral}]$	$k_{IL6_{off}} \cdot [sR\_IL6_{peripheral}]$
IL-6 trans degradation	17	$[sR\_IL6_{peripheral}] \rightarrow \varnothing$	$k_{IL6_{deg}} \cdot [sR\_IL6_{peripheral}]$
gp130 trans association	18	$[sR\_IL6_{peripheral}] + [gp130_{peripheral}] \rightarrow [sR\_IL6\_gp130_{peripheral}]$	$k_{gp130_{on}} \cdot [sR\_IL6_{peripheral}] \cdot [gp130_{peripheral}]$
gp130 trans dissociation	19	$[sR\_IL6\_gp130_{peripheral}] \rightarrow [sR\_IL6_{peripheral}] + [gp130_{peripheral}]$	$k_{gp130_{off}} \cdot [sR\_IL6\_gp130_{peripheral}]$
gp130 trans internalisation	20	$[sR\_IL6\_gp130_{peripheral}] \rightarrow \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [sR\_IL6\_gp130_{peripheral}]$
sgp130 synthesis	21	$\varnothing \rightarrow [sgp130_{peripheral}]$	$k_{sgp130_{peripheral\ synth}}$
sgp130 degradation	22	$[sgp130_{peripheral}] \rightarrow \varnothing$	$k_{sgp130_{deg}} \cdot [sgp130_{peripheral}]$
sgp130 association	23	$[sR\_IL6_{peripheral}] + [sgp130_{peripheral}] \rightarrow [sR\_IL6\_sgp130_{peripheral}]$	$k_{gp130_{on}} \cdot [sR\_IL6_{peripheral}] \cdot [sgp130_{peripheral}]$
sgp130 dissociation	24	$[sR\_IL6\_sgp130_{peripheral}] \rightarrow [sR\_IL6_{peripheral}] + [sgp130_{peripheral}]$	$k_{gp130_{off}} \cdot [sR\_IL6\_sgp130_{peripheral}]$
sgp130 trans degradation	25	$[sR\_IL6\_sgp130_{peripheral}] \rightarrow \varnothing$	$k_{sR_{deg}} \cdot [sR\_IL6\_sgp130_{peripheral}]$
IL-6Rα cis activation	26	$[R\_IL6\_gp130_{peripheral}] \rightarrow [mR_{active_{peripheral}}]$	$k_{R_{activation}} \cdot [R\_IL6\_gp130_{peripheral}]$
IL-6Rα cis deactivation	27	$[mR_{active_{peripheral}}] \rightarrow [R\_IL6\_gp130_{peripheral}]$	$\frac{V_M R_{dephos} \cdot [mR_{active_{peripheral}}]}{K_M R_{dephos} + [mR_{active_{peripheral}}]}$
IL-6Rα trans activation	28	$[sR\_IL6\_gp130_{peripheral}] \rightarrow [sR_{active_{peripheral}}]$	$k_{R_{activation}} \cdot [sR\_IL6\_gp130_{peripheral}]$
IL-6Rα trans deactivation	29	$[sR_{active_{peripheral}}] \rightarrow [sR\_IL6\_gp130_{peripheral}]$	$\frac{V_M R_{dephos} \cdot [sR_{active_{peripheral}}]}{K_M R_{dephos} + [sR_{active_{peripheral}}]}$
IL-6Rα cis internalisation	30	$[mR_{active_{peripheral}}] \rightarrow \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [mR_{active_{peripheral}}]$
IL-6Rα trans internalisation	31	$[sR_{active_{peripheral}}] \rightarrow \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [sR_{active_{peripheral}}]$

Table 3.12: Peripheral reactions for IL-6 signalling components.

Description	ID	Reaction	Rate
IL-6 synthesis	1	$\emptyset \rightarrow [IL6_{liver}]$	$k_{synth_{IL6}}$
IL-6 degradation	2	$[IL6_{liver}] \rightarrow \emptyset$	$k_{IL6_{deg}} \cdot [IL6_{liver}]$
mIL-6Rα synthesis	3	$\emptyset \rightarrow [R_{liver}]$	$k_{R_{liver synth}}$
mIL-6Rα internalisation	4	$[R_{liver}] \rightarrow \emptyset$	$k_{R_{int}} \cdot [R_{liver}]$
gp130 synthesis	5	$\emptyset \rightarrow [gp130_{liver}]$	$k_{gp130_{synth}}$
gp130 internalisation	6	$[gp130_{liver}] \rightarrow \emptyset$	$k_{gp130_{int}} \cdot [gp130_{liver}]$
IL-6 cis association	7	$[IL6_{liver}] + [R_{liver}] \rightarrow [R\_IL6_{liver}]$	$k_{IL6_{on}} \cdot [IL6_{liver}] \cdot [R_{liver}]$
IL-6 cis dissociation	8	$[R\_IL6_{liver}] \rightarrow [IL6_{liver}] + [R_{liver}]$	$k_{IL6_{off}} \cdot [R\_IL6_{liver}]$
IL-6 cis internalisation	9	$[R\_IL6_{liver}] \rightarrow \emptyset$	$k_{R\_IL6_{deg}} \cdot [R\_IL6_{liver}]$
gp130 cis association	10	$[R\_IL6_{liver}] + [gp130_{liver}] \rightarrow [R\_IL6\_gp130_{liver}]$	$k_{gp130_{on}} \cdot [R\_IL6_{liver}] \cdot [gp130_{liver}]$
gp130 cis dissociation	11	$[R\_IL6\_gp130_{liver}] \rightarrow [R\_IL6_{liver}] + [gp130_{liver}]$	$k_{gp130_{off}} \cdot [R\_IL6\_gp130_{liver}]$
gp130 cis internalisation	12	$[R\_IL6\_gp130_{liver}] \rightarrow \emptyset$	$k_{R\_IL6\_gp130_{int}} \cdot [R\_IL6\_gp130_{liver}]$
sIL-6Rα synthesis	13	$\emptyset \rightarrow [sR_{liver}]$	$k_{sR_{liver synth}}$
sIL-6Rα degradation	14	$[sR_{liver}] \rightarrow \emptyset$	$k_{sR_{deg}} \cdot [sR_{liver}]$
IL-6 trans association	15	$[IL6_{liver}] + [sR_{liver}] \rightarrow [sR\_IL6_{liver}]$	$k_{IL6_{on}} \cdot [IL6_{liver}] \cdot [sR_{liver}]$
IL-6 trans dissociation	16	$[sR\_IL6_{liver}] \rightarrow [IL6_{liver}] + [sR_{liver}]$	$k_{IL6_{off}} \cdot [sR\_IL6_{liver}]$
IL-6 trans degradation	17	$[sR\_IL6_{liver}] \rightarrow \emptyset$	$k_{sR\_IL6_{deg}} \cdot [sR\_IL6_{liver}]$
gp130 trans association	18	$[sR\_IL6_{liver}] + [gp130_{liver}] \rightarrow [sR\_IL6\_gp130_{liver}]$	$k_{gp130_{on}} \cdot [sR\_IL6_{liver}] \cdot [gp130_{liver}]$
gp130 trans dissociation	19	$[sR\_IL6\_gp130_{liver}] \rightarrow [sR\_IL6_{liver}] + [gp130_{liver}]$	$k_{gp130_{off}} \cdot [sR\_IL6\_gp130_{liver}]$
gp130 trans internalisation	20	$[sR\_IL6\_gp130_{liver}] \rightarrow \emptyset$	$k_{R\_IL6\_gp130_{int}} \cdot [sR\_IL6\_gp130_{liver}]$
sgp130 synthesis	21	$\emptyset \rightarrow [sgp130_{liver}]$	$k_{sgp130_{liver synth}}$
sgp130 degradation	22	$[sgp130_{liver}] \rightarrow \emptyset$	$k_{sgp130_{deg}} \cdot [sgp130_{liver}]$
sgp130 association	23	$[sR\_IL6_{liver}] + [sgp130_{liver}] \rightarrow [sR\_IL6\_sgp130_{liver}]$	$k_{gp130_{on}} \cdot [sR\_IL6_{liver}] \cdot [sgp130_{liver}]$
sgp130 dissociation	24	$[sR\_IL6\_sgp130_{liver}] \rightarrow [sR\_IL6_{liver}] + [sgp130_{liver}]$	$k_{gp130_{off}} \cdot [sR\_IL6\_sgp130_{liver}]$
sgp130 trans degradation	25	$[sR\_IL6\_sgp130_{liver}] \rightarrow \emptyset$	$k_{sR_{deg}} \cdot [sR\_IL6\_sgp130_{liver}]$
IL-6Rα cis activation	26	$[R\_IL6\_gp130_{liver}] \rightarrow [mR_{active_{liver}}]$	$k_{Ractivation} \cdot \dfrac{[R\_IL6\_gp130_{liver}]}{K_M Ractivation + [R\_IL6\_gp130_{liver}]}$
IL-6Rα cis deactivation	27	$[mR_{active_{liver}}] \rightarrow [R\_IL6\_gp130_{liver}]$	$V_M Rdephos \cdot \dfrac{[mR_{active_{liver}}]}{K_M Rdephos + [mR_{active_{liver}}]}$
IL-6Rα trans activation	28	$[sR\_IL6\_gp130_{liver}] \rightarrow [sR_{active_{liver}}]$	$k_{Ractivation} \cdot \dfrac{[sR\_IL6\_gp130_{liver}]}{K_M Ractivation + [sR\_IL6\_gp130_{liver}]}$
IL-6Rα trans deactivation	29	$[sR_{active_{liver}}] \rightarrow [sR\_IL6\_gp130_{liver}]$	$V_M Rdephos \cdot \dfrac{[sR_{active_{liver}}]}{K_M Rdephos + [sR_{active_{liver}}]}$
IL-6Rα cis internalisation	30	$[mR_{active_{liver}}] \rightarrow \emptyset$	$k_{R\_IL6\_gp130_{int}} \cdot [mR_{active_{liver}}]$
IL-6Rα trans internalisation	31	$[sR_{active_{liver}}] \rightarrow \emptyset$	$k_{R\_IL6\_gp130_{int}} \cdot [sR_{active_{liver}}]$
STAT3 phosphorylation	32	$[R_{active_{liver}}] + [STAT3_{liver}] \rightarrow [R_{active_{liver}}] + [pSTAT3_{liver}]$	$\dfrac{V_M STAT_{phos} \cdot [STAT3_{liver}]}{K_M STAT_{phos} + [STAT3_{liver}]} \cdot [R_{active_{liver}}]$
pSTAT3 dephosphorylation	33	$[pSTAT3_{liver}] \rightarrow [STAT3_{liver}]$	$\dfrac{V_M STAT_{dephos} \cdot [pSTAT3_{liver}]}{K_M STAT_{dephos} + [pSTAT3_{liver}]}$
CRP Synthesis	34	$[pSTAT3_{liver}] \rightarrow [CRP_{intracellular}] + [pSTAT3_{liver}]$	$\dfrac{V_M Prot_{synth} \cdot [pSTAT3_{liver}]}{K_M Prot_{synth} + [pSTAT3_{liver}]}$
CRP Secretion	35	$[CRP_{intracellular}] \rightarrow [CRP_{liver}]$	$k_{CRP_{sec}} \cdot [CRP_{liver}]$
CRP Degradation	36	$[CRP_{liver}] \rightarrow \emptyset$	$k_{CRP_{deg}} \cdot [CRP_{liver}]$

Table 3.13: Liver reactions for IL-6 signalling components.

Description	ID	Reaction	Rate
IL-6 synthesis	1	$\varnothing \rightarrow [IL6_{synovium}]$	$k_{synth_{IL6}}$
IL-6 degradation	2	$[IL6_{synovium}] \rightarrow \varnothing$	$k_{IL6_{deg}} \cdot [IL6_{synovium}]$
mIL-6Rα synthesis	3	$\varnothing \rightarrow [R_{synovium}]$	$k_{R_{synovium_{synth}}}$
mIL-6Rα internalisation	4	$[R_{synovium}] \rightarrow \varnothing$	$k_{R_{int}} \cdot [R_{synovium}]$
gp130 synthesis	5	$\varnothing \rightarrow [gp130_{synovium}]$	$k_{gp130_{synth}}$
gp130 internalisation	6	$[gp130_{synovium}] \rightarrow \varnothing$	$k_{gp130_{int}} \cdot [gp130_{synovium}]$
IL-6 cis association	7	$[IL6_{synovium}] + [R_{synovium}] \rightarrow [R\_IL6_{synovium}]$	$k_{IL6_{on}} \cdot [IL6_{synovium}] \cdot [R_{synovium}]$
IL-6 cis dissociation	8	$[R\_IL6_{synovium}] \rightarrow [IL6_{synovium}] + [R_{synovium}]$	$k_{IL6_{off}} \cdot [R\_IL6_{synovium}]$
IL-6 cis internalisation	9	$[R\_IL6_{synovium}] \rightarrow \varnothing$	$k_{R\_IL6_{deg}} \cdot [R\_IL6_{synovium}]$
gp130 cis association	10	$[R\_IL6_{synovium}] + [gp130_{synovium}] \rightarrow [R\_IL6\_gp130_{synovium}]$	$k_{gp130_{on}} \cdot [R\_IL6_{synovium}] \cdot [gp130_{synovium}]$
gp130 cis dissociation	11	$[R\_IL6\_gp130_{synovium}] \rightarrow [R\_IL6_{synovium}] + [gp130_{synovium}]$	$k_{gp130_{off}} \cdot [R\_IL6\_gp130_{synovium}]$
gp130 cis internalisation	12	$[R\_IL6\_gp130_{synovium}] \rightarrow \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [R\_IL6\_gp130_{synovium}]$
sIL-6Rα synthesis	13	$\varnothing \rightarrow [sR_{synovium}]$	$k_{sR_{synovium_{synth}}}$
sIL-6Rα degradation	14	$[sR_{synovium}] \rightarrow \varnothing$	$k_{sR_{deg}} \cdot [sR_{synovium}]$
IL-6 trans association	15	$[IL6_{synovium}] + [sR_{synovium}] \rightarrow [sR\_IL6_{synovium}]$	$k_{IL6_{on}} \cdot [IL6_{synovium}] \cdot [sR_{synovium}]$
IL-6 trans dissociation	16	$[sR\_IL6_{synovium}] \rightarrow [IL6_{synovium}] + [sR_{synovium}]$	$k_{IL6_{off}} \cdot [sR\_IL6_{synovium}]$
IL-6 trans degradation	17	$[sR\_IL6_{synovium}] \rightarrow \varnothing$	$k_{IL6_{deg}} \cdot [sR\_IL6_{synovium}]$
gp130 trans association	18	$[sR\_IL6_{synovium}] + [gp130_{synovium}] \rightarrow [sR\_IL6\_gp130_{synovium}]$	$k_{gp130_{on}} \cdot [sR\_IL6_{synovium}] \cdot [gp130_{synovium}]$
gp130 trans dissociation	19	$[sR\_IL6\_gp130_{synovium}] \rightarrow [sR\_IL6_{synovium}] + [gp130_{synovium}]$	$k_{gp130_{off}} \cdot [sR\_IL6\_gp130_{synovium}]$
gp130 trans internalisation	20	$[sR\_IL6\_gp130_{synovium}] \rightarrow \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [sR\_IL6\_gp130_{synovium}]$
sgp130 synthesis	21	$\varnothing \rightarrow [sgp130_{synovium}]$	$k_{sgp130_{synovium_{synth}}}$
sgp130 degradation	22	$[sgp130_{synovium}] \rightarrow \varnothing$	$k_{sgp130_{deg}} \cdot [sgp130_{synovium}]$
sgp130 association	23	$[sR\_IL6_{synovium}] + [sgp130_{synovium}] \rightarrow [sR\_IL6\_sgp130_{synovium}]$	$k_{gp130_{on}} \cdot [sR\_IL6_{synovium}] \cdot [sgp130_{synovium}]$
sgp130 dissociation	24	$[sR\_IL6\_sgp130_{synovium}] \rightarrow [sR\_IL6_{synovium}] + [sgp130_{synovium}]$	$k_{gp130_{off}} \cdot [sR\_IL6\_sgp130_{synovium}]$
sgp130 trans degradation	25	$[sR\_IL6\_sgp130_{synovium}] \rightarrow \varnothing$	$k_{sR_{deg}} \cdot [sR\_IL6\_sgp130_{synovium}]$
IL-6Rα cis activation	26	$[R\_IL6\_gp130_{synovium}] \rightarrow [mR_{active_{synovium}}]$	$k_{R_{activation}} \cdot [R\_IL6\_gp130_{synovium}]$
IL-6Rα cis deactivation	27	$[mR_{active_{synovium}}] \rightarrow [R\_IL6\_gp130_{synovium}]$	$\dfrac{V_{MR_{dephos}} \cdot [mR_{active_{synovium}}]}{K_{MR_{dephos}} + [mR_{active_{synovium}}]}$
IL-6Rα trans activation	28	$[sR\_IL6\_gp130_{synovium}] \rightarrow [sR_{active_{synovium}}]$	$k_{R_{activation}} \cdot [sR\_IL6\_gp130_{synovium}]$
IL-6Rα trans deactivation	29	$[sR_{active_{synovium}}] \rightarrow [sR\_IL6\_gp130_{synovium}]$	$\dfrac{V_{MR_{dephos}} \cdot [sR_{active_{synovium}}]}{K_{MR_{dephos}} + [sR_{active_{synovium}}]}$
IL-6Rα cis internalisation	30	$[mR_{active_{synovium}}] \rightarrow \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [mR_{active_{synovium}}]$
IL-6Rα trans internalisation	31	$[sR_{active_{synovium}}] \rightarrow \varnothing$	$k_{R\_IL6\_gp130_{int}} \cdot [sR_{active_{synovium}}]$
STAT3 phosphorylation	32	$[R_{active_{synovium}}] + [STAT3_{synovium}] \rightarrow [R_{active_{synovium}}] + [pSTAT3_{synovium}]$	$\dfrac{k_{cat}STAT_{phos} \cdot [STAT3_{synovium}] \cdot [R_{active_{synovium}}]}{K_M STAT_{phos} + [STAT3_{synovium}]}$
pSTAT3 dephosphorylation	33	$[pSTAT3_{synovium}] \rightarrow [STAT3_{synovium}]$	$\dfrac{V_M STAT_{dephos} \cdot [pSTAT3_{synovium}]}{K_M STAT_{dephos} + [pSTAT3_{synovium}]}$

Table 3.14: Synovium reactions for IL-6 signalling components.

Description	ID	Reaction	Rate
IL6 in	D2	$[IL6_{serum}] \rightarrow [IL6_{peripheral}]$	$k_{dist_{12}} \cdot [IL6_{serum}]$
IL6 out	D2	$[IL6_{peripheral}] \rightarrow [IL6_{serum}]$	$k_{dist_{21}} \cdot [IL6_{peripheral}]$
sR in	D2	$[sR_{serum}] \rightarrow [sR_{peripheral}]$	$k_{dist_{12}} \cdot [sR_{serum}]$
sR out	D2	$[sR_{peripheral}] \rightarrow [sR_{serum}]$	$k_{dist_{21}} \cdot [sR_{peripheral}]$
sgp130 in	D2	$[sgp130_{serum}] \rightarrow [sgp130_{peripheral}]$	$k_{dist_{12}} \cdot [sgp130_{serum}]$
sgp130 out	D2	$[sgp130_{peripheral}] \rightarrow [sgp130_{serum}]$	$k_{dist_{21}} \cdot [sgp130_{peripheral}]$
sR_IL6 in	D2	$[sR\_IL6_{serum}] \rightarrow [sR\_IL6_{peripheral}]$	$k_{dist_{12}} \cdot [sR\_IL6_{serum}]$
sR_IL6 out	D2	$[sR\_IL6_{peripheral}] \rightarrow [sR\_IL6_{serum}]$	$k_{dist_{21}} \cdot [sR\_IL6_{peripheral}]$
sR_IL6_sgp130 in	D2	$[sR\_IL6\_sgp130_{serum}] \rightarrow [sR\_IL6\_sgp130_{peripheral}]$	$k_{dist_{12}} \cdot [sR\_IL6\_sgp130_{serum}]$
sR_IL6_sgp130 out	D2	$[sR\_IL6\_sgp130_{peripheral}] \rightarrow [sR\_IL6\_sgp130_{serum}]$	$k_{dist_{21}} \cdot [sR\_IL6\_sgp130_{peripheral}]$
IL6 in	D3	$[IL6_{serum}] \rightarrow [IL6_{liver}]$	$k_{dist_{13}} \cdot [IL6_{serum}]$
IL6 out	D3	$[IL6_{liver}] \rightarrow [IL6_{serum}]$	$k_{dist_{31}} \cdot [IL6_{liver}]$
sR in	D3	$[sR_{serum}] \rightarrow [sR_{liver}]$	$k_{dist_{13}} \cdot [sR_{serum}]$
sR out	D3	$[sR_{liver}] \rightarrow [sR_{serum}]$	$k_{dist_{31}} \cdot [sR_{liver}]$
sgp130 in	D3	$[sgp130_{serum}] \rightarrow [sgp130_{liver}]$	$k_{dist_{13}} \cdot [sgp130_{serum}]$
sgp130 out	D3	$[sgp130_{liver}] \rightarrow [sgp130_{serum}]$	$k_{dist_{31}} \cdot [sgp130_{liver}]$
sR_IL6 in	D3	$[sR\_IL6_{serum}] \rightarrow [sR\_IL6_{liver}]$	$k_{dist_{13}} \cdot [sR\_IL6_{serum}]$
sR_IL6 out	D3	$[sR\_IL6_{liver}] \rightarrow [sR\_IL6_{serum}]$	$k_{dist_{31}} \cdot [sR\_IL6_{liver}]$
sR_IL6_sgp130 in	D3	$[sR\_IL6\_sgp130_{serum}] \rightarrow [sR\_IL6\_sgp130_{liver}]$	$k_{dist_{13}} \cdot [sR\_IL6\_sgp130_{serum}]$
sR_IL6_sgp130 out	D3	$[sR\_IL6\_sgp130_{liver}] \rightarrow [sR\_IL6\_sgp130_{serum}]$	$k_{dist_{31}} \cdot [sR\_IL6\_sgp130_{liver}]$
IL6 in	D4	$[IL6_{serum}] \rightarrow [IL6_{synovium}]$	$k_{dist_{14}} \cdot [IL6_{serum}]$
IL6 out	D4	$[IL6_{synovium}] \rightarrow [IL6_{serum}]$	$k_{dist_{41}} \cdot [IL6_{synovium}]$
sR in	D4	$[sR_{serum}] \rightarrow [sR_{synovium}]$	$k_{dist_{14}} \cdot [sR_{serum}]$
sR out	D4	$[sR_{synovium}] \rightarrow [sR_{serum}]$	$k_{dist_{41}} \cdot [sR_{synovium}]$
sgp130 in	D4	$[sgp130_{serum}] \rightarrow [sgp130_{synovium}]$	$k_{dist_{14}} \cdot [sgp130_{serum}]$
sgp130 out	D4	$[sgp130_{synovium}] \rightarrow [sgp130_{serum}]$	$k_{dist_{41}} \cdot [sgp130_{synovium}]$
sR_IL6 in	D4	$[sR\_IL6_{serum}] \rightarrow [sR\_IL6_{synovium}]$	$k_{dist_{14}} \cdot [sR\_IL6_{serum}]$
sR_IL6 out	D4	$[sR\_IL6_{synovium}] \rightarrow [sR\_IL6_{serum}]$	$k_{dist_{41}} \cdot [sR\_IL6_{synovium}]$
sR_IL6_sgp130 in	D4	$[sR\_IL6\_sgp130_{serum}] \rightarrow [sR\_IL6\_sgp130_{synovium}]$	$k_{dist_{14}} \cdot [sR\_IL6\_sgp130_{serum}]$
sR_IL6_sgp130 out	D4	$[sR\_IL6\_sgp130_{synovium}] \rightarrow [sR\_IL6\_sgp130_{serum}]$	$k_{dist_{41}} \cdot [sR\_IL6\_sgp130_{synovium}]$

Table 3.15: Distribution reactions for IL-6 signalling components.

Description	ID	Reaction	Rate
mAb Absorption	A1	$[Ab_{depot}] \to [Ab_{serum}]$	$k_{abs} \cdot [Ab_{depot}]$
serum linear elimination	A2	$[Ab_{serum}] \to \varnothing$	$k_{el} \cdot [Ab_{serum}]$
serum mIL-6Rα association	A3	$[Ab_{serum}] + [R_{serum}] \to [Ab\_R_{serum}]$	$k_{Ab_{on}} \cdot [Ab_{serum}] \cdot [R_{serum}]$
serum mIL-6Rα dissociation	A4	$[Ab\_R_{serum}] \to [Ab_{serum}] + [R_{serum}]$	$k_{Ab_{off}} \cdot [Ab\_R_{serum}]$
serum mAb:mIL-6Rα degradation	A5	$[Ab\_R_{serum}] \to \varnothing$	$k_{el} \cdot [Ab\_R_{serum}]$
serum mAb:mIL-6Rα internalisation	A6	$[Ab\_R_{serum}] \to \varnothing$	$k_{Ab\_R_{int}} \cdot [Ab\_R_{serum}]$
serum sIL-6Rα association	A7A	$[Ab_{serum}] + [sR_{serum}] \to [Ab\_sR_{serum}]$	$k_{Ab_{on}} \cdot [Ab_{serum}] \cdot [sR_{serum}]$
serum sIL-6Rα dissociation	A8A	$[Ab\_sR_{serum}] \to [Ab_{serum}] + [sR_{serum}]$	$k_{Ab_{off}} \cdot [Ab\_sR_{serum}]$
serum mAb:sIL-6Rα degradation	A9A	$[Ab\_sR_{serum}] \to \varnothing$	$k_{Ab\_sR_{deg}} \cdot [Ab\_sR_{serum}]$
serum IL-6 association	A7B	$[Ab_{serum}] + [IL6_{serum}] \to [Ab\_IL6_{serum}]$	$k_{Ab_{on}} \cdot [Ab_{serum}] \cdot [IL6_{serum}]$
serum IL-6 dissociation	A8B	$[Ab\_IL6_{serum}] \to [Ab_{serum}] + [IL6_{serum}]$	$k_{Ab_{off}} \cdot [Ab\_IL6_{serum}]$
serum mAb:IL-6 degradation	A9B	$[Ab\_IL6_{serum}] \to \varnothing$	$k_{Ab\_IL6_{deg}} \cdot [Ab\_IL6_{serum}]$
mAb distribution to peripheral	D2$_{Ab}$	$[Ab_{serum}] \to [Ab_{peripheral}]$	$k_{12} \cdot [Ab_{serum}]$
mAb distribution from peripheral	D2$_{Ab}$	$[Ab_{peripheral}] \to [Ab_{serum}]$	$k_{21} \cdot [Ab_{peripheral}]$
mAb:sIL-6Rα distribution to peripheral	D2$_{Ab}$	$[Ab\_sR_{serum}] \to [Ab\_sR_{peripheral}]$	$k_{12} \cdot [Ab\_sR_{serum}]$
mAb:sIL-6Rα distribution from peripheral	D2$_{Ab}$	$[Ab\_sR_{peripheral}] \to [Ab\_sR_{serum}]$	$k_{21} \cdot [Ab\_sR_{peripheral}]$
mAb:IL-6 distribution to peripheral	D2$_{Ab}$	$[Ab\_IL6_{serum}] \to [Ab\_IL6_{peripheral}]$	$k_{12} \cdot [Ab\_IL6_{serum}]$
mAb:IL-6 distribution from peripheral	D2$_{Ab}$	$[Ab\_IL6_{peripheral}] \to [Ab\_IL6_{serum}]$	$k_{21} \cdot [Ab\_IL6_{peripheral}]$
peripheral linear elimination	A2	$[Ab_{peripheral}] \to \varnothing$	$k_{el} \cdot [Ab_{peripheral}]$
peripheral mIL-6Rα association	A3	$[Ab_{peripheral}] + [R_{peripheral}] \to [Ab\_R_{peripheral}]$	$k_{Ab_{on}} \cdot [Ab_{peripheral}] \cdot [R_{peripheral}]$
peripheral mIL-6Rα dissociation	A4	$[Ab\_R_{peripheral}] \to [Ab_{peripheral}] + [R_{peripheral}]$	$k_{Ab_{off}} \cdot [Ab\_R_{peripheral}]$
peripheral mAb:mIL-6Rα degradation	A5	$[Ab\_R_{peripheral}] \to \varnothing$	$k_{el} \cdot [Ab\_R_{peripheral}]$
peripheral mAb:mIL-6Rα internalisation	A6	$[Ab\_R_{peripheral}] \to \varnothing$	$k_{Ab\_R_{int}} \cdot [Ab\_R_{peripheral}]$
peripheral sIL-6Rα association	A7A	$[Ab_{peripheral}] + [sR_{peripheral}] \to [Ab\_sR_{peripheral}]$	$k_{Ab_{on}} \cdot [Ab_{peripheral}] \cdot [sR_{peripheral}]$
peripheral sIL-6Rα dissociation	A8A	$[Ab\_sR_{peripheral}] \to [Ab_{peripheral}] + [sR_{peripheral}]$	$k_{Ab_{off}} \cdot [Ab\_sR_{peripheral}]$
peripheral mAb:sIL-6Rα degradation	A9A	$[Ab\_sR_{peripheral}] \to \varnothing$	$k_{Ab\_sR_{deg}} \cdot [Ab\_sR_{peripheral}]$
peripheral IL-6 association	A7B	$[Ab_{peripheral}] + [IL6_{peripheral}] \to [Ab\_IL6_{peripheral}]$	$k_{Ab_{on}} \cdot [Ab_{peripheral}] \cdot [IL6_{peripheral}]$
peripheral IL-6 dissociation	A8B	$[Ab\_IL6_{peripheral}] \to [Ab_{peripheral}] + [IL6_{peripheral}]$	$k_{Ab_{off}} \cdot [Ab\_IL6_{peripheral}]$
peripheral mAb:IL-6 degradation	A9B	$[Ab\_IL6_{peripheral}] \to \varnothing$	$k_{Ab\_IL6_{deg}} \cdot [Ab\_IL6_{peripheral}]$

Table 3.16: Serum and peripheral compartment antibody reactions.

1, serum; 2, peripheral; 3, liver; 4, synovium

Description	ID	Reaction	Rate
mAb distribution to liver	$D2_{Ab}$	$[Ab_{serum}] \rightarrow [Ab_{liver}]$	$k_{13} \cdot [Ab_{serum}]$
mAb distribution from liver	$D2_{Ab}$	$[Ab_{liver}] \rightarrow [Ab_{serum}]$	$k_{31} \cdot [Ab_{liver}]$
mAb:sIL-6Rα distribution to liver	$D2_{Ab}$	$[Ab\_sR_{serum}] \rightarrow [Ab\_sR_{liver}]$	$k_{13} \cdot [Ab\_sR_{serum}]$
mAb:sIL-6Rα distribution from liver	$D2_{Ab}$	$[Ab\_sR_{serum}] \rightarrow [Ab\_sR_{liver}]$	$k_{31} \cdot [Ab\_sR_{liver}]$
mAb:IL-6 distribution to liver	$D2_{Ab}$	$[Ab\_IL6_{serum}] \rightarrow [Ab\_IL6_{liver}]$	$k_{13} \cdot [Ab\_IL6_{serum}]$
mAb:IL-6 distribution from liver	$D2_{Ab}$	$[Ab\_IL6_{liver}] \rightarrow [Ab\_IL6_{serum}]$	$k_{31} \cdot [Ab\_IL6_{liver}]$
liver linear elimination	$A2$	$[Ab_{liver}] \rightarrow \emptyset$	$k_{el} \cdot [Ab_{liver}]$
liver mIL-6Rα association	$A3$	$[Ab_{liver}] + [R_{liver}] \rightarrow [Ab\_R_{liver}]$	$k_{Ab_{on}} \cdot [Ab_{liver}] \cdot [R_{liver}]$
liver mIL-6Rα dissociation	$A4$	$[Ab\_R_{liver}] \rightarrow [Ab_{liver}] + [R_{liver}]$	$k_{Ab_{off}} \cdot [R_{liver}]$
liver mAb:mIL-6Rα degradation	$A5$	$[Ab\_R_{liver}] \rightarrow \emptyset$	$k_{el} \cdot [Ab\_R_{liver}]$
liver mAb:mIL-6Rα internalisation	$A6$	$[Ab\_R_{liver}] \rightarrow \emptyset$	$k_{Ab\_R_{int}} \cdot [Ab\_R_{liver}]$
liver sIL-6Rα association	$A7A$	$[Ab_{liver}] + [sR_{liver}] \rightarrow [Ab\_sR_{liver}]$	$k_{Ab_{on}} \cdot [Ab_{liver}] \cdot [sR_{liver}]$
liver sIL-6Rα dissociation	$A8A$	$[Ab\_sR_{liver}] \rightarrow [Ab_{liver}] + [sR_{liver}]$	$k_{Ab_{off}} \cdot [Ab\_sR_{liver}]$
liver mAb:sIL-6Rα degradation	$A9A$	$[Ab\_sR_{liver}] \rightarrow \emptyset$	$k_{Ab\_sR_{deg}} \cdot [Ab\_sR_{liver}]$
liver IL-6 association	$A7B$	$[Ab_{liver}] + [IL6_{liver}] \rightarrow [Ab\_IL6_{liver}]$	$k_{Ab_{on}} \cdot [Ab_{liver}] \cdot [IL6_{liver}]$
liver IL-6 dissociation	$A8B$	$[Ab\_IL6_{liver}] \rightarrow [Ab_{liver}] + [IL6_{liver}]$	$k_{Ab_{off}} \cdot [Ab\_IL6_{liver}]$
liver mAb:IL-6 degradation	$A9B$	$[Ab\_IL6_{liver}] \rightarrow \emptyset$	$k_{Ab\_IL6_{deg}} \cdot [Ab\_IL6_{liver}]$
mAb distribution to synovium	$D2_{Ab}$	$[Ab_{serum}] \rightarrow [Ab_{synovium}]$	$k_{14} \cdot [Ab_{serum}]$
mAb distribution from synovium	$D2_{Ab}$	$[Ab_{synovium}] \rightarrow [Ab_{serum}]$	$k_{41} \cdot [Ab_{synovium}]$
mAb:sIL-6Rα distribution to synovium	$D2_{Ab}$	$[Ab\_sR_{serum}] \rightarrow [Ab\_sR_{synovium}]$	$k_{14} \cdot [Ab\_sR_{serum}]$
mAb:sIL-6Rα distribution from synovium	$D2_{Ab}$	$[Ab\_sR_{synovium}] \rightarrow [Ab\_sR_{serum}]$	$k_{41} \cdot [Ab\_sR_{synovium}]$
mAb:IL-6 distribution to synovium	$D2_{Ab}$	$[Ab\_IL6_{serum}] \rightarrow [Ab\_IL6_{synovium}]$	$k_{14} \cdot [Ab\_IL6_{serum}]$
mAb:IL-6 distribution from synovium	$D2_{Ab}$	$[Ab\_IL6_{synovium}] \rightarrow [Ab\_IL6_{serum}]$	$k_{41} \cdot [Ab\_IL6_{synovium}]$
synovium linear elimination	$A2$	$[Ab_{synovium}] \rightarrow \emptyset$	$k_{el} \cdot [Ab_{synovium}]$
synovium mIL-6Rα association	$A3$	$[Ab_{synovium}] + [R_{synovium}] \rightarrow [Ab\_R_{synovium}]$	$k_{Ab_{on}} \cdot [Ab_{synovium}] \cdot [R_{synovium}]$
synovium mIL-6Rα dissociation	$A4$	$[Ab\_R_{synovium}] \rightarrow [Ab_{synovium}] + [R_{synovium}]$	$k_{Ab_{off}} \cdot [Ab_{R_{synovium}}]$
synovium mAb:mIL-6Rα degradation	$A5$	$[Ab\_R_{synovium}] \rightarrow \emptyset$	$k_{el} \cdot [Ab\_R_{synovium}]$
synovium mAb:mIL-6Rα internalisation	$A6$	$[Ab\_R_{synovium}] \rightarrow \emptyset$	$k_{Ab\_R_{int}} \cdot [Ab\_R_{synovium}]$
synovium sIL-6Rα association	$A7A$	$[Ab_{synovium}] + [sR_{synovium}] \rightarrow [Ab\_sR_{synovium}]$	$k_{Ab_{on}} \cdot [Ab_{synovium}] \cdot [sR_{synovium}]$
synovium sIL-6Rα dissociation	$A8A$	$[Ab\_sR_{synovium}] \rightarrow [Ab_{synovium}] + [sR_{synovium}]$	$k_{Ab_{off}} \cdot [Ab\_sR_{synovium}]$
synovium mAb:sIL-6Rα degradation	$A9A$	$[Ab\_sR_{synovium}] \rightarrow \emptyset$	$k_{Ab\_sR_{deg}} \cdot [Ab\_sR_{synovium}]$
synovium IL-6 association	$A7B$	$[Ab_{synovium}] + [IL6_{synovium}] \rightarrow [Ab\_IL6_{synovium}]$	$k_{Ab_{on}} \cdot [Ab_{synovium}] \cdot [IL6_{synovium}]$
synovium IL-6 dissociation	$A8B$	$[Ab\_IL6_{synovium}] \rightarrow [Ab_{synovium}] + [IL6_{synovium}]$	$k_{Ab_{off}} \cdot [Ab\_IL6_{synovium}]$
synovium mAb:IL-6 degradation	$A9B$	$[Ab\_IL6_{synovium}] \rightarrow \emptyset$	$k_{Ab\_IL6_{deg}} \cdot [Ab\_IL6_{synovium}]$

Table 3.17: Liver and synovium compartment antibody reactions.

Description	Name	Value	Units	Source
Tissue IL-6 synthesis	$k_{IL6_{synth}}$	0.0165	nM h^{-1}	†
IL-6 degradation	$k_{IL6_{deg}}$	0.164	h^{-1}	A
sIL-6Rα degradation	$k_{sR_{deg}}$	0.11	h^{-1}	B
sgp130 degradation	$k_{sgp130_{deg}}$	0.11	h^{-1}	B
CRP degradation	$k_{CRP_{deg}}$	0.0365	h^{-1}	C
gp130 association	$k_{gp130_{on}}$	1.0	$\text{nM}^{-1}\,\text{h}^{-1}$	D, E, F
gp130 dissociation	$k_{gp130_{off}}$	0.05	h^{-1}	D, E, F
IL-6 association	$k_{IL6_{on}}$	1.0	$\text{nM}^{-1}\,\text{h}^{-1}$	G, H, I, J
IL-6 dissociation	$k_{IL6_{off}}$	0.5	h^{-1}	G, H, I, J
gp130 synthesis	$k_{gp130_{synth}}$	2.77 (0.29)	nM h^{-1}	†
basal gp130 internalisation	$k_{gp130_{int}}$	0.35	h^{-1}	K, L
basal mIL-6Rα internalisation	$k_{R_{int}}$	0.25	h^{-1}	M
IL-6 driven mIL-6Rα internalisation	$k_{R\text{-}IL6_{int}}$	0.35	h^{-1}	M
IL-6 driven gp130 internalisation	$k_{R\text{-}IL6\text{-}gp130_{int}}$	2.77	h^{-1}	N, O, P
Receptor activation	$k_{R_{activation}}$	155	h^{-1}	Q
Receptor dephosphorylation	$V_m R_{dephos}$	0.525	h^{-1}	Q
MM constant receptor dephosphorylation	$K_M R_{dephos}$	155.3	nM	Q
IL-6Rα shedding	$k_{R_{shedding}}$	1.02×10^{-6}	h^{-1}	†
V_M STAT3 phosphorylation	$k_{cat} STAT_{phos}$	5280	h^{-1}	†
K_M STAT3 phosphorylation	$K_M STAT_{phos}$	0.0164	nM	†
V_M pSTAT3 dephosphorylation	$V_M STAT_{dephos}$	144000	nM h^{-1}	†
K_M pSTAT3 dephosphorylation	$K_M STAT_{dephos}$	31.4	nM	†
Serum T cells	$Tcells_{serum}$	7.83×10^8	cells l^{-1}	*R, S, T
Serum sIL-6Rα synthesis	$k_{sR_{serum_{synth}}}$	0.00327	nM h^{-1}	*
Peripheral T cells	$Tcells_{peripheral}$	8.65×10^{10}	cells l^{-1}	*U
Peripheral mIL-6Rα synthesis	$k_{R_{peripheral_{synth}}}$	8.675×10^{-3}	nM h^{-1}	*V
Peripheral sIL-6Rα synthesis	$k_{sR_{peripheral_{synth}}}$	0.361	nM h^{-1}	†
Serum mIL-6Rα synthesis	$k_{R_{serum_{synth}}}$	7.85×10^{-5}	nM h^{-1}	*
Serum sgp130 synthesis	$k_{sgp130_{serum_{synth}}}$	0.0986	nM h^{-1}	†
Peripheral sgp130 synthesis	$k_{sgp130_{peripheral_{synth}}}$	1	nM h^{-1}	†
Liver mIL-6Rα synthesis	$k_{R_{liver_{synth}}}$	0.254	nM h^{-1}	*N
Liver sIL-6Rα synthesis	$k_{sR_{liver_{synth}}}$	0.365	nM h^{-1}	†
Liver sgp130 synthesis	$k_{sgp130_{liver_{synth}}}$	1.23	nM h^{-1}	†
V_M Protein synthesis	$V_m prot_{synth}$	1.36×10^8	nM h^{-1}	†
K_M Protein MM constant	$K_M prot_{synth}$	15100	nM	†
Synovial T cells	$Tcells_{synovium}$	3.92×10^8	cells l^{-1}	*T
Synovial IL-6 synthesis	$k_{IL6_{synovium_{synth}}}$	0.226	nM h^{-1}	†
Synovial sIL-6Rα synthesis	$k_{sR_{synovium_{synth}}}$	0.291	nM h^{-1}	†
Synovial mIL-6Rα synthesis	$k_{R_{synovium_{synth}}}$	3.93×10^{-5}	nM h^{-1}	*
Synovial sgp130 synthesis	$k_{sgp130_{synovium_{synth}}}$	2.44×10^{-6}	nM h^{-1}	†
PD flow parameter	Q_{PD}	0.255	l h^{-1}	†
Peripheral PD distribution rate	$k_{dist_{12}}$	0.146	h^{-1}	†
Liver PD distribution rate	$k_{dist_{13}}$	0.146	h^{-1}	†
Synovium PD distribution rate	$k_{dist_{14}}$	0.00988	h^{-1}	†
Peripheral PD distribution rate	$k_{dist_{21}}$	0.0227	h^{-1}	†
Liver PD distribution rate	$k_{dist_{31}}$	0.142	h^{-1}	†
Synovium PD distribution rate	$k_{dist_{41}}$	0.115	h^{-1}	†
CRP secretion rate	$k_{CRP_{secretion}}$	0.116	h^{-1}	W

Sources: *A*, (Weber et al., 1993); *B*, (Jacobs et al., 1993); *C*, (Pepys and Hirschfield, 2003); *D*, (Richards et al., 2006); *E*, (Jostock et al., 2001); *F*, (Schroers, 2005); *G*, (Heinrich et al., 1998); *H*, (Gearing et al., 1992); *I*, (Weiergraber et al., 1995); *J*, (Dittrich et al., 1994); *K*, (Thiel et al., 1998); *L*, (Blanchard et al., 2001); *M*, (Fujimoto et al., 2015); *N*, (Zohlnhöfer et al., 1992); *O*, (Nesbitt and Fuller, 1992); *P*, (Dittrich et al., 1996); *Q*, (Dwivedi et al., 2014); *R*, (Horneff et al., 1991); *S*, (Mélet et al., 2013); *T*, (Moradi et al., 2014); *U*, (Mascio et al., 2009); *V*, (Bongioanni et al., 2000); *W*, (Macintyre et al., 1985). *, Calculated *a priori* using evidence from literature. †, Estimated parameters using data from literature.

Table 3.18: IL-6 signalling parameter values and sources.

Description	Name	Value	Units	Source
Serum volume	$serum$	3.5	l	A
Peripheral volume	$peripheral$	11.25	l	*
Liver volume	$liver$	1.8	l	B, C, D
Synovium volume	$synovium$	0.15	l	E, F, G
mAb mass	Ab_kDa	145	kDa	H
mAb bioavailability	$bioavailability$	0.81 (SC)	N/A	I
mAb absorption	k_{abs}	0.0091 (SC)	h^{-1}	A
mAb elimination rate	k_{el}	0.0012257	h^{-1}	J
mAb association rate	$k_{Ab_{on}}$	1	$nM^{-1}\,h^{-1}$	K L
mAb dissociation rate	$k_{Ab_{off}}$	0.001 - 1	$nM^{-1}\,h^{-1}$	K, L
mAb:sIL-6Rα elimination rate	$k_{Ab\_sR_{deg}}$	0.00976	h^{-1}	†
mAb:mIL-6Rα internalisation rate	$k_{Ab\_R_{deg}}$	0.35	h^{-1}	M
mAb:IL-6 elimination rate	$k_{Ab\_IL6_{deg}}$	0.0012257	h^{-1}	J
PK flow parameter	Q	0.0354	$l\,h^{-1}$	†
Peripheral PK distribution rate	k_{12}	0.0101	h^{-1}	†
Peripheral PK distribution rate	k_{21}	0.00314	h^{-1}	†
Liver PK distribution rate	k_{13}	0.0101	h^{-1}	†
Liver PK distribution rate	k_{31}	0.0197	h^{-1}	†
Synovium PK distribution rate	k_{14}	6.86×10^{-4}	h^{-1}	†
Synovium PK distribution rate	k_{41}	0.0320	h^{-1}	†
Synovium perfusion fraction	V_{4p}	0.0678	N/A	†

Sources: *A*, (*RoActemra: EPAR - Product Information*); *B*, (Kwon et al., 2001); *C*, (Heinemann et al., 1999); *D*, (Kan and Hopkins, 1979); *E*, (R.W et al., 2007); *F*, (Buckwalter, 2007); *G*, (Simkin et al., 1995); *H*, (Mihara et al., 2011); *I*, (*CHMP Assessment report RoActemra. EMA/CHMP/606295/2013, 20 February 2014.*); *J*, (Paul, 2008); *K*, (Mihara et al., 2005); *L*, (*U.S Food and Drug Administration, BLA: 125276*); *M*, (Fujimoto et al., 2015); *, Calculated *a priori* using evidence from literature. †, Estimated parameters using data from literature.

Table 3.19: Antibody parameter values and sources.

3.3.2 *Validation*

To validate the model, simulations were compared with independent datasets not previously included in estimation. These cover the PK and PD of both TCZ and SRK. The PK data to both mAbs were relatively well predicted, and single and multiple-dose regimens for TCZ were readily simulated using the final parameter set (Figure 3.11). This provides evidence that the model suitably predicts TCZ kinetics. This was reasonably expected as TCZ PK data were were used to parametrise the model. However, the PK of an anti–IL-6 mAb such as SRK emerged from the model suggesting that the model can be applied to similar mAbs and molecules (Figure 3.12).

On the other hand, the PD data were less well-simulated. The extent of CRP suppression was evidently under-predicted in TCZ simulations and the elevation of sIL-6Rα in response to TCZ binding was over-estimated (Figures 3.13A and B). Furthermore, while the initial dose of TCZ dissociated IL-6 from IL-6Rα in good agreement with experimental data, the final peak was over-predicted by approximately 3-fold. This was observed despite the success of simulations in adhering to experimental trough measurements (the concentrations of IL-6 prior to administering a new dose) (Figure 3.13C). Additionally, the simulated suppression of CRP following SRK administration shows that although the model captures the depth of CRP suppression in RA
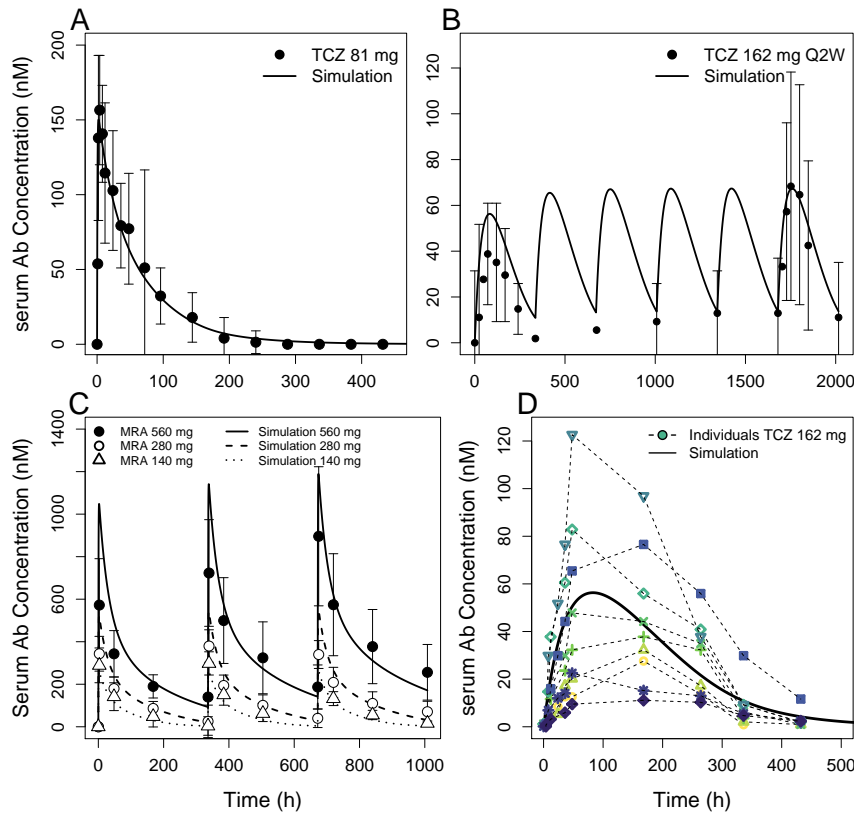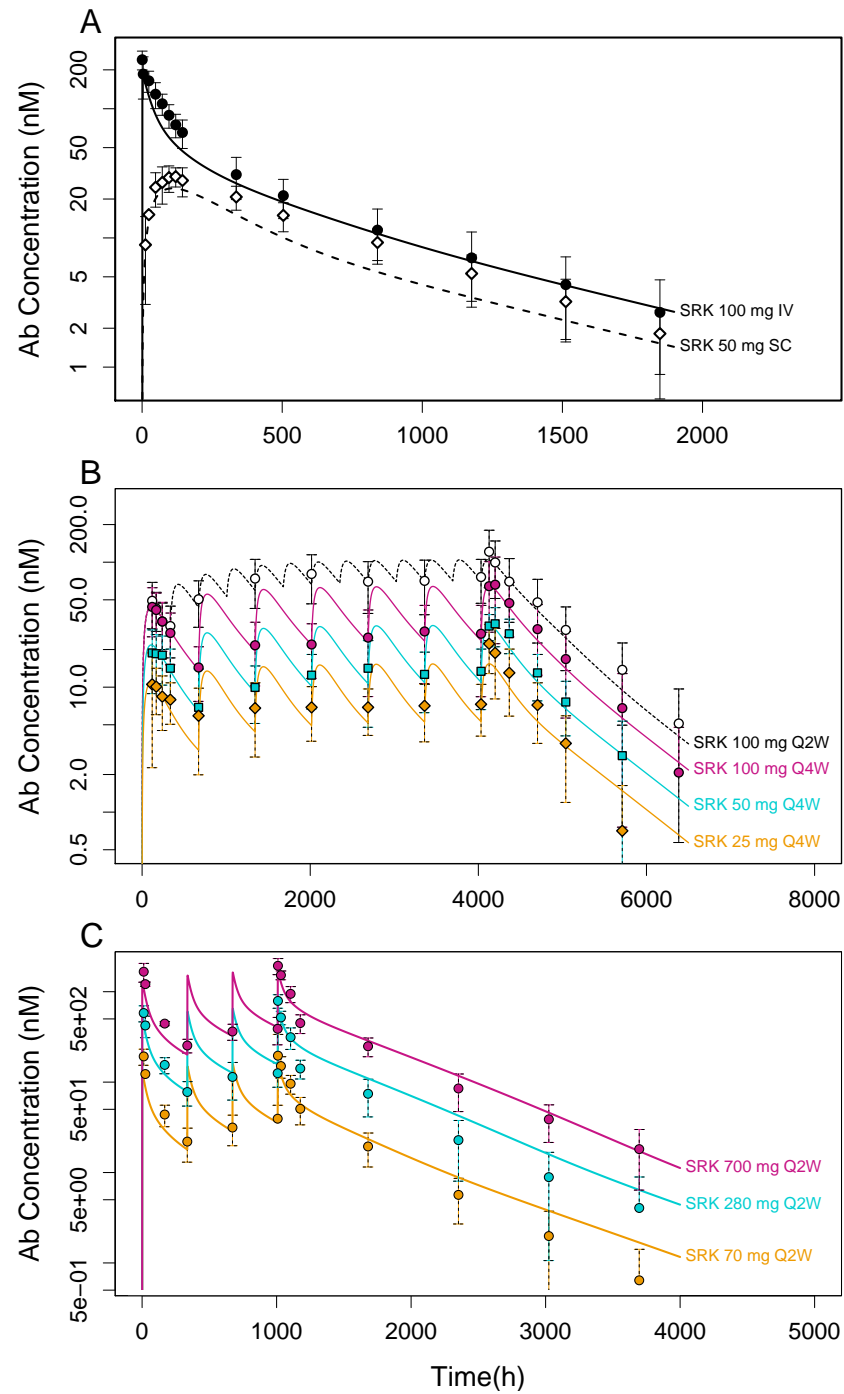
Figure 3.11: TCZ PK Validations. Simulated TCZ administration matches the serum PK profiles of several dose regimens. **A**, single-dose IV (Zhang et al., 2013a); **B**, multiple-dose SC (Zhang et al., 2013b); **C**, multiple-dose IV (Nishimoto et al., 2003); and **D**, single-dose SC regimens (Ohta et al., 2013) were captured by the whole model.

patients, the simulation fails to capture the slow speed of CRP recovery seen in clinical data (Figure 3.14).

Shown in these simulations, the model captures and predicts the PK of mAbs reasonable well. However, there are quantitative differences between experimental data and that of the model in the case of PD dynamics. This error is largely within the experimental error with mean differences of approximately 522, 52 and 46 per cent from the mean experimental data (CRP, sIL-6Rα and IL-6 respectively). These differences between simulation and clinical data may be explained by multiple factors associated with the structure and parametrisation of the model.

Figure 3.12: SRK PK Validations. Simulated SRK administration matches the serum PK profiles of several dose regimens in RA patients. **A**, single-dose administration (Zhuang et al., 2016); **B**, multiple-dose SC administration (Smolen et al., 2014); **C**, multiple-dose IV administration in patients with lupus (Szepietowski et al., 2013).
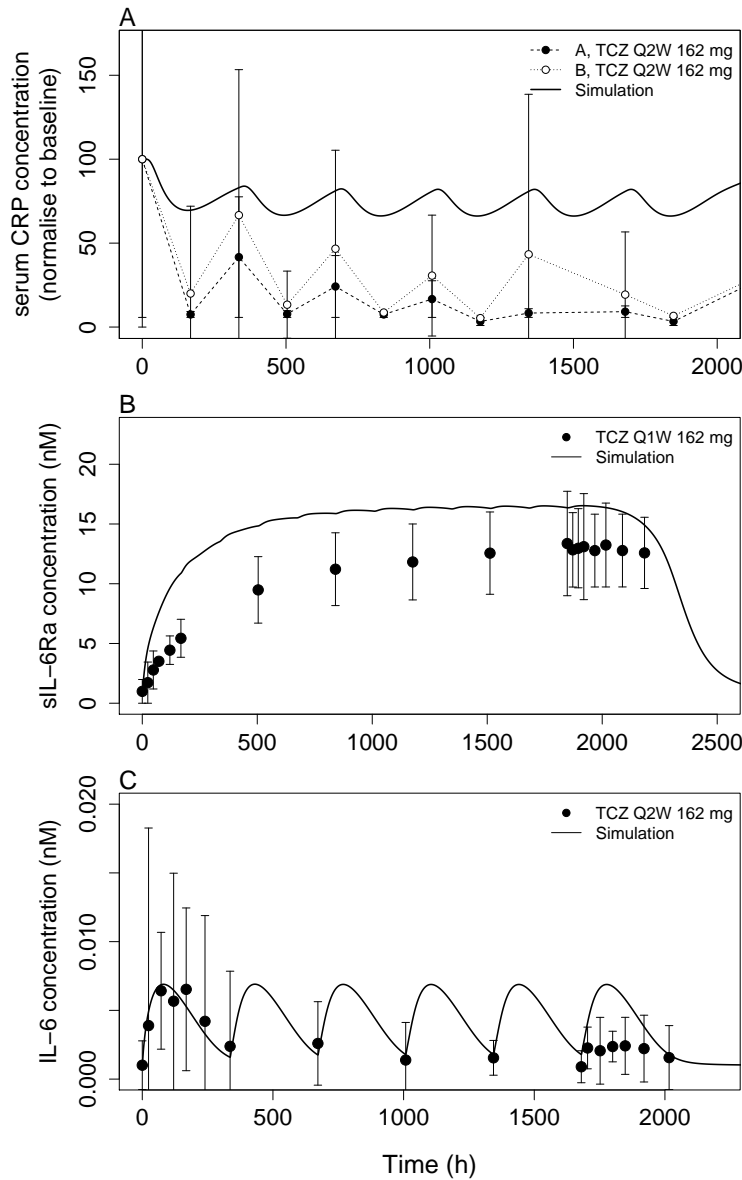
Figure 3.13: TCZ PD Validations. Simulated TCZ administration over several SC dose regimens attempts to match the serum PD biomarker responses observed in patients. **A**, CRP suppression from multiple doses; **B**, elevation of sIL-6Rα during dosing; and **C**, a rise in IL-6 (Zhang et al., 2013b).

3.3.2.1 *Can we improve the pharmacodynamics?*

Throughout the model development, some work was undertaken to understand the limitations of the model when reproducing the PD as well try and improve the fit. One thing of note was was the threshold for STAT3 phosphorylation by activating IL-6Rα when examining the output for CRP. The dose–response curve for pSTAT3 suggested that physiological hepatic levels of IL-6 within the liver (approximately 5
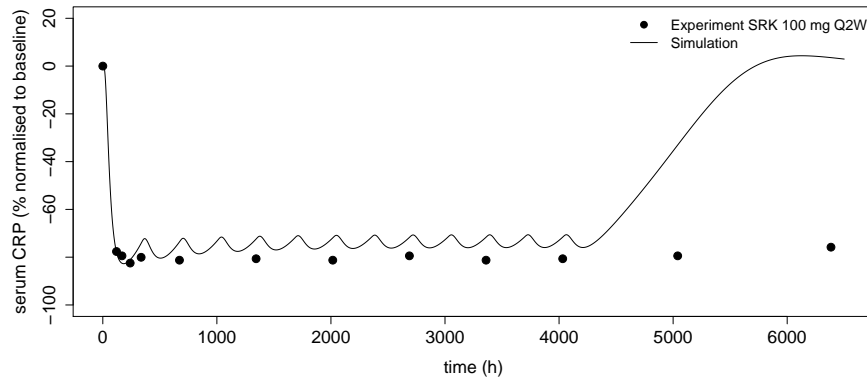
Figure 3.14: SRK PD Validations. Simulated CRP suppression during SRK SC administration (Smolen et al., 2014).

pM) would have very little response when using the phosphorylation curve I fitted to *in vitro* data (Figure 3.7B). The result of this is that the subsequent parametrisation of CRP synthesis plays solely upon the lower tail of the sigmoidal STAT3 phosphorylation curve in Figure 3.7B and remains too stiff to be modulated by mAb perturbations. This may be fixed by re-evaluating the STAT3 phosphorylation sub-model using *in vivo* or *ex vivo* data if available to gather quantitative evidence of the relationship between the hepatic IL-6 and pSTAT3. Other improvements to PD may be made by including an IL-6-dependent feedback mechanism on IL-6 production. This would act to both lower the IL-6 synthesis parameter, fixing the sustained response of both IL-6 and sIL-6Rα as observed in Figures 3.13B and 3.13C.

3.3.3 *Prediction*

To make predictions on the relative merits and caveats of IL-6 blocking strategies, we ran simulations. These compared the effects that the targets have upon a mAb-based therapy as well as the ability of each target to elicit a clinical response. Further predictions are made with specific comparisons of TCZ and SRK.

3.3.3.1 *Comparing drug targets for monoclonal antibody therapy*

Predictions compare the target druggability by assessing how either the receptor or ligand affects mAb PK. Simulations of an equal affinity mAb at a K_D of 1 nM demonstrate that targeting mIL-6Rα as opposed to IL-6 results in a much lower area under the curve and steady-state trough concentration (Figure 3.15). Furthermore, the difference in $t_{1/2} - \beta$ between the two mAb targets is apparent and nonlinear kinetics were also observed at the lower end of the mAb concentration. Nonlinear elimination of the anti–IL-6Rα mAb was expected as the K_D used represents the approximate affinity of TCZ.
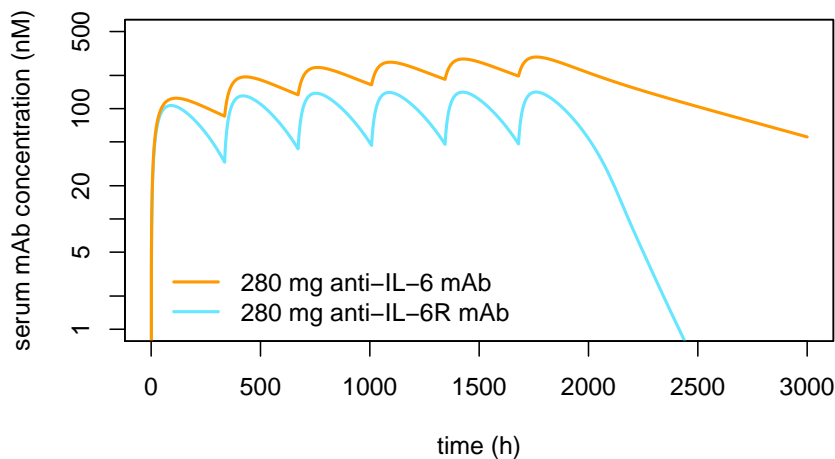
Figure 3.15: The effect of drug target on PK. Simulations were run for 12 weeks with a twice-monthly SC administration of mAb. A K_D of 1 nM was chosen for each mAb.

We made further simulations to predict the behaviour of the system in response to mAb administration. In clinic, the response to a mAb which targets either IL-6 or IL-6Rα can be seen through the suppression of CRP in patients. In our model, Dose–response predictions reveal that blockades of either IL-6 or IL-6Rα are able to achieve full CRP suppression which suggests that either target is viable (Figure 3.16a). However, the model predicts that the potency differs between

both strategies; the comparison of acute phase protein response using a mAbs with a one nM K_D suggests that targeting IL-6Rα with a mAb would be more beneficial than targeting IL-6, and the suppression of CRP is achieved at a lower dose when targeting IL-6Rα compared with IL-6 (Figure 3.16a). Furthermore, the trough to peak range in *serum* CRP after antibody administration appears to be wider for anti–IL-6 therapy at lower doses but dampens as the dosage increases. For anti–IL-6 therapy these pulses appear to maintain their magnitude (Figure 3.16a).

However, upon modifying the affinity of the mAbs we can see a change in the potency and behaviour of CRP suppression in both strategies. Decreasing the off-rate ($k_{Ab_{off}}$) over several orders of magnitude yields a typical shift to the left for the anti–IL-6 mAb but not substantially for the anti–IL-6Rα mAb (Figure 3.16b). Instead, the effect of a higher affinity in the anti–IL-6Rα mAb alters the response behaviour from a gradual sigmoidal curve to a dose-dependent switch (Figure 3.16b). In other words the leftwards shift is otherwise limited by a dose-threshold. This effect could be caused by the necessity of an anti–IL-6Rα mAb to first saturate sIL-6Rα as suggested in the *original model* (Dwivedi et al., 2014). In assuming this however, the contribution of classical signalling to STAT3 phosphorylation in hepatocytes must be greater than that of trans-signalling in order to regulate the output of CRP.
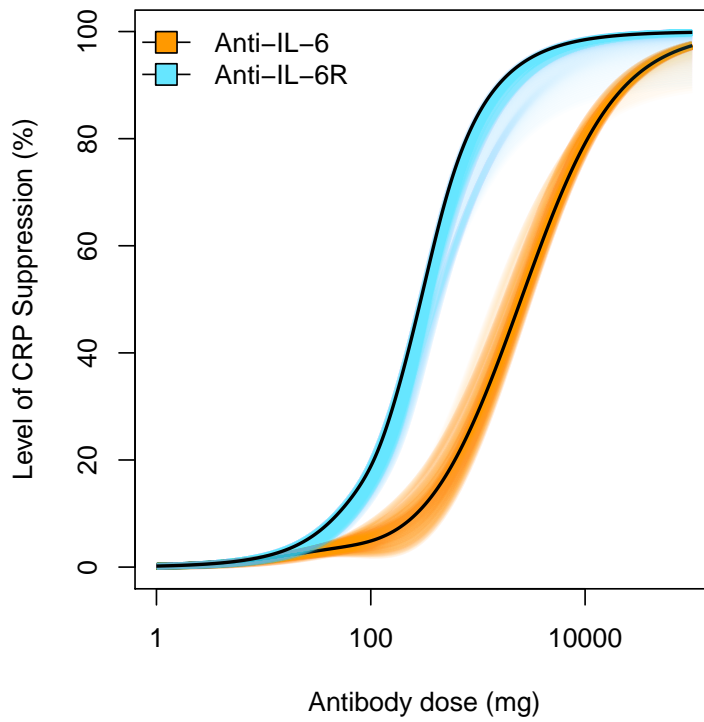
(a) Equal affinity dose–response comparisons.
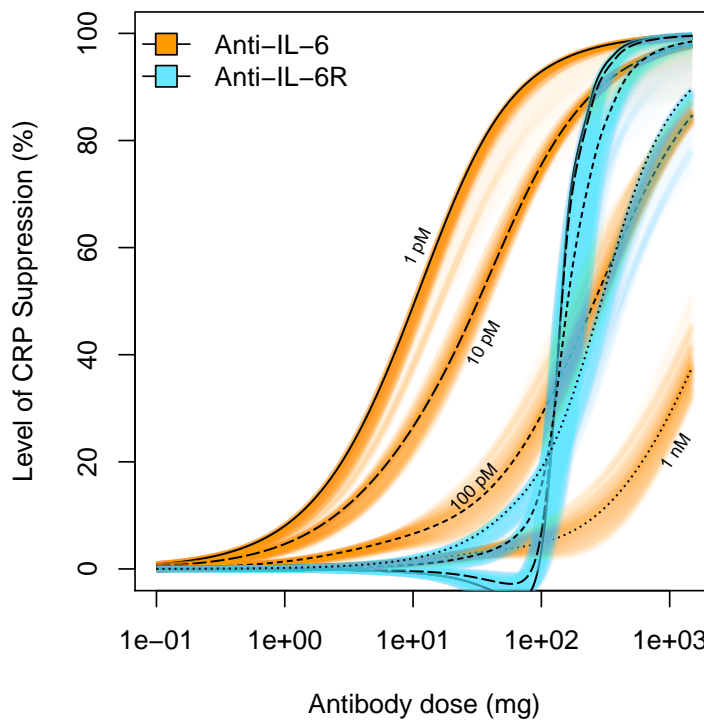
(b) Multiple affinity dose–response comparisons.

Figure 3.16: CRP was fully suppressed by the inhibition of either target. Simulations were run for 12 weeks with a twice-monthly SC administration of a mAb with the labelled K_D. Saturation of colour is the density of response values over the course of treatment. Lines are the final trough values after 12 doses.

3.3.3.2   *The majority of an anti–IL-6R mAb is bound to sIL-6Rα.*

To determine whether the pool of free sIL-6Rα was responsible for this dose threshold, the concentrations of hepatic bound and free IL-6Rα complexes were monitored over a range of simulated anti–IL-6Rα mAb doses. The model illustrated that the high affinity anti–IL-6Rα mAb has to first bind to and saturate the pool of soluble receptors before the drug can bind to the pool of mIL-6Rα. This is apparent because the mAb has an equal affinity for both membrane and soluble forms of the receptor. The model showed that a significant shift from free to bound mIL-6Rα only occurred when less than ten per cent of sIL-6Rα remains unbound (Figure 3.17). Additionally, the model also showed that on the approach to this critical dose-threshold, the concentration of active mIL-6Rα actually increases by 40 per cent (Figure 3.17). This suggested that a below-threshold dose of a high-affinity mAb targeting IL-6Rα may increase the signalling via classical pathways.
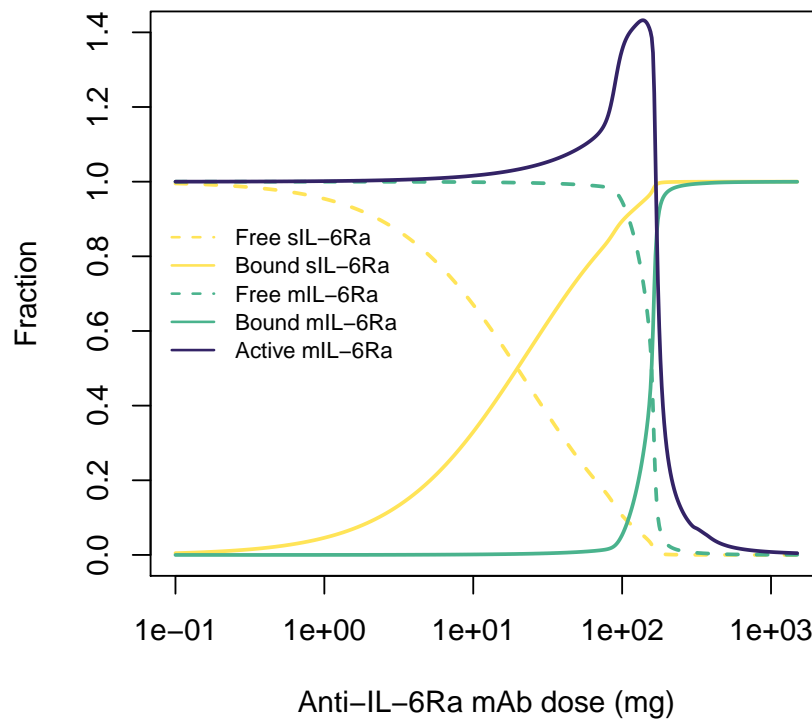
Figure 3.17: Free and mAb-bound IL-6Rα concentrations in the *liver* as a fraction of the total concentration. Active mIL-6Rα was normalised to RA the pre-treatment concentration (t =0). Simulations were run for 12 weeks with a fortnightly administration (Q2W) SC administration of a high-affinity anti–IL-6Rα mAb.

To further assess how the sequestration of TCZ by sIL-6Rα affects the dynamics of cell-surface receptor components in the liver, hepatic classical and trans-signalling components were monitored over a range of high affinity anti–IL-6Rα mAb doses (Figure 3.18, note the log scale). In the absence of antibody, mIL-6Rα and sIL-6Rα were mostly free. Simulations predicted that, at larger concentrations, the anti–IL-6Rα mAb competes with IL-6 for sIL-6Rα and which slows the degradation of the total pool of sIL-6Rα. Figure 3.18 shows that the total concentration of sIL-6Rα increased after exposure to the anti–IL-6Rα mAb, which agrees with clinical studies (Figure 3.13). The total concentration of sIL-6Rα peaks after 100 mg when sIL-6Rα is fully saturated by the mAb. At the same dose however, the total concentration of mIL-6Rα was seen to decrease when it suddenly became saturated by the mAb. This increases the rate of receptor internalisation as it is bound by the antibody. Furthermore, the active membrane-bound receptor was seen to disappear completely after dosing with 100 mg of a high affinity anti–IL-6Rα mAb. Sequestration of the anti–IL-6Rα mAb by sIL-6Rα is augmented by an increasing total concentration of the receptor. This is driven by the a reduction in sIL-6Rα degradation when bound to the antibody. Combined with the sustained synthesis reaction of sIL-6Rα, a new equilibrium was reached after continuous mAb administration (Figure 3.18).
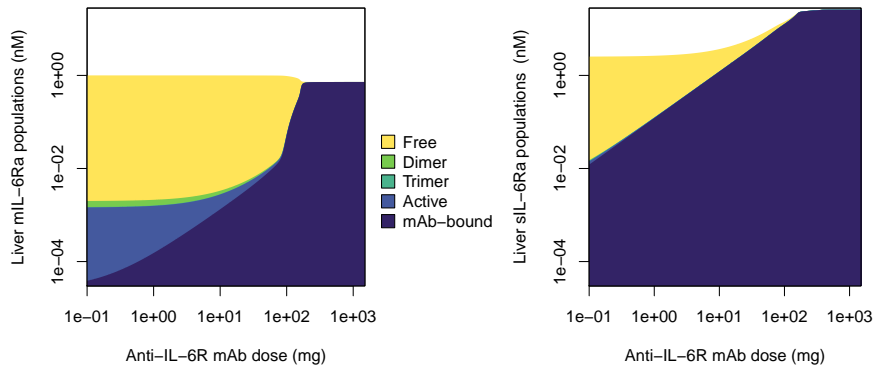
Figure 3.18: Hepatic classical (left) and trans-signalling (right) receptor saturation. Simulations were run for 12 weeks with a twice-monthly SC administration of a high-affinity mAb (K_D = 1 pM).
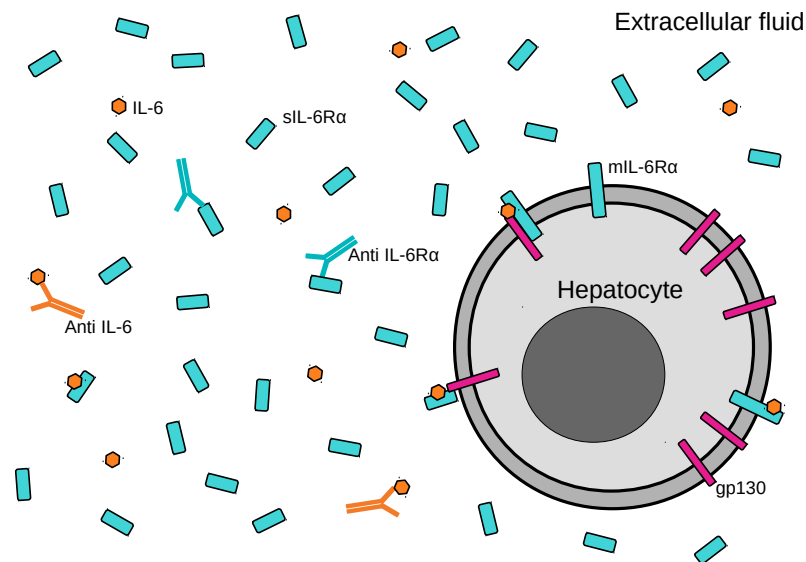
Figure 3.19: A diagram of the proposed saturation problem seen by comparing mAbs against IL-6Rα and IL-6. A mAb with IL-6Rα as a target may have to first bind to and saturate the overwhelming populations of rapidly synthesised sIL-6Rα across multiple tissues prior to binding to hepatic mIL-6Rα. Competitive binding to IL-6 on the other hand affects both classical and trans-signalling in unison. Furthermore, this effect is exacerbated because of differences between low IL-6 and high sIL-6Rα concentrations in extracellular space; the ratio between sIL-6Rα and IL-6 is several orders of magnitude greater still than that depicted.

3.3.3.3 *Comparing sirukumab and tocilizumab.*

As the majority of mAbs are competitive inhibitors, the therapeutic difference between two mAbs of the same class is often associated with their affinities. We simulated the dose–response of CRP using mAbs with affinities similar to those seen in TCZ and SRK with K_Ds of one nM and and ten pM respectively. The model predicted that the SRK-like mAb was more potent in terms of end-of-course CRP response and a full reduction of serum CRP was achieved at a much lower dose of SRK compared with TCZ (Figure 3.20). The serum EC_{50}s of SRK and TCZ were 23.1 nM and 88.3 nM respectively, delivered by doses of 32 mg (SRK) and 293 mg (TCZ) SC, Q2W for 12 weeks.
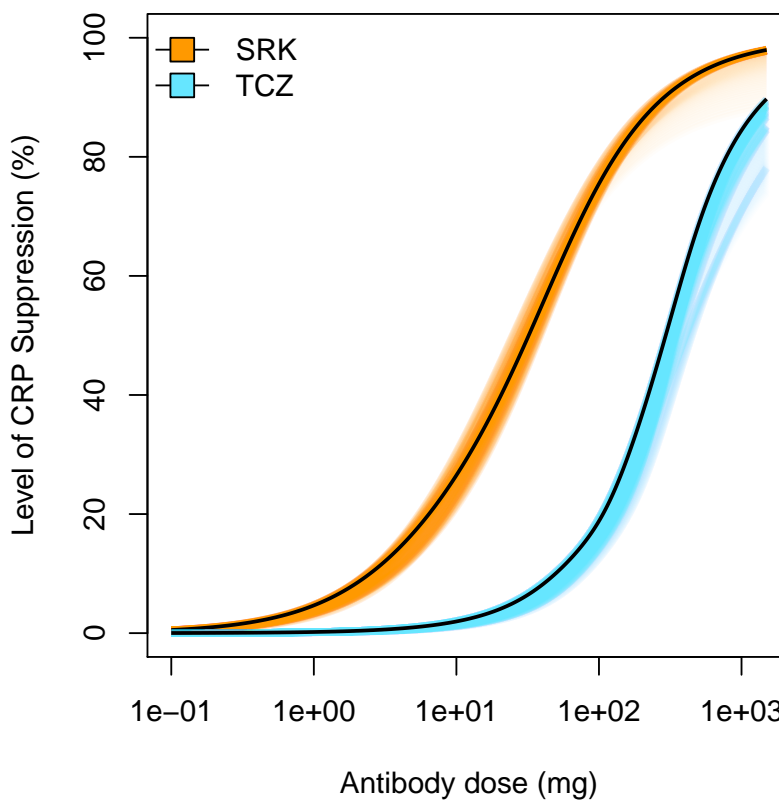
Figure 3.20: Dose–response of TCZ compared with SRK. Simulations were run for 12 weeks with a twice-monthly SC administration of a mAb with an affinity in the range of TCZ or SRK. The K_Ds used here were one nM and ten pM respectively. Colour saturation is the density distribution of response values over the course of treatment. Lines are the final trough values after 12 doses.

Because we see that the potency of SRK is greater than TCZ (Figure 3.20), it may be possible to use a substantially lower dose for SRK and still achieve the same level of CRP suppression as TCZ. Using the same

dose schedule and route of administration as TCZ, the model predicts
that an SRK-like mAb can achieve the same CRP response as TCZ with
a dose of 12 mg (Figure 3.21). This is half the lowest strength of the
proposed set of doses for SRK used in clinical trials but administered
more frequently. In making this comparison I assume that CRP is a a
valid surrogate biomarker for clinical efficacy in anti–IL-6 therapy for
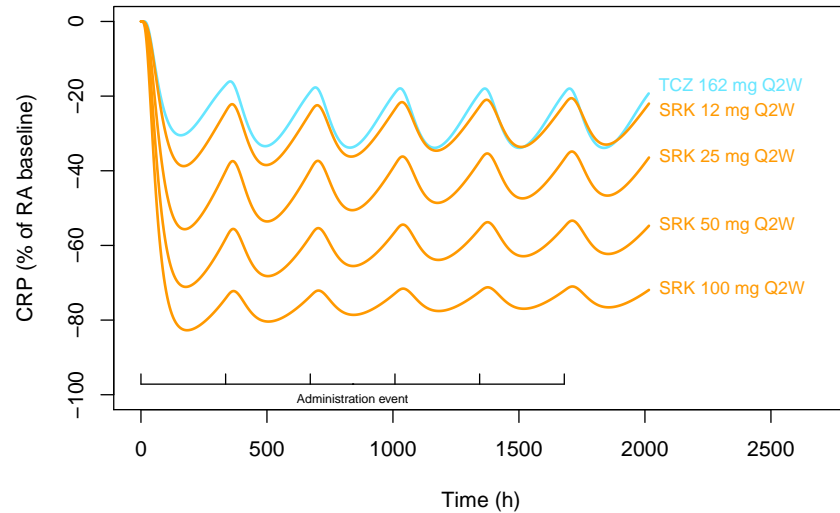treating RA.

Figure 3.21: Lower doses of SRK suppresses CRP similarly to TCZ. Simula-
tions were run for 12 weeks with a twice-monthly SC adminis-
tration of SRK or TCZ.

To see the effect of each mAb within the target tissue, the *syn-
ovium*, we assessed the concentrations of IL-6 and the trans-signalling
complex sIL-6Rα. This may give a more relevant comparison of the
therapeutic effect in arthritis. We noticed that the administration of
either SRK or TCZ reduced the concentration of synovial IL-6-bound
sIL-6Rα (Figures 3.22a and 3.22b). However, this effect was greater
using SRK, with a ten-fold reduction in synovial [sIL-6Rα:IL-6] by the
sixth dose (Figure 3.22b). During treatment with both mAbs, the *serum*
concentration of the trans–signalling complex was seen to fluctuate
around the RA baseline concentration. However, upon the cessation
of therapy, the SRK treatment resulted in a dramatic increase in serum
[sIL-6Rα:IL-6] (Figure 3.22b). Interestingly, this may be caused by an-
other observation; Figures 3.22a and 3.22b also show that both mAbs
elevate serum IL-6 during treatment. While this effect has been ob-
served for TCZ in clinical studies (Zhang et al., 2013b), it may be sur-
prising following the administration of SRK as SRK is a selective IL-6
antagonist. The nature of the IL-6 response to SRK suggests a rebound

effect which occurs as the course of therapy is terminated. This is presumably caused by the binding of IL-6 to SRK, and the protection from degradation that SRK bestows upon IL-6 as the first-order elimination rate of the complex is lower than that of the cytokine.
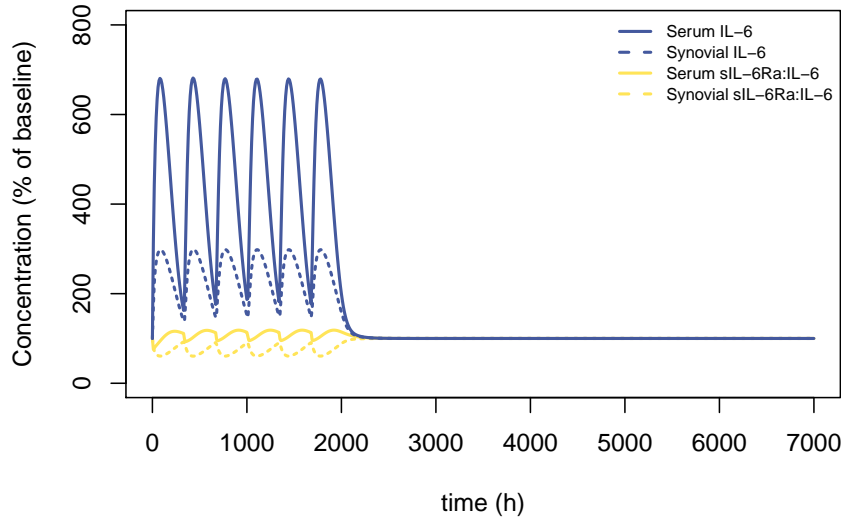
(a) The effect of TCZ upon trans-signalling
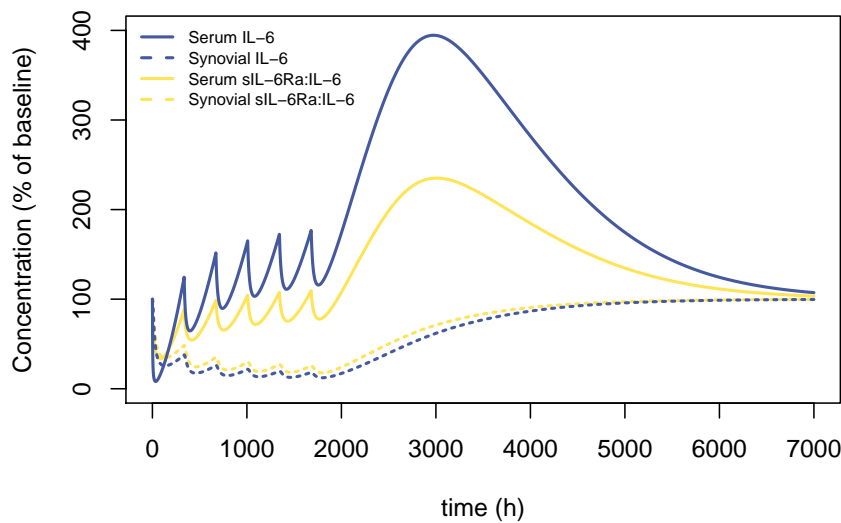
(b) The effect of SRK upon trans-signalling

Figure 3.22: Simulated dose regimen for 12 weeks with a twice-monthly SC administration of (a), TCZ at 162 mg or; (b), SRK at 50 mg.

We further explored this phenomenon by producing a a dose–concentration curve of SRK against *serum* IL-6 after 12 weeks of administration (Figure 3.23). This illustrates that if a rebound were to occur following SRK treatment it may be seen even at low doses and at lower frequen-

cies. Furthermore, doses above 100 mg appear to mitigate against this IL-6 rebound immediately after treatment. However, the timing of IL-6 measurement is key to the simulation in Figure 3.23; Figure 3.22b suggests that the peak rebound concentration of IL-6 after treatment occurs after a month while Figure 3.23 shows the concentration immediately following the 12-week course of SRK therapy (2016 hours). Therefore it is highly plausible that higher simulated doses of anti–IL-6 mAbs still cause a rebound effect, increasing the concentration of IL-6 weeks after therapy.
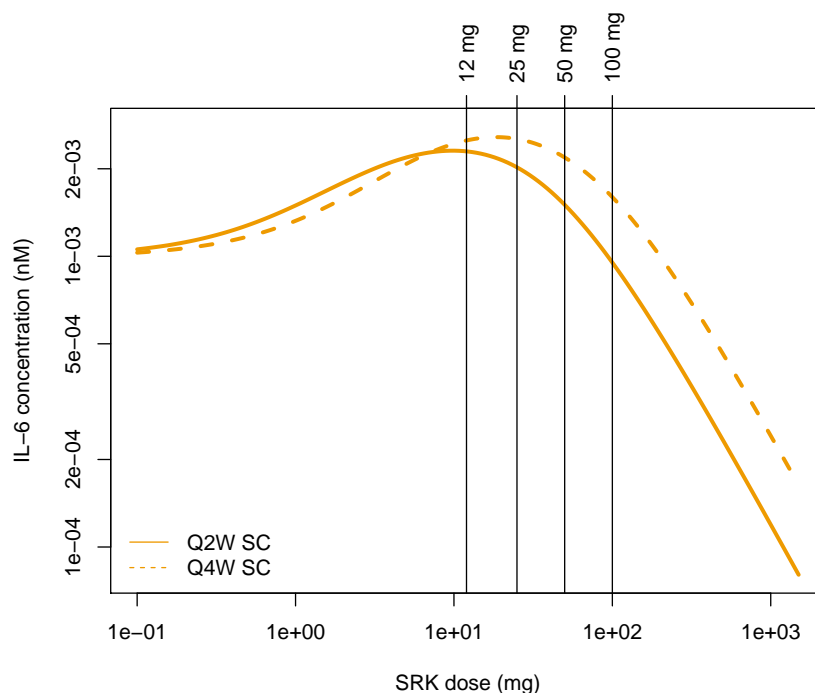
Figure 3.23: The response of *serum* IL-6 to a potent anti–IL-6 mAb. Administration of an anti–IL-6 mAb may cause an IL-6 rebound. Simulated dose–response of an SC administered SRK-like mAb increases the concentration of IL-6 as measured after 12 weeks.

3.4   DISCUSSION

A QSP model was re-purposed to address questions on druggable target selection and disease mechanisms within the scope of RA. The CD model by Dwivedi et al. (2014) was broken down into its individual modules and parametrised using published *in vitro*, *in vivo* and clinical data from cell lines, animals and arthritic patients. The process outlines a case study in the reuse of QSP models, the types of answers sought, and the questions raised in probing our knowledge of biology. The effort in re-purposing this model from CD to RA was not without

errors. The validation simulations question the model's structure and parametrisation and the results of estimation steps suggested identifiability issues. Furthermore, the process was more laborious than anticipated due to the absence of key quantifiable data.

Despite the challenges faced, the model was able to fit the PK data of TCZ as well as predict the PK of SRK demonstrating that the assumptions, structure, and parametrisation were at least sound enough to simulate novel data. Moreover, the model forms emergent predictions relating to target selection between IL-6Rα and IL-6 as well as comparisons of the mAbs TCZ and SRK. Finally, these predictions generate a novel hypothesis of IL-6 rebound in SRK therapy which can be tested in clinical studies as a followup.

3.4.1   *Successes and challenges in model parametrisation.*

During the process of re-purposing and parametrisation, reactions were modified, parameter values altered and identifiability issues were raised. Primarily, to explore the causes of TMDD in TCZ, a large peripheral compartment was populated with the receptor turnover reactions which gives evidence for the receptor's role as a large part of TCZ's saturable elimination. Additionally, the receptor turnover submodel was re-calibrated using further experimental values from literature. To account for the differences between diseases, all biological parameter values were altered during the construction of the model and assumptions were made to the best of the knowledge surrounding the disease mechanisms of RA as opposed to CD (see Materials and Methods). In turn, the parametrisation also explored and often agreed with hypotheses driven by experimental evidence.

3.4.1.1   *TCZ pharmacokinetics*

Previous PK models have used MM approximation for nonlinear elimination routes of TCZ although the molecular basis of this effect was not covered by their model's granularity (Gibiansky and Frey, 2011). Another model discusses that the nonlinearity is reminiscent of a target-mediated elimination via mIL-6Rα (Frey et al., 2010). At this time there was no strong evidence that the target for TCZ, mIL-6Rα, internalises efficiently without binding to gp130. Nevertheless, it was still considered that TCZ internalises alongside the receptor (Ritchie et al., 2013). The history of IL-6 signalling studies implicates gp130 as the main player in receptor endocytosis which suggested that, if TMDD was occurring, TCZ could bind in a complex with mIL-6Rα and gp130 and internalise through gp130-mediated routes. Addressing this was a more recent study which defines the mechanisms of TCZ TMDD (Fujimoto et al., 2015). Fujimoto et al. (2015) demonstrated that not only did mIL-6Rα-dependent internalisation occur with TCZ at a similar rate to IL-6-bound mIL-6Rα, but also that it does so independently of

gp130. The internalisation is mediated by an mIL-6Rα cytoplasmic motif through clathrin-dependent pathways. It is uncertain whether TMDD is driven only by the classical signalling component of IL-6 signalling or also by trans–signalling too. These may have different degradation or internalisation rates as modelled by Gibiansky and Frey (2011) or affinities for TCZ (*U.S Food and Drug Administration, BLA: 125276*).

The dose-dependent nonlinear clearance kinetics of TCZ were readily reproduced in simulation by using parameter values derived only from published studies, including mIL-6Rα concentrations based solely from interstitial T cell counts, hepatocyte numbers at physiological levels, and by assuming that the internalisation of TCZ-bound mIL-6Rα occurs at a similar rate to IL-6-bound mIL-6Rα. Therefore, our PK model as described confirms that TMDD via mIL-6Rα is a likely explanation for the nonlinear elimination of TCZ *in vivo*, although sIL-6Rα-mediated sequestration and elimination may also contribute to the overall effect. The parameter set acquired during estimation of the whole model suggested that the complex of TCZ and sIL-6Rα degrades faster than TCZ on its own which implies a role for the pool of soluble receptors in TMDD.

3.4.1.2 *Receptor Dynamics and Turnover*

The parameters for the receptor dynamics were re-estimated and differed from those of the *original model* to account for several factors. Experimental evidence suggested different basal internalisation rates for both mIL-6Rα and gp130. Furthermore, the ratio of gp130 and mIL-6Rα are likely to be different in wild-type hepatocytes than cells which artificially over-express the receptor such as those used for the CD model's estimation. Wild-type human hepatocyte data would certainly disperse the ambiguity.

The parametrisation of the receptor dynamics model provides an interesting view on receptor cell-surface expression profiles. The first observation is that the cell-surface expression of gp130 was estimated to be higher than that of mIL-6Rα in fitting HepG2 data which agrees with a variety of experimental findings. For example, early investigations showed that cell-surface gp130 is expressed at much higher levels than mIL-6Rα in HepG2 cells (Hibi et al., 1990; Taga et al., 1989) and that the amount of mRNA for mIL-6Rα is far lower than gp130 in all patient groups in chronic *liver* disease (Lemmers et al., 2009). The evidence implies that it may indeed be reasonable for gp130 to be more highly expressed than mIL-6Rα in human hepatocytes. On this note, the expression ratios of IL-6 receptors in HepG2 cells may be different to hepatocytes *in vivo*; similar strengths of fluorescence were observed in IL-6-stimulated hepatocytes when comparing the surface expression of both gp130 and mIL-6Rα (Memoli et al., 2010). Additionally, in largely different cell lines such as human multiple myeloma cell line (U266) cells or transfected jurkat T cells, the ratio of gp130

to mIL-6Rα was less than one (Hibi et al., 1990). Clearly there is a wide spread of mIL-6Rα expression on the cell-surface within different cell populations. This is exemplified in a similar study instead using mIL-6Rα-transfected COS-7 cells where the rate of internalisation appears to be much slower (Dittrich et al., 1996). The parameter scan altering gp130 synthesis in the materials and methods section implies a slowing of IL-6 internalisation in lower ratios of gp130 to mIL-6Rα (Figure 3.3), which mimics the experiments by Dittrich et al. (1996).

3.4.1.3 *Intracellular signalling and STAT3*

The goal of parametrising the STAT3 phosphorylation reactions was to simulate the dose–response profiles of pSTAT3 after IL-6 administration seen in *in vitro* experiments. In doing so, our model at least describes a phenomenological model of IL-6-mediated cellular response and measurable outcome via STAT3 phosphorylation.

Issues were raised in the practical identifiability of our model as a number of parameter's minima were uniformly distributed across a large range of values. I argue that the our parameter estimation was successful in fitting to the data from multiple cell lines by reproducing the behaviour of STAT3 phosphorylation *in vitro* in response to IL-6 stimulation. The dose–response curve generated using the model results in a reasonable fit to multiple cell-lines and independent experiments. We assumed that the broad coverage of data using different cell types describes a general pSTAT3 response. Furthermore, the model shows a documented time-dependent effect of IL-6 stimulation. Experiments *in vitro* show that a short-lived peak of pSTAT3 occurs rapidly after the administration of IL-6 (Ara et al., 2013; Niemand et al., 2003; Thiel et al., 2000) (Figure 3.24). This has also been successfully modelled with more complex intracellular signalling systems biology models by including negative feedback mechanisms (Moya et al., 2011; Qi et al., 2013; Singh et al., 2006). However, in our simplified model, the peak occurred a little later than in these experiments.

The *original model* suggested that a 3-fold increase of pSTAT3 is observed between Healthy and Disease patients. This is illustrated by a variety of studies in inflammatory disease states in both in CD3+ cells or rat hepatocytes after IL-6 stimulation (Anderson et al., 2015; Henkel et al., 2011). However, within this reaction scheme, this behaviour was not identifiable. No parameter set would produce the basal pSTAT3 levels at steady-state under normal IL-6 concentrations. This suggests that the structure of STAT3 phosphorylation within model does not fully capture the dynamic range of STAT3 present in biology. This may be solved by the addition of a basal phosphorylation reaction which phosphorylates STAT3 independently of IL-6. IL-6-independent STAT3 phosphorylation may be the cause for unstimulated cell pSTAT3 responses seen in experiments (Henkel et al., 2011). Furthermore it
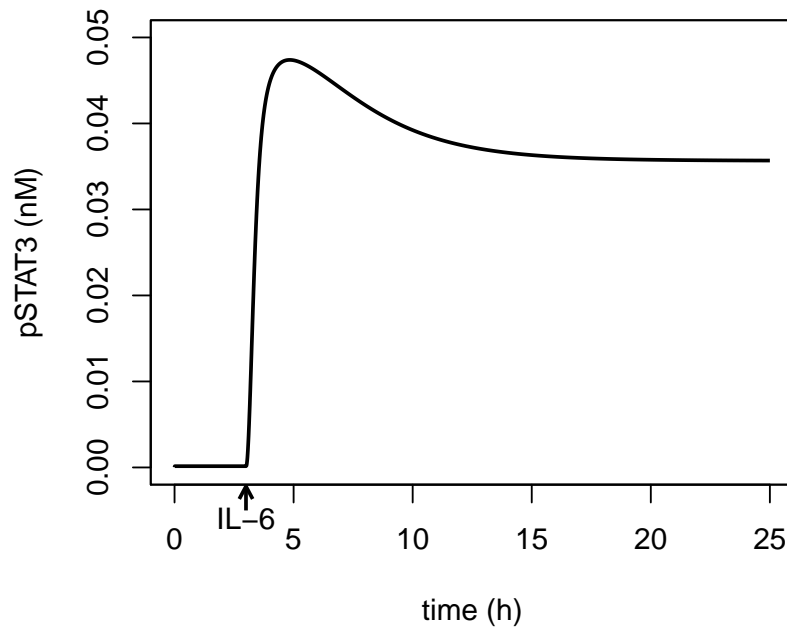
Figure 3.24: IL-6 administration results in a rapid transient peak in pSTAT3 concentrations.

is questionable that basal phosphorylation of STAT3 is caused by IL-6 as available published experiments on STAT3 phosphorylation suggest that the picomolar physiological range of IL-6 would have very little effect indeed upon the range of STAT3 phosphorylation, at least according to *in vitro studies*. For instance, the *in vitro* EC50 of IL-6 was 0.1, 0.9, and 0.046 nM in HuH7, human Cal33, and 7TD1 cell lines respectively (Casanovas et al., 2014; Gough et al., 2016; Simard et al., 2014). Therefore, although IL-6-independent phosphorylation is likely, it is currently outside of the model scope. Lastly, the number of molecules and the concentration of STAT3 in the model is most likely underestimated. Future studies would build upon estimating a more realistic value for the concentrations and revision of structure alongside true human hepatocyte STAT3 data.

3.4.1.4 *Tissue compartments, CRP and steady-state concentrations*

Without *in vivo* time course data across multiple tissues in humans it was difficult to determine suitable values for the distribution rates. In a similar fashion to the PK parametrisation process, a single macro parameter Q_{PD} was used for all components but a more realistic approach could be applied in using a PBPK model with measured or estimated flow parameters for each compartment. Furthermore, the availability of multi-compartmental time course data for signalling

components would certainly assist in parametrising the distribution reactions for a more mechanistic model.

No value for the rate of sgp130 synthesis was found which further minimised the objective function. Furthermore, the lack of parameter identifiability and model sensitivity to the estimations of sgp130 synthesis were consistent with the previous conclusions regarding sgp130 as a poor therapeutic protein (Dwivedi et al., 2014) (Figure 3.9). Additionally, the estimation of $k_{R_{shedding}}$ suggests that the current model structure offers little place for CRP-mediated mIL-6Rα shedding. I argue that the extent of CRP-mediated shedding *in vivo* by hepatocytes is not fully understood and CRP is only one of several reported causes of mIL-6Rα solubilisation (Jones, 2001; Lust et al., 1992; Matthews et al., 2003; Schumacher et al., 2015). Perhaps the specific contribution of CRP to mIL-6Rα shedding is smaller than theorised in this situation. That said, a biological mechanism must exist which is responsible for the increase in sIL-6Rα in elevated IL-6 patients such as those seen with RA.

In validating the model with TCZ data, the elevation in IL-6 or sIL-6Rα was not properly simulated. One reason may be that the model does not account for any regulation upon the rate of IL-6 synthesis. The model instead assumes a fixed IL-6 synthesis for generating RA steady-state concentrations. However, clinical data suggested that IL-6 increases transiently but tends towards or below the pre-treatment concentrations after several weeks of therapy. This is evident as the magnitude of the experimental peak after multiple doses is smaller than the simulated peak implying a reduction in IL-6 *in vivo* (Figure 3.13C). An improvement to the disease model would be to include a mechanism for the disease progression which incorporates a positive feedback mechanism for IL-6 synthesis responsible for an increased IL-6 synthesis rate in the disease state. The ability to reverse this progression could then be simulated.

The reason that simulated elevation of sIL-6Rα may not fit the experimental data is that the synthesis and degradation parameters are unknown and therefore may not truly represent those apparent in nature. The degradation rate of sIL-6Rα was fixed prior to estimation utilising approximate *serum* half-life values for similar soluble cytokine receptors (Jacobs et al., 1993). Although this is a reasonable assumption, a multitude of unknown biological reactions may influence this observed half-life and alter the model parameters mediating sIL-6Rα's steady-state concentration. However, contrasting with the *original model*, our parameter set fitted the elevation in sIL-6Rα more closely where the previous model may have been fitted without sufficient CD data.

The rate constants for CRP synthesis were difficult to estimate and the ability of TCZ to suppress the acute phase response was underpredicted. A reduction in CRP was observed which only qualitatively

represented experimental data. One reason may be the response threshold of the STAT3 module; physiological RA concentrations of IL-6 in the *liver* are in the range of one pM to ten pM. However, data from *in vitro* studies used in validation suggest that STAT3 is barely phosphorylated at these disease, let alone healthy, concentrations of IL-6. The model may be too insensitive to simulate reduction in pSTAT3 following TCZ administration at this dose. Alternatively the receptor activation reactions may not be parametrised effectively which together highlights a lack of data for these reactions and greater identifiability issues within the model. Finally, the simulated suppression of CRP was further limited by the fixed synthesis rate of IL-6 in the model without a feedback mechanism.

3.4.2 *Model predictions*

The model was used to predict the different effects of targeting either mIL-6Rα or IL-6 in RA using mAbs. Furthermore, the model used these predictions to compare the perturbative capacity of two mAbs which are either in clinical use (TCZ) or currently in development (SRK) for treating IL-6-mediated diseases such as RA. The two different targets regulated the PK of each mAb as well as the behaviour of the response. Both of which have implications for their use as druggable targets.

3.4.2.1 *Druggable target selection*

The differences in between targeting either IL-6Rα and IL-6 are apparent firstly by observing the PK profiles. Our model demonstrates that the target is the largely responsible for this effect as simulations of equal affinity mAbs show dramatically different PK profiles (Figure 3.15). Targeting IL-6Rα opens a rapid elimination route through receptor internalisation for drugs in contrast to IL-6. This is a typical phenomenon often seen in receptor-targeting mAbs such as TCZ and is of consequence as linear elimination kinetics are easier to predict, monitor and subsequently dose in patients.

Overall, targeting the receptor is seen to be strongly dose-dependent in terms of pharmacokinetics and the rate of elimination, which limits the duration of drug–target localisation. Additionally, the response behaviour is sub-optimal as a large dose has to be achieved to pass the critical dose-response threshold. Targeting the ligand on the other hand provides a predictable linear PK and, as the response is affinity-dependent, there is the potential for improving the binding and therefore potency throughout optimisation phases in drug development.

Selecting between druggable targets within a biological network is a challenging task and there are many variables to consider. Based on these simulations it would be wiser to select IL-6 as a target instead of IL-6Rα to advance along a drug development pipeline. That said, in the case of IL-6 signalling, trans-signalling may be respons-

ible for the deleterious effects (Garbers et al., 2015; Rose-John, 2012). If so, this preferential saturation of sIL-6Rα instead of mIL-6Rα by an anti–IL-6Rα mAb may be favourable for mitigating against both the on-target toxicity and the blanket blockade of IL-6 that may attenuate normal function.

To address mechanistic toxicity, the model could be improved by including specific toxicological modules. These could include and explore immune system reactions by further incorporating neutrophil activity to address the risk of infections, the rise in low-density and high-density lipoproteins, or perhaps even an increased risk in malignancy as seen following anti–IL-6Rα treatment (Hennigan and Kavanaugh, 2008).

3.4.2.2   *Tocilizumab or sirukumab*

From a pre-clinical perspective, a fresh pair of targets to choose between offers a decision-making task with the promises of compound modification or optimisation of formula further down the developmental pipeline. In this project instead, the focus was upon the mAb TCZ, with demonstrable success in clinic against SRK, a mAb which has yet to be approved for normal clinical use.

PHARMACOKINETICS AND PHARMACODYNAMICS   The linear elimination kinetics of SRK compared with the nonlinear kinetics of TCZ suggests that the dosing of SRK may be easier to scale for clinical use. Nonlinear kinetics are often caused by RME and can be problematic for drug dosing.

Furthermore, in terms of cost and ease of use to patients and health care professionals, a high-affinity mAb with linear kinetics such as SRK also has the benefit of a reduced cost per dose as well as the potential for a reduced subcutaneous injection volume. However there may even be further room for dose optimisation in SRK with regimens not explored in clinical trials. We considered that clinical trials of SRK in arthritis patients using dose regimens of 50 mg monthly administration (Q4W) and 100 mg Q2W recorded a CRP response of 98 per cent. This suggests that the treatment may be refined to include smaller doses.

One outcome of the model was to suggest that the linear PK and high-affinity properties of SRK can be used to reduce dose magnitude and increase frequency. The model predicts that SRK achieves the same CRP suppression response as TCZ using lower doses than those used in clinical trials (Figures 3.21 and 3.20).

It is important to note that the CRP response was not fully fitted or validated using the clinical data for TCZ. However, at the dose selected for the anti–IL-6Rα mAb, TCZ has been shown to be effective in treating the disease as monotherapy and can be measured through its suppression of CRP. Therefore, a comparison can be made on the

basis that TCZ is clinically viable using this dose and we can thus infer equivalence for SRK from a simulated reduction in *serum* CRP.

THE SAFETY OF ANTI–IL-6 THERAPEUTICS   The anti–IL-6 mAb SRK failed to be approved by the FDA for its use in treating RA in humans as there were some concerns with safety at the proposed dose regimens (Johnson & Johnson, 2017). The exact causes of this were not clear. However, patient deaths may have been attributed to major adverse cardiovascular events (MACE) while a number of infections and infestations were also reported (FDA Briefing Document, Arthritis Advisory Committee Meeting, August 2017). Moreover, MACE were seen to be higher in the trials without a correlation with lipids and no other explanation was offered. That said, a recent clinical study assessing the use of SRK in RA showed that no MACE were reported (Takeuchi et al., 2018). However, this could have been be due to a lower relative incidence of MACE in Japanese patients in comparison with the global population.

Although the model was not developed with a dedicated safety sub-model or set of reactions, the results provide substrate for a hypothesis on the adverse effects patterns observed with anti–IL-6 therapy. Our simulations showed that the rebound in IL-6 was mirrored by a rise in circulating sIL-6Rα:IL-6. Upon careful inspection, it was apparent that a rebound only occurred in the *serum* compartment, leaving the *synovium* with lower concentrations of free IL-6 and sIL-6Rα:IL-6. This could be best explained by arthritic synovial tissues having concentrations of IL-6 at orders of magnitude above other tissues. Therefore, perhaps IL-6 is transported out of the *synovium* and into the *serum* by strong anti–IL-6 mAb binding, augmented by the difference between mAb influx and efflux rates through the synovial capsule. The disparity between the *synovium* and *serum* concentrations of IL-6 and sIL-6Rα:IL-6 after treatment makes for a plausible and compelling hypothesis: the therapeutic effects of anti–IL-6 mAbs in RA patients may be seen quite readily in the reduction in synovial IL-6 signalling and pSTAT3 but the increase in *serum* unbound as well as bound IL-6 may account for unknown systemic side effects.

For instance, the increase in sIL-6Rα:IL-6 might contribute towards side effects in the broader context of anti–IL-6 therapy by considering the evidence that places trans-signalling as a the pathological facet of IL-6 signalling (Rose-John, 2012). Moreover, increasing the concentration of systemic trans-signalling is potentially relevant in cardiovascular events (Fontes et al., 2015; Morieri et al., 2017; Ziegler et al., 2018). It is also interesting to note that clinical evidence presented to the FDA arthritis advisory committee revealed that the incidence of MACE was seen mostly in the 50 mg Q4W group whereas the the incidence in the 100 mg group was similar to placebo (Vratsanos, August 2,

2017). These clinical results agree with our IL-6 rebound simulations in Figure 3.23 where the post-therapy IL-6 rebound at 50 mg doses was higher than at 100 mg doses.

In light of these simulations I propose that the potential link between anti–IL-6 therapies and trans–signalling-mediated side effects could be further examined and confirmed in experiments. Firstly however, we must debate whether this simulated rebound translates into a biological setting.

The first argument against the existence of an IL-6 rebound in peripheral compartments is the model structure as IL-6 synthesis in the model is fixed to produce the RA steady-state. This is an important assumption to consider as the synovial secretion of IL-6 is proportional to the severity of inflammation. Therefore, because IL-6 synthesis is fixed in our model, we do not account for the relationship between IL-6 secretion rates and the level of synovial inflammation. By factoring this relationship in our model, we may instead find that the IL-6 synthesis rate decreases throughout the course of treatment which in turn could reduce the accumulation of cytokine–antibody complexes. A second argument against this hypothesis is that the degradation rate of the mAb:IL-6 complex in the model is assumed to be no different than the elimination rate of mAb. A lack of contrary evidence for this resulted in this being the assumed value. A third possible reason for disputing these predictions is that anti–IL-6 mAbs have been seen to reduce depressive symptoms in RA patients during clinical trials (Aletaha et al., 2017; Lindqvist et al., 2009). This could be taken as evidence that systemic IL-6 signalling may indeed be reduced after anti–IL-6 therapy, refuting an IL-6 rebound. However, perhaps anti–IL-6 therapy scrubs IL-6 from a CNS compartment as described for the *synovium* here. This might also reduce only local concentrations of CNS IL-6 if subject to constraints of compartmental permeability towards the molecular species in the *serum*.

Existing observations of the rebound phenomena with mAb-based therapies substantiate this hypothesis. Evidence suggests that the mAb infliximab elicits a positive three to four-fold change in circulating TNF-α, the cytokine that it is highly selective for (Chung, 2003; Koller-Strametz et al., 1998). Another example is seen after treatment with the anti–interleukin 5 (IL-5) compound, SCH55700 (Kim et al., 2004). Kim et al. (2004) go on to discuss that the prolongation of IL-5 likely depends upon the molar ratio of the cytokine and the monoclonal antibody, and that the rebound results from the decreased clearance of the drug-target complex. Also noted is that a higher dose of the anti–IL-5 antibody prevents the apparent rebound of IL-5, similarly seen in figure 3.23. Finally, this IL-6 rebound occurs in patients with cu-

taneous lupus erythematosus when treated using SRK (Szepietowski et al., 2013).

3.4.3   *Concluding remarks*

The various merits and caveats of drug targets may not be teased solely from *in vitro* studies. Even *in vivo* studies may have trouble in defining a mechanistic basis of PK, PD and toxicity associated with potential targets.

In addressing the aim of predicting target druggability, the model initially presents IL-6 as a better candidate for mAb-based therapy. This was due to tunable PD behaviour and less of an impact upon drug PK. In contrast, the IL-6Rα PD response was seen to be a dose-dependent switch in cases of high drug affinity and nonlinear PK was seen when targeting IL-6Rα. Simulations predict that, by using an anti–mIL-6Rα mAb, larger systemic concentrations are required for a significant clinical effect while an anti–IL-6 mAb would need a substantially smaller concentration for the same response.

When we compared the two drugs, the model suggested that dose regimens for SRK lower than those currently used in clinical trials may have a similar therapeutic effect to TCZ but no conclusion could be made about specific measures of toxicity or safety between the two mAbs. That said, a rebound effect with anti–IL-6 therapy emerged which was revealed by studying the results of the re-purposed model. The offloading of the active trans-signalling complex in high concentrations was unexpected and may begin to explain the nature behind FDA concerns in using anti–IL-6 mAbs.

While the SIRROUND trials still continue for the use of SRK in the treatment of RA (ClinicalTrials.gov; NCT01856309, 2013 - [cited 2018 March]), another clinical study is carried out for its use in treating major depressive disorder (ClinicalTrials.gov; NCT02473289, 2015 - [cited 2018 March]) which may reveal more information on the likelihood of an IL-6 rebound.

Finally, A key component in driving the use of QSP is the ease at which it is applied. By demonstrating that new models need not be made from scratch, an attempt was made to show that reuse of QSP models can be used for pharmacological predictions in a limited timeframe. However many challenges were faced in obtaining parameter values, re-using code and simulating experiments. As speed is of particular importance to making strategic decisions within pharmaceutical research, this highlights the necessity that these processes need to be standardised and streamlined for QSP to be integrated effectively in drug development pipelines.

4

EXTENDING THE CLASSIFICATION MODEL FOR PATIENT RESPONSE PREDICTION.

4.1 INTRODUCTION

One hundred years after what was arguably the greatest pandemic in human history, infections caused by the influenza virus are still a problem today. One concern now with influenza is the annual winter epidemics which occur across the globe. The World Health Organization (WHO) state that the winter outbreak is responsible for severe illness in three to five million individuals and the death of between 291 000 and 646 000 patients world-wide (Iuliano et al., 2017). The disease burden is high in patient groups such as those who are either pregnant, elderly, very young, or immunocompromised. The burden in healthy patients is otherwise relatively low. It is these groups that a vaccination program is targeted at. For example, a vaccination rate of 75 per cent is recommended for patients over 65 and those in other at-risk categories (Palache et al., 2014).

4.1.1 *Influenza vaccination*

The influenza virus was first isolated in 1933 and the first large-scale studies on patient vaccination were carried out in 1942 (Francis et al., 1945; Smith et al., 1933). This study demonstrated that vaccines were effective in protecting against subsequent infection using a bivalent vaccination against type A and type B influenza. Now the flu vaccination is reformulated every year to accommodate for the prevalent strains (Santos et al., 2015).

EFFICACY AND SAFETY   In reducing the risk of infection in the adult population, the efficacy in 18 to 49 year olds was 70 per cent (Monto et al., 2009). However, at-risk populations such as the elderly or immuno-compromised patients do not respond as well as healthy individuals and the vaccination is not as effective. When including older patient groups (18 to 65 year olds) the pooled efficacy of trivalent inactivated vaccines was shown to be 59 per cent in one meta-analysis (Osterholm et al., 2012). Although confirmation of reduced efficacy in elderly patients has been marred by lack of statistical confidence (Trucchi et al., 2015), one study indicates that an response rate of 42 per cent was seen in patients over 60 years of age (mean age = 69.5) using a live attenuated influenza vaccine (Villiers

et al., 2009). There are many studies on population wide response to vaccination and these represent just a few.

Vaccination is the primary method of effective protection against influenza infection by providing individual and herd immunity. However, occasional adverse events have been associated with vaccination (Stratton et al., 2011). The majority of non-serious adverse reactions are headaches, runny nose and other mild flu-like symptoms (Belshe et al., 2004; Musana et al., 2004). A serious adverse reaction, post-immunisation anaphylaxis can occur in 1.3 individuals per million with the majority of cases occurring less than four hours after vaccination (McNeil et al., 2016). Evidence also implicated vaccination with an oculo-respiratory syndrome (Skowronski et al., 2003). This was seen to in 2.9 per cent in vaccinated patients during the 2001-2002 vaccination season in Canada (Scheifele et al., 2003).

MEASURES OF EFFECTIVENESS   The effectiveness of influenza vaccination can be interpreted using serological measures of immunogenicity, commonly the production of antibodies to viral haemagglutinin (HA) (Hobson et al., 1972). Laboratory measurements of these antibodies are obtained by performing a haemagglutination inhibition (HAI) assay where patient serum samples are serially diluted until there is no more (if any) inhibition of red blood cell agglutination (Reber and Katz, 2013; Zacour et al., 2016). The induction of a four-fold increase in the titre from baseline can be associated with a two-fold decrease in the risk of infection (Benoit et al., 2015). This four-fold increase in titre between pre- and post-vaccination is known as seroconversion, one universal measure of vaccine response to vaccination (Reber and Katz, 2013; Talbot et al., 2012). The rate of seroconversion was estimated to be 70 per cent in patients across different strains (Seidman et al., 2012).

The relationship between age and response The production of HA antibodies and the resulting seroconversion in patients is driven by B cells and plasma cells within the blood and bone marrow. Evidence suggests the anti-HA antibody response peaks rapidly within one week of vaccination secreted by plasma cells. Another peak is seen later between two and three weeks and is attributed to IgG+ secreting memory B cells (Wrammert et al., 2008).

The response of CD4+ T cells may be also be coordinated with antibody secretion as a result of vaccination (Tan et al., 2017); An expansion of haemagglutinin tetramer-positive (tet+) CD4+ T cells is seen in patients which is seen to correlate with HAI measurements (Nayak et al., 2012). Therefore it may be important to understand the effects of vaccination upon gene expression in the HA-specific CD4+ T cell population.

4.1.2 *Predicting seroconversion*

The real-terms protective effect of vaccine-induced HAI titres has been predicted using post-vaccination measurements. A dilution passing 1 in 40 has said to confer a 50 per cent reduction in the risk of influenza infection in healthy patients and remains as a benchmark (Hobson et al., 1972). A more recent dose-response model takes this a step further to quantify the response for antibody titres measured after vaccination (Huang et al., 2017). However, the prerequisite is that patients have to be patients have to be vaccinated to predict whether they will be protected. A more ideal model would also be able to predict the patient response prior to vaccination itself in an effort to mitigate against unnecessary injections.

METHODS OF PREDICTING SEROCONVERSION   Linear regression models have been used to predict HAI response and seroprotection rates in patients. One highly predictive linear regression model showed that pre-vaccination CD4+ T cells were the best age-independent predictor of non-seroprotection (HAI titre below 1:40) (Jürchott et al., 2016). Jürchott et al. (2016) also show that, surprisingly, their influenza-specific (CD40L+) activated CD4+ T cells did not contribute towards prediction. Another study even demonstrated that a good mood, or at least the unmeasured factors associated with a patient's positivity, was the greatest predictor patient response in a regression model (Ayling et al., 2018). Non-clinical variables such as residence have also been studied which suggests that institutional residence affects seroconversion rates between strains (Seidman et al., 2012). An extensive differential expression analysis of PBMCs was carried out by Nakaya et al. (2011) which used a discriminant analysis algorithm to predict HAI response, demonstrating a high predictive accuracy (90 per cent) in independent trials. This is the largest validated predictor which uses a large variety of cells, however, a more specific cell type may work just as well for response categorisation.

AIMS   As part of a larger project, this research describes the use of a classification model of patient seroconversion response which uses the transciptomic profiles of serum tet+ CD4+ memory T cells prior to vaccination. The study aimed to examine the gene expression of immune cells using mRNA sequencing, to determine the key genetic differences between tet+ CD4+ memory T cells from seronegative and seropositive patient samples and return a gene panel for response classification.

4.2   MATERIALS AND METHODS

4.2.1   *Patients and samples*

Two influenza season cohorts of patients were used in this study. Samples were accessed through the Cambridge BioResource and consisted of two ages groups, young and old. Vaccination formula and key sample information can be found in Table 4.1. Both cohorts received the current season's influenza vaccination and blood samples were collected before, one week and six weeks after vaccination (denoted day zero, day seven and day 42). Influenza (A/H1N1/Cal09) tet+ memory (CD45RA$^-$) T cells were isolated using fluorescence-activated cell sorting (FACS). The experimental side of this research prior to sequencing was carried out by Danika Hill at the Babraham Institute, Cambridge, UK.

	E1 (2014/2015)	E2 (2014/2015)[†]
Seropositive	14	4
Seronegative	9	6
Age (years)	48.5 ± 20.5	60.4 ± 2.5

Influenza season 2014/2015 Trivalent Vaccination: A/California/7/2009 (H1N1)pdm09, A/Texas/50/2012 (H3N2)-like virus, B/Massachusetts/2/2012-like virus. [†], the age range for this group was 57–64

Table 4.1: Sample characteristics.

4.2.2   *Software and packages*

R (R Core Team, 2016) was used to perform the data cleaning, normalisation, statistical analyses and classification as well as graphics. Differential expression analysis was performed using package DESeq2 (Love et al., 2014). The random forest was performed using the implementation in randomForest package (Liaw and Wiener, 2002). Intercohort correction was carried out using ComBat (Johnson et al., 2007).

4.2.3   *Supervised random forest*

Random forests were generated over the top 1 000 DESeq-ranked list of genes. A supervised random forest is given a class list as well as training and test data to generate decision trees. Trees were initially seeded with a random sample of genes from the input data and the most important gene features across multiple forests were harvested. Each feature selection was tested with four-fold cross validation.

4.2.4 Unsupervised clustering

Unsupervised clustering of data generates distinct clusters of samples by attempting to determine distinct true class distributions using a serious of parameters. In this research, non-squared euclidean distances were calculated between sample and the clusters were computed using Ward's algorithm.

4.2.5 Sequencing and read count processing

Single-end mRNA sequencing was carried out, The expected library size was 100 base pairs. Genes with zero counts in all samples were removed. For classification and principal component analysis (PCA), read counts were normalised using regularised log2 transformation which accounts for library and sample sizes. For differential expression analysis, the non-normalised counts were used.

4.2.6 Classification

The output from the differential expression analysis was ranked by both absolute fold change and significance scores using equation 4.1.

$$score = \frac{|foldchange|}{significance} \tag{4.1}$$

The probability of a patient response using the NBC is calculated as the product of both prior and posterior probabilities 4.2. The prior probability is the probability that response category C_k occurs. The posterior probability is the probability that the genes' log RNA read counts $(x_1, x_2, ...x_j)$ in $|X|$ are observed in class C_k as is modelled using a univariate probability density function (PDF). In using a univariate distribution, the classifier assumes variable independence.

$$P(C_k|(x_1, x_2, ..., x_j)) \propto P((x_1, x_2, ..., x_j)|C_k) \cdot P(C_k) \tag{4.2}$$

The model is trained by taking the mean μ_{jk} and standard deviation σ_{jk} of the log read counts of genes x_{ijk} in samples of each class. To make a prediction, a new patient sample is a given vector of log read counts d, $(x_1, x_2, ...x_j)$. The likelihood that the read count x_j is from a given class C_k is then calculated as a function of the Gaussian distribution parametrised to μ_{jk} and standard deviation σ_{jk} (4.3).

$$\mathcal{L}(x_j|C_k) = \frac{1}{\sqrt{2\pi\sigma_{jk}^2}} \cdot e^{-\frac{(x_j - \mu_{jk})^2}{2\sigma_{jk}^2}} \tag{4.3}$$

A MAP decision is then made between the classes after calculating the product sum of each class' gene-expression likelihoods and the prior probabilities defined by 4.4.

$$P(C_k|d) \propto P(C_k) \cdot \prod_{j=1}^{|X|} \mathcal{L}(x_j|C_k)^{x_{ij}}$$
(4.4)

As a result, the classifier selects the highest probability / likelihood patient response as the category for vector of log read counts provided to it.

4.3 RESULTS

For training the model, the first cohort's (E1) pre-vaccination tet+ CD4+ memory T cells were examined. Although here the interest is in developing an age-independent predictor, age is likely to correlate with response.

4.3.1 *The relationship between age and response*

Measured by the fold-change in HAI titre before and after vaccination, the response was seen to decrease with age in E1 samples (Figure 4.1). The median fold-change increase in HAI measured from day zero to day 7 was 10 and 1.5 in young and old patients respectively. The difference between age groups was significant ($p <$ 0.01, Kruskal–Wallis test; normality rejected with Shapiro–Wilk test, $p < 1 \times 10^{-3}$). Moreover, there were 11 positive young responders compared with three old responders determined by a four-fold change in HAI titre after vaccination. This suggests a correlation of age with a drop in seroconversion rates.
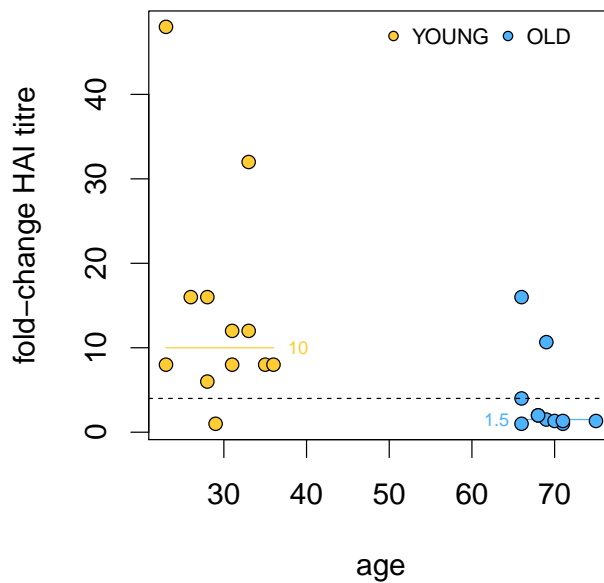
Figure 4.1: The dashed line shows the four-fold change threshold for categorising a positive response patient. Medians values of the age groups are shown.

4.3.2   *Principal component analysis using unfiltered genes.*

To explore the transcription profiles of tet+ CD4+ memory T cells, differential expression analysis was carried out on E1 samples. The list of gene features was cleaned by removing dropout genes. The primary aim of this was to compare T cells between seropositive and seronegative patients and so a standard pairwise comparison was carried out between groups without adjusting for age. No genes were determined to be statistically significant after adjusting for false discovery rate following a Benjamini-Hochberg procedure. A PCA using this list was performed to observe the variance, if any, caused by response category and age in tet+ cells prior to vaccination (Figure 4.2).
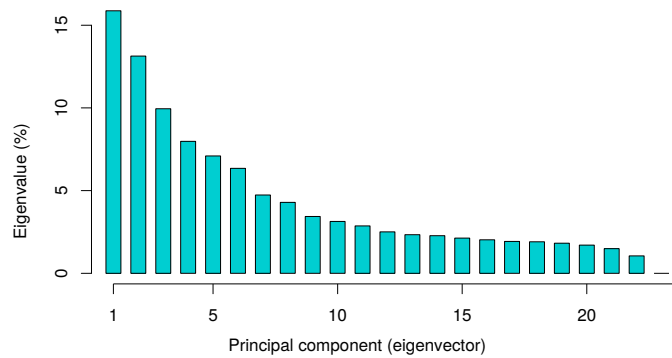
(a) Principal component variance.
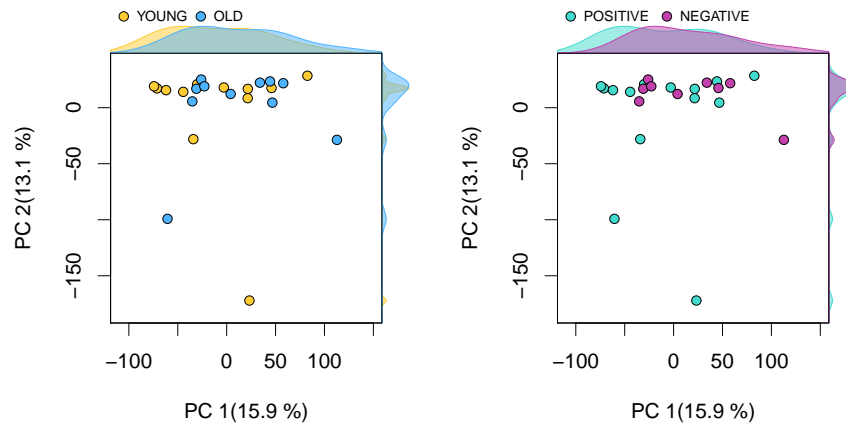
(b) Age categories (left) and response categories (right).

Figure 4.2: PCA of E1 (n = 23) day zero samples consisting of 11274 gene-features.

Neither age nor response explain the two largest components of variance (making up 29 per cent of the variance) in gene expression within E1 samples (Figure 4.2). To find a principal component (PC) that contains the response category the most, the median normalised PCA scores of each class were subtracted from each other within each

component. The largest difference between age groups was PC 1 followed by PC 3 and the largest difference for response categories was seen in PC 1 followed by PC 23 (Figure 4.3).
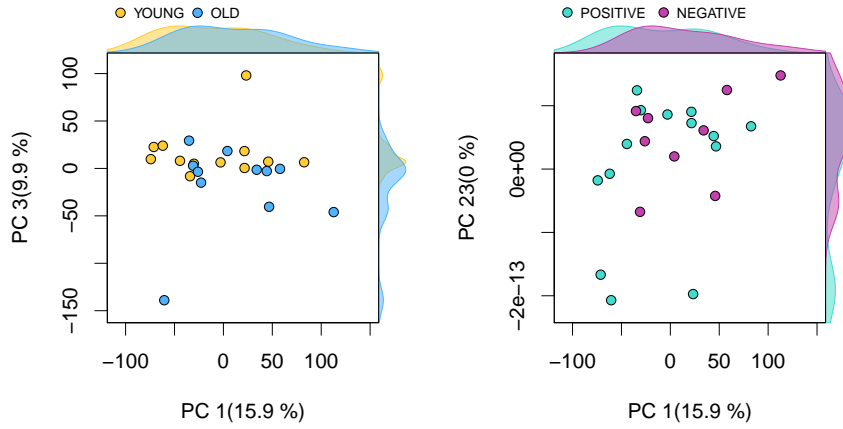
Figure 4.3: PCA components with maximal distance between age and response categories in cohort E1.

4.3.3 *Dimensionality reduction of genes in training an NBC*

To narrow down the genes and reveal any potential underlying transcriptional features of memory T cells which regulate HAI response, a classification model was used, applying dimensionality reduction methods.
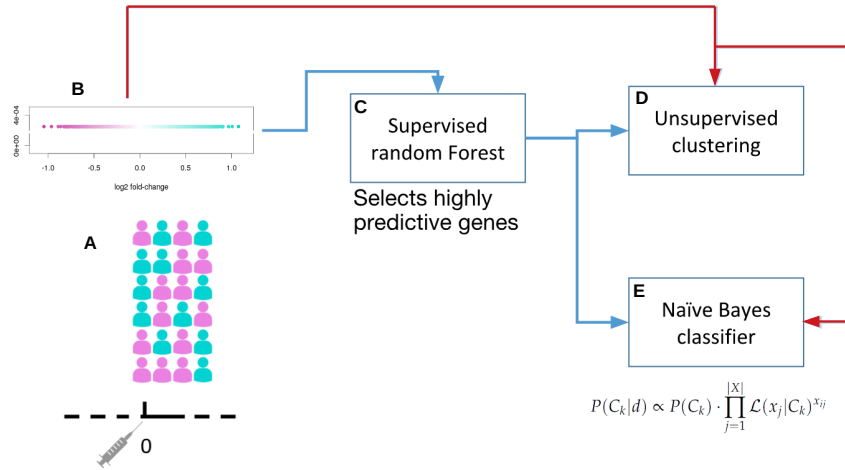
Figure 4.4: A, Patient data was taken before vaccination; B, differential expression results reveal no significant differences in gene expression; C, supervised random forests discern between serological classes and re-order the ranked list of genes; D, unsupervised hierarchical clustering using the output genes gives an unbiased view of the gene selection; E, the naïve Bayes classifier model is formed with an optimal and reduced-dimension subset of genes.
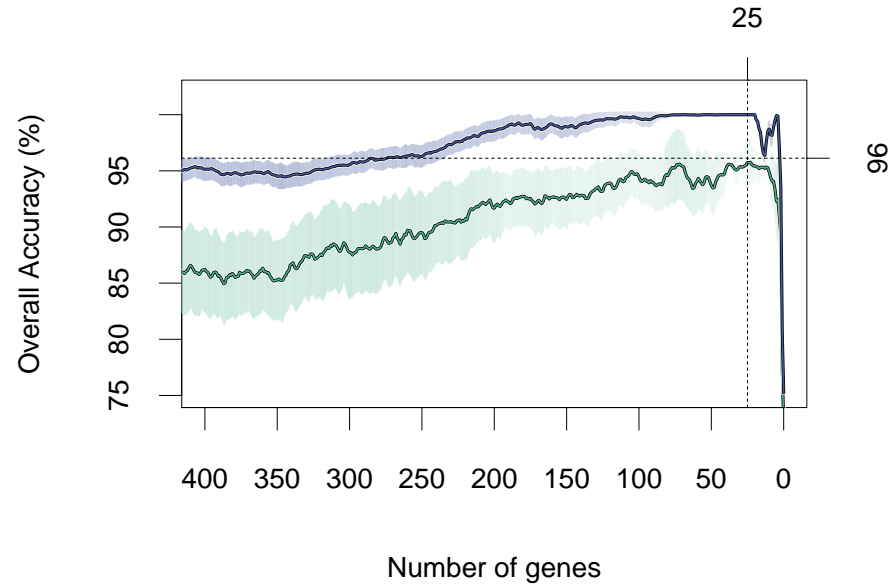
(a) Reducing the gene dimensions of the differential expression analysis.
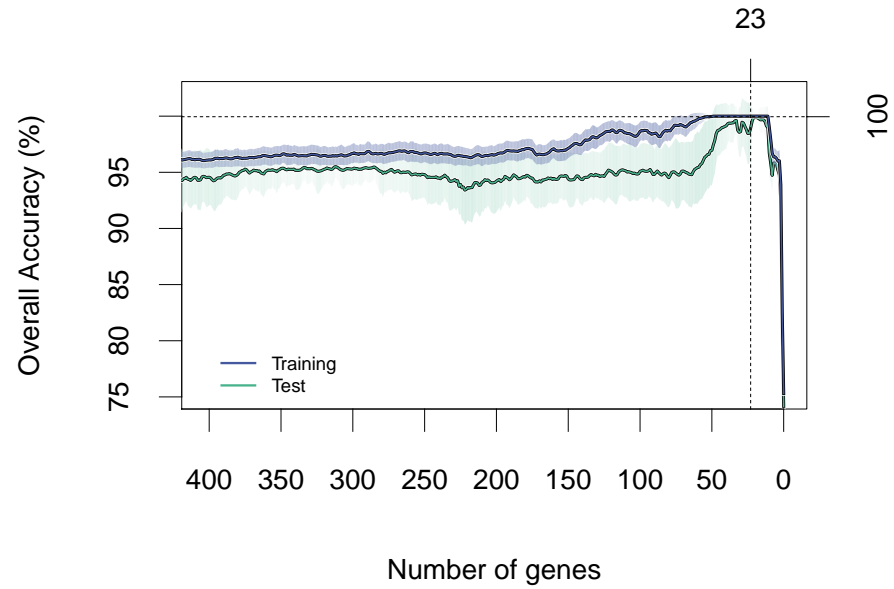
(b) Reducing the gene dimensions of the differential expression analysis after feature selection with random forests.

Figure 4.5: Plots show the means of four-fold cross validation over 100 random samples. The shade and area show the standard deviation across NBCs.

Gene-rank positions were obtained by sorting the differential expression analysis by fold-change and unadjusted significance between E1 response categories. We assumed that this ranked the genes by their importance as predictors. For each sequence of genes, descending through the highest ranking differential expression results, one hundred Gaussian NBCs were trained and tested in four-fold cross

validation over randomly selected partitions from E1 samples. The genes list was cut off at a maximum of 1 000 genes as a larger set of features was not expected to increase accuracy. Using this preliminarily ranked list of differentially expressed genes following the red work flow in Figure 4.4, the NBC achieved an accuracy of 96.12 per cent ± 2.53 per cent in test partitions by selecting only the top 25 genes (Figure 4.5a).

In parallel, feature selection was performed over these top-ranking 1 000 genes using a supervised random forest algorithm (blue trajectory in Figure 4.4). This re-ordered the genes by the frequency of appearance within the ten highest importance genes from each of 1 000 forests. One hundred NBCs were also trained over each descending sequence of re-ordered genes. The random forest optimisation resulted in an accuracy of 99 per cent within 11 genes and the optimal subset of 23 gene features resulted in an overall accuracy of 99.95 per cent ± 0.5 per cent providing a highly-predictive subset of features between E1 samples (Figure 4.5b).

Unsupervised hierarchical clustering of the data using only the 23 predictive genes demonstrates that these genes effectively separate between the seropositive and seronegative classes (Figure 4.6). Furthermore, a follow-up PCA using only these 23 genes demonstrates that the largest source of variance, component one, is now undoubtedly explained by the response category (Figure 4.7). Patient age and response category appear to be associated as a large overlap is seen between the positive responders and young patients. However, the spread of old patients was not fully accounted for in component one which suggests that the gene subset is at least partially age-independent (Figure 4.7).
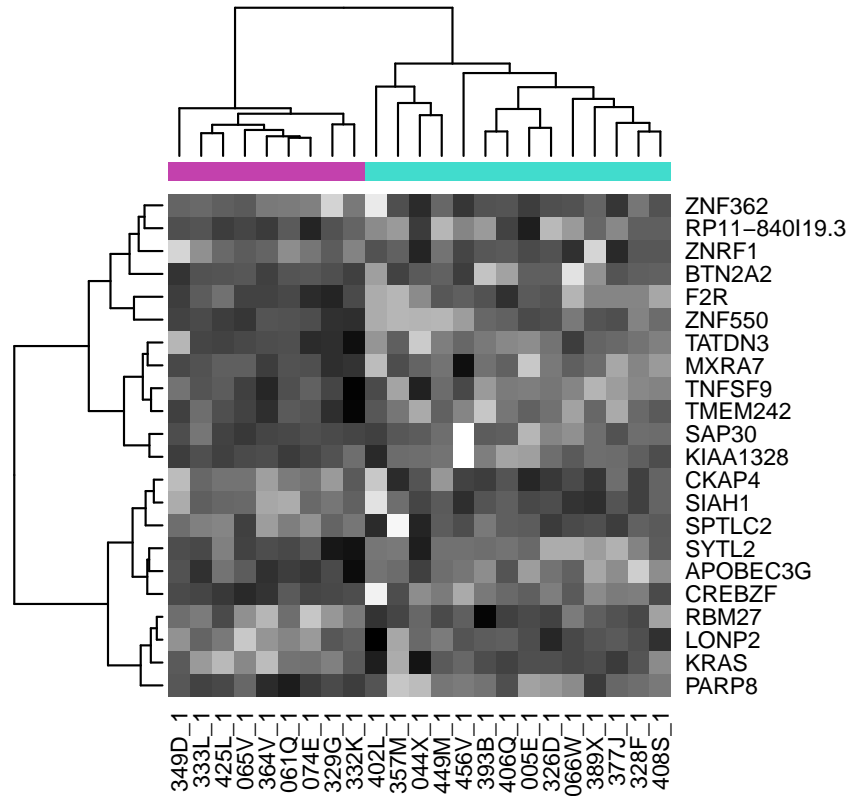
Figure 4.6: Magenta coloured dendrogram labels are seronegative and turquoise labels are seropositive.
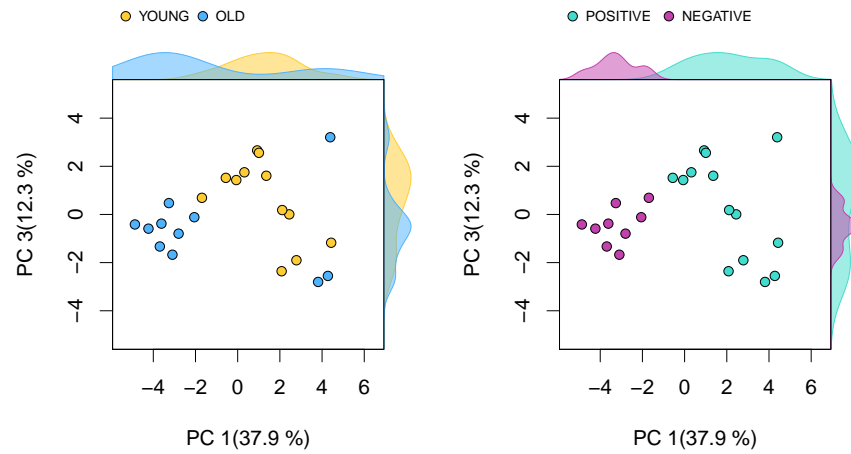
Figure 4.7: PCA of E1 day zero samples using the optimal predictive gene subset.

4.3.4 *Predicting patient response*

The goal was then to test the classifier by predicting an independent cohort, E2. These samples were collected from patients receiving the same vaccine formula in the same year. However, they were sequenced at a later date. A large inter-cohort effect was observed and cohort corrections were carried out to remove this effect seen in PCA (Figure 4.8a). The NBC was then trained with the corrected E1 data and subsequently used to predict the seroconversion responses of E2 patients. This was carried out blind to avoid the possibility of bias. Despite the solid prediction of E1 using cross-validation methods, the overall accuracy of E2 predictions was 30 per cent with 90 per cent of the predictions being seropositive. The PCA (Figure 4.8b) confirms this over-estimation of seroconversion as the majority of E2 are clustered with E1 seropositive samples. Examination over why the classifier performs as it does on the new data shows that most predictive genes in E2 are expressed in a manner complementary to E1 (Figure 4.9).

Many of the genes selected by the random forest (RF) and NBC optimisation process have been highlighted in previous studies on viral response. F2R plays and anti-viral role in innate immunity and F2R (PAR-1)$^{-/-}$ mice show increased viral load and inflammation after influenza infection (Antoniak et al., 2013). However, the knockout of F2R has also increased the survival of mice after influenza infection as it might reduce epithelial barrier integrity (Khoufache et al., 2012). MXRA7 was found to be expressed at a lower level individuals resilient to Enterotoxigenic Escherichia coli infection (Yang et al., 2016). SAP30 is seen to play a role in Rift Valley fever virus as its activity may cause increased virulence (Terasaki et al., 2016). Upregulation of SYTL2 was seen in pigs infected with a strain of African swine fever viruse (Jaing et al., 2017). It is reportedly associated with the immunological synapse. CREBZF has been implicated in a number of anti-viral responses (Zhang et al., 2012). TATDN3 has previously been seen as significant between HIV and healthy patients (Wu et al., 2015). SPTLC2 is responsive in rhinovirus infection amongst others (Çalışkan et al., 2015). One interesting gene is LONP2. The expression of LONP2 is increased in aged antibody secreting cells in mice (Kannan et al., 2016). The PARP8 gene remains part of a larger family of proteins which regulate virus response (Kuny and Sullivan, 2016). CKAP4 was shown to be upregulated in chicken embryos in response to a strain of influenza (Li et al., 2017). A cytodine deaminase, APOBEC3G, a protein coding gene, is an inhibitor of retrovirus replication and retrotransposon mobility. This gene correlated with seroresponse in a previous analysis but was only highly expressed in natural killer cells and not discussed (Nakaya et al., 2011). Stav-

rou and Ross (2015) review APOBEC3 proteins and point out that APOBEC3G may be a dominant anti-viral protein in CD4+ T cells (Gillick et al., 2012).
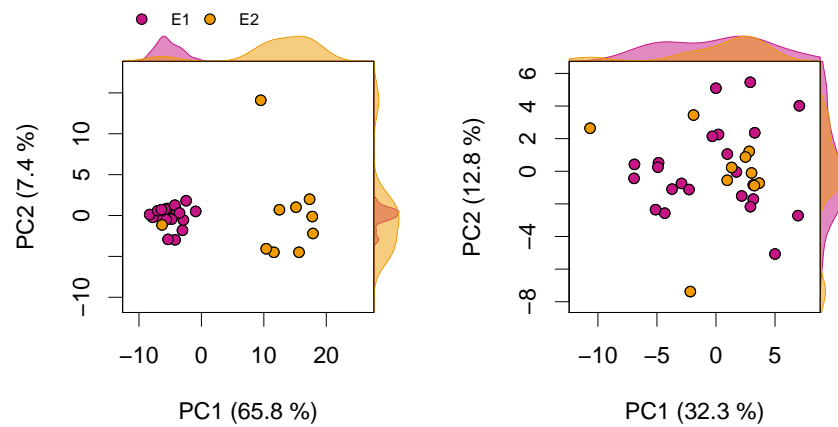
(a) PCAs of E1 and E2 using the 23 predictor genes before (left) and after (right) cohort correction
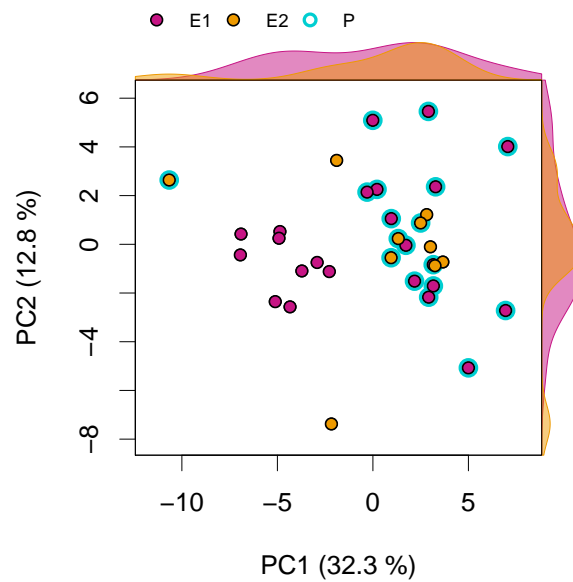
(b) Seroconversion of E1 and E2 samples across the component explained by the predictive genes.

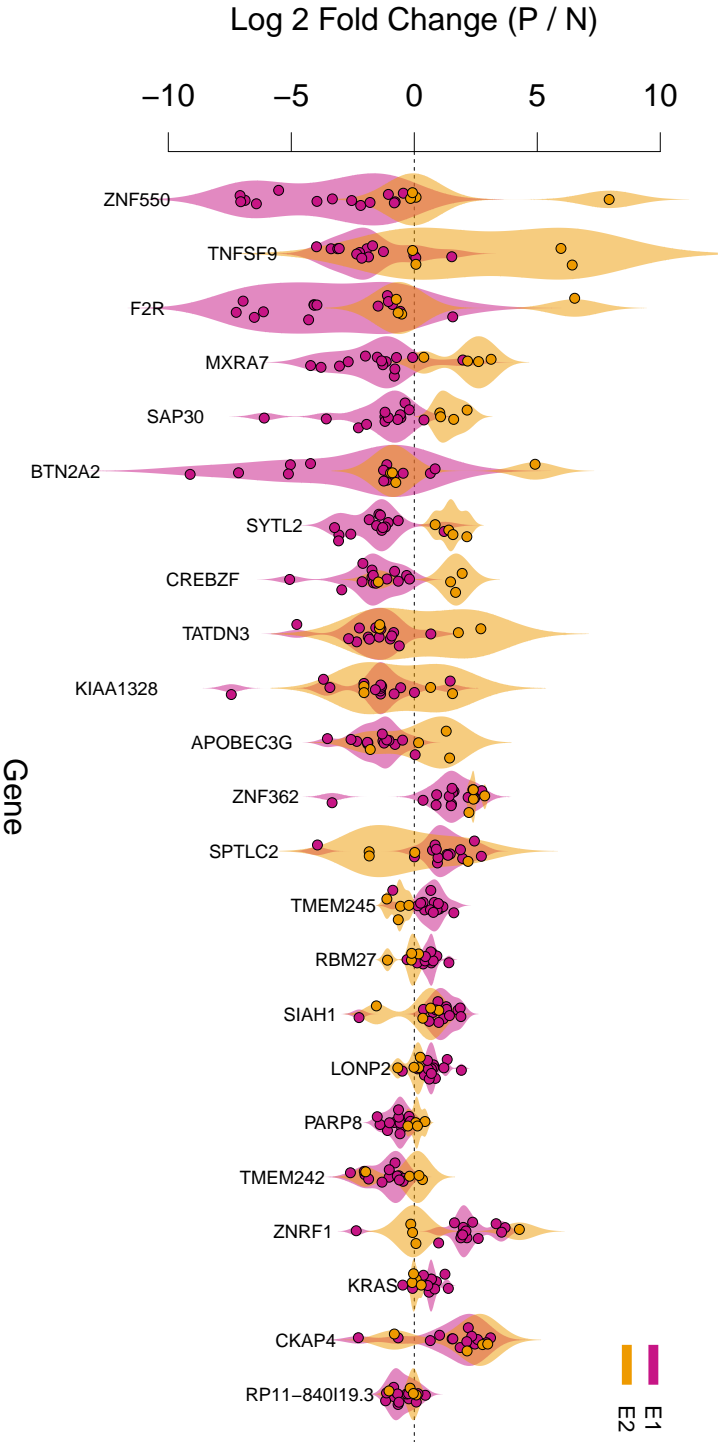Figure 4.8: True-positive seroresponders are highlighted by P.

Figure 4.9: Regularised log2 RNA-seq counts from seropositive patients were subtracted from the geometric mean of seronegative patients in each cohort.

4.4 DISCUSSION

Understanding the mechanisms behind the protective effects of vaccination is of major importance in the prevention of seasonal epidemics and occasional pandemics. Previous studies have identified T cell expansion as a major correlate of serological response in patients and have outlined predictive sets of genes between a large selection of PBMCs (Nakaya et al., 2011; Nayak et al., 2012; Tan et al., 2017).

One subset of cells not specifically explored yet were tet+ CD4+ memory T cells. An attempt was made to reveal a genetic determinant of response to vaccination in circulating tet+ CD4+ memory T cells using differential expression analysis. However, only non-significant results were seen in transcriptional differences between seropositive and seronegative patient's cells. One possible reason for non-significance with this study could be the low power due to a small sample are limitation the low power of this experiment was limited due to sample sizes.

4.4.1 *Feature selection and Classification*

This research shows a novel work flow that despite the lack of significance in differentially expressed genes, selects features from RNA-seq data using an RF and provides a non-arbitrary cut-off for biological analysis using the NBC. This demonstrates one approach to RNA-seq dimensionality reduction and classification with potential use in predicting patient response to medication.

The use of the Gaussian NBC is justified given that RNA-seq data are transformed correctly. Classifying RNA-seq data was shown to be reasonably effective using classifiers such as RF, SVM or variants of discriminant analysis upon data transformed by variance stabilisation or regularised log transformations (Zararsız et al., 2017). However, authors also describe that an alternative method to deal with inherent over-dispersion of in mRNA counts data is to use discrete counts based classifiers. Additionally a multivariate distribution NBC can also be used for linear-discrimination which was shown to be more effective in a number of datasets (Gama, 2000). Another improvement could be made by using different or adjusted measures of effectiveness as normal serological endpoints overestimate the effectiveness in inactivated vaccines (Petrie et al., 2011). Furthermore, the change in HAI resulting from vaccination is altered by the baseline abundance of HAI. Adjustments for this effect may normalise the patient responses and improve the classifier result (Beyer et al., 2004). It would be interesting to use the mNBC classifier from chapter two for this purpose as it assumes a multinomial, hence discretised, distribution. The use of this approach on publicly available data is required to validate the current approach further.

CHALLENGES ACROSS COHORTS   Within the first cohort, the classifier accuracy was 100 per cent with a panel of 23 genes. However, the prediction of second cohort was much less than 30 per cent which could suggest a non-random result; the model actively mis-classifies samples. Possible reasons for issues faced within this research were the low number of samples, a large variance between cohorts and individual samples as well as a number of uncontrolled biases in the training data, specifically a young positive responder bias. We also saw uneven distributions of age and response within and between cohorts. Sample size could be one of the most important factors to successful classification models using RNA-seq data (Zararsız et al., 2017) which is exaggerated by high variance in human sequencing data. To address this, larger studies could be carried out to better understand and predict the role of pre-vaccination tet+ CD4+ memory T cells. Figure 4.9 illustrates how the model actively mis-classifies patient responses. It shows that the fold-change of selected gene mRNA levels between positive and negative patient's tends to be largely opposite of the other cohort. This results in sample's gene expression being incorrectly assigned within the wrong class distribution by the predictive model.

It was initially assumed that an age-independent predictor of response could be found. However, it might be seen that age is a strong component of the response and a large intersection is seen between seropositive patients and young patients (Figures 4.7 & 4.1). In fact, only one young patient was seronegative in our training data which resulted in a particularly underpowered analysis for this sub-population. Experimental differences could may also be a reason that the classifier struggled to reproduce the precise classification of cohort E1 across cohort E2 samples. For example, the use of a different vaccine between the two cohort years may have contributed to the inter–cohort effects seen here. Furthermore, as large inter-cohort differences were observed, cohort correction using ComBat is advised for for small batches but may impose certain unwanted transformations of data with unbalanced sample groups (Nygaard et al., 2015). Additionally, it may be there is simply no real tangible difference between responsive and non-responsive patients using this cell type in a predictive model of vaccination.

4.4.2 *Gene selection and prediction of vaccination*

The gene selection made through classification is heavily associated with regulation of host-viral interactions. Interestingly, the trend in the training data E1 suggests that the majority of the anti-viral genes were expressed more in seronegative rather than seropositive patients. Unfortunately the results poor significance after multiple-test corrections gives little weight to these observations. As discussed age may

have biased the gene selection. This is potentially seen with LONP2 as its higher expression in positive cells may mimic its increased expression in aged cells as seen in mice (Kannan et al., 2016). Furthermore, the expression of APOBEC3G, one of the top predictors of seroconversion, was significantly higher in old patients compared with young patients at day zero when examined with a differential expression analysis (not shown).

In light of my analysis, perhaps pre-existing circulating tet+ CD4+ memory T cells convey little predictive power for estimating serological response to vaccination. Previous regression modelling has shown that influenza-specific T cells were not important compared with the total naive T cell population for prediction (Jürchott et al., 2016). Additional challenges faced in applying this model is that previous exposure to influenza may play a role in classifying response categories as the fold-change of antibody titres between pre-vaccination and post-vaccination are less pronounced in those who have high HAI titres at the time of injection (Francis et al., 1945; Seidman et al., 2012). Furthermore, patients with zero detectable antibody prior to vaccination were supposedly less susceptible to infection (Hobson et al., 1972). This poses an issue when matching genotype to phenotype when determining if a patient responds to vaccination.

In terms of translating pre-vaccination genetic variables to actual protection in real terms, models predicting HAI titre and the protective effect from a given titre value could be combined. One such future model could combine modules of pre-vaccination predictions of HAI and another with a protective dose response for HAI such as the model Huang et al. (2017). This could be a systems pharmacology model as none to my knowledge exist for this purpose.

4.4.3   *Concluding remarks*

I have successfully outlined an approach to classify vaccination response categories based on RNA-seq data. This was applied to predicting influenza vaccination response in patients but fell short when validated by independent test data. This might have been caused by a limited sample size combined with a great inter-patient variability. Successful prediction of vaccination responses can be important for screening patients in high-risk groups as well those who may experience more serious adverse reactions. However, the magnitude of most adverse reactions in influenza vaccination is low. Therefore it this context it could also be applied for resource-saving measures. For example, predictors of vaccine efficacy could prove useful in that herd immunity could be achieved by targeting only high-responders within a population. By evaluating an individual's priority to be vaccinated, a proactive vaccine vigilance program may then recommend the most responsive patients for vaccination to achieve the fraction of

population immunity required for population-level immunity. More broadly, the potential to predict the patient response rate for other more harmful drugs would be highly desirable and classification models such as this could be essential in advent of a more personalised genomic medical system.

One of the aims of systems pharmacology is to predict the response to therapeutics in patients and inform decisions on dose regimens and formulation. by sampling a patient's molecular profile we can inform models for predicting outcomes and communicate individual and population-wide treatment and prevention strategies for diseases. Additionally, systems pharmacology models are interested in understanding the underlying disease mechanisms. To that end I argue that we can also derive the critical molecular mechanisms responsible for conferring vaccination efficacy by applying feature reduction and classification to feed into networks and explore disease mechanisms.

DISCUSSION, CONCLUSIONS AND OUTLOOK

5.1 INTRODUCTION

The research presented in this thesis set out to explore the systems pharmacology modelling and whether its use can help us understanding disease mechanisms and makes decisions between druggable targets.

We uncovered a retrospective spread of computational pharmacology models, exemplified a case of systems pharmacology models in target selection by modelling disease mechanisms, and examined a classification model for predicting response to therapeutics was examined. Two crucial questions were addressed by the study and help to assess whether systems pharmacology models are able to reduce drug attrition:

1. Can systems pharmacology models explore disease mechanisms?

2. Can systems pharmacology models predict drug targets?

5.2 MAIN FINDINGS

The discussions situated at the end of each chapter hold a more specific and topical debate on the results, limitations of methodology and value of findings. However, in the true spirit of systems pharmacology, the individual chapters present pieces of a combined puzzle greater than merely the sum of their parts. Together, the findings show that systems pharmacology models can be and are useful in exploring disease mechanisms.

CLINICAL FOCUS DIRECTS AN EVOLVING MODELLING LANDSCAPE: The findings in chapter two suggest that firstly, the diseases which are explored by the models up until now have been relevant to today's picture of medical needs with only a few exceptions. This observation is important as a clear dialogue between clinical scientists and modellers is key in understanding which disease mechanisms are most important for exploring for impacting drug attrition.

Secondly, the rising knowledge of complex biological systems is mirrored by a rise in model complexity. The success in preventing drug failure that the previous generation of predictive models may have played a part in could be bolstered by the newer generation of mechanistic systems pharmacology models.

SYSTEMS PHARMACOLOGY IS BUILT BY UNDERSTANDING THE DISEASE:
Beyond demonstrating the re-usability of methods, chapter four shows that disease mechanisms are derived from mechanistic understanding. Classification models that predict patient response highlight key players which drive efficacy in drugs. Variation in response is often modulated multi-factorially and require systems approaches to understand.

SYSTEMS PHARMACOLOGY MODELS RESULT IN EMERGENT HYPOTHESES:
The IL-6 model of arthritis in chapter three shows that systems pharmacology models built to model biological mechanisms may reveal overlooked properties of the disease. Offloading a soluble receptor in high concentrations by sequestering its cognate ligand was unforeseen and not documented in clinical trials. Moreover, the work in developing this model aligns both chapters two and four in realising that the conjunction of clinical drive and amassed mechanistic knowledge are necessary motivators of constructing systems pharmacology models to understand disease mechanisms.

Furthermore, we find that systems pharmacology models can predict response outcomes to aid optimal target selection, a requirement for developing safer more efficacious therapy.

SYSTEMS PHARMACOLOGY MAKES PRE-CLINICAL DECISIONS: In chapter three, the QSP model of TCZ and SRK therapy in RA is the comparison of two targets within a single system. The dichotomy of receptor versus ligand is a perfect example of how dynamic systems pharmacology models can be used in target pre-clinical selection. We show that the comparison between targets could take place before drugs are pushed into large-scale clinical trials without fitting to drug-specific data. Drug viability can then be inferred for making judgment calls in continuing development of a drug over a competitor or gold-standard of therapy. Furthermore, the model shows that systems pharmacology model-derived hypotheses on target-related toxicity may suggest alternative biomarkers or measures in clinical studies. These evaluations would be important in making go or no-go decisions and reduce late-stage attrition.

NETWORKS CAN ASSESS MULTIPLE TARGETS: Predicting patient response using a classification model in chapter four may be used to identify strong drivers of therapeutic response. This shows one case for determining the multiple components of complex networks that drive drug efficacy which could lead to developing systems pharmacology models and therapeutic targets.

5.3 THEORETICAL IMPLICATIONS.

The work undertaken here gives an organic view of the evolution into systems pharmacology models however more work needs to be carried out to reach the substantial goals of reduced attrition in drug development.

In this study, a desired outcome was to reveal the merits of systems pharmacology models in making clinical decisions but the current assessment using clinical trials data was too course-grained to derive any pre-clinical context. Although the use of pharmacometrics in regulatory decisions has shown remarkable success (Bhattaram et al., 2007; Bhattaram et al., 2005; Lee, 2014), the effectiveness of systems pharmacology models to aid decisions in pre-clinical phases still needs to be properly assessed. It may be difficult to achieve this as both published and unpublished accounts of modelling need to be considered.

The hypothesis-driven construction of systems pharmacology models can stem from connecting models of patient response to the underlying variables taken from *in vitro* studies such as RNA sequencing. We generated a work-flow for one half of the puzzle; evaluating the power of underlying genes in which predict patient response from the levels of their mRNA in circulating T cells. This is one approach of developing networks which can be transferred into systems pharmacology models (Thiel et al., 2016). Inconsistencies between the data sets inaccurate predictions do not reduce the value of these methods and call for more data.

The model constructed for IL-6-mediated RA may improve our understanding of intricacies in targeting IL-6 in arthritis and other conditions. This work built upon a QSP model and was further validated by fitting to and predicting a broad array of data. Previous models in this disease and drug topic have been population PK models and have used estimations for nonlinear elimination. The model in chapter three however demonstrates that an understanding of the disease and a mechanistic interpretation using experimentally derived measures can simulate the TMDD seen in clinical data. The model examined the general use of monoclonal antibodies and proposed challenges that may come from high affinity drugs. One implication of this model is that we may now have a testable hypothesis for the safety concerns of MACE in SRK therapy.

5.4 PUSHING FOR CHANGE.

The paradigm of successful drug discovery that relies upon serendipity has long since past and researchers strive to generate new drugs for existing diseases using a wide range of methods. The utility of QSP and systems pharmacology as a framework will likely be a keystone

in the multi-disciplinary endeavour that is drug discovery and development, helping to reduce attrition. The thoughts contained here are not without an echo in the field (Androulakis, 2016) and many computational pharmacologists are looking forward to fully using QSP throughout the drug discovery pipeline.

The landscape generated in chapter two suggests that the majority of software used are proprietary and operated through industrial efforts. This may hinder the uptake of avid younger pharmacologists in pursuit of interdisciplinary science. A push to overcome this major hurdle is needed for embracing a systems pharmacology framework by expanding QSP training in institutions of education as discussed in the white paper (Sorger et al., 2011). Moreover, as the field becomes a bonefide player in the wider scientific community, a more solid definition will form, increasing accessibility of QSP.

Additionally, following the re-purposing of the IL-6 model in chapter three, it is apparent that model standards and source code are important aspects for systems pharmacologists to be able to share their models and knowledge. A push for change would be to increase the uptake of model exchange formats such as PharmML (Swat et al., 2015) in pharmacometrics or SBML in systems biology as well increase use of standard guidelines for annotation and model documentation like MIRIAM (Le Novère et al., 2005). Interestingly, we did not see the terms SBML or PharmML once in the systems pharmacology modelling landscape I outlined in chapter 2 which is further evidence that although systems modelling communities have coordinated to build these standards, there is no evidence that they are adopted within systems pharmacology.

5.5 EXPANDING THE RESEARCH.

One future project for the landscape of systems pharmacology models could be to generate a modelling literature library where researchers can search for models using a pre-classified index of PubMed. Experience in curation within BioModels and filtering through literature to find models has shown that it can be tedious for researchers to find mathematical models. This could be linked to model repositories such as The Drug Disease Model Resources (DDMoRe) repository (Harnisch et al., 2013) or BioModels and help to further the practice of re-usability in modelling research. This could also serve to improve model accuracy through public consensus of class labelling and be implemented as a dynamic picture of systems pharmacology and biology modelling landscapes. Additionally, this could provide a platform for crowd-sourced contributions towards modelling ontologies such as MAMO and improve annotation accuracy.

Future research would be to validate the findings of the IL-6 model in chapter three. Quite simply, data from clinical studies measuring

serum sIL-6Rα or IL-6 may entirely refute the hypothesis for safety concerns which, in itself would be one outcome of the model. This could lead to extending the model structure to incorporate known aspects of the drug toxicity. Furthermore, the current trial and use of anti–IL-6 therapy including SRK, in indications such as major depressive disorder could provide a new opportunity for model use and refinement.

Finally, although very much in its infancy, the influenza response classifier using RNA-seq data in circulating flu-specific memory T cells should be further validated in independent cohorts. Extension of the model into a dynamical network could generate an interesting systems pharmacology model in vaccination.

5.6 CONCLUSION

Systems pharmacology models are vital for interpreting the rising complexities in our knowledge of biology. Moreover, their use in turning these interpretations into predictions for developing medicines is seen by their application in probing druggable targets.

Their success in the clinical setting is beginning to show and yet their application in the preclinical setting is where they will most likely shine; it may still be too early to tell if systems pharmacology models can really impact attrition in drug discovery.

However, this research addressed two key questions with findings suggesting that in fully modelling the disease mechanisms, systems pharmacology models may predict the druggable targets for drug development and indeed, reduce the chances of attrition.

BIBLIOGRAPHY

Agoram, Balaji M., Steven W. Martin and Piet H. van der Graaf (2007). 'The role of mechanism-based pharmacokinetic–pharmacodynamic (PK–PD) modelling in translational research of biologics'. In: *Drug Discovery Today* 12.23-24, pp. 1018–1024. DOI: 10.1016/j.drudis.2007.10.002.

Ajmera, I, M Swat, C Laibe, N Le Novère and V Chelliah (2013). 'The impact of mathematical modeling on the understanding of diabetes and related complications'. In: *CPT: Pharmacometrics & Systems Pharmacology* 2.7, e54. DOI: 10.1038/psp.2013.30.

Akaike, H. (1974). 'A new look at the statistical model identification'. In: *IEEE Transactions on Automatic Control* 19.6, pp. 716–723. DOI: 10.1109/tac.1974.1100705.

Aletaha, Daniel, Clifton O Bingham, Yoshiya Tanaka, Prasheen Agarwal, Regina Kurrasch, Paul P Tak and Sharon Popik (2017). 'Efficacy and safety of sirukumab in patients with active rheumatoid arthritis refractory to anti-TNF therapy (SIRROUND-T): a randomised, double-blind, placebo-controlled, parallel-group, multinational, phase 3 study'. In: *The Lancet* 389.10075, pp. 1206–1217. DOI: 10.1016/s0140-6736(17)30401-4.

Allerheiligen, S R B (2010). 'Next-Generation Model-Based Drug Discovery and Development: Quantitative and Systems Pharmacology'. In: *Clin Pharmacol Ther* 88.1, pp. 135–137. DOI: 10.1038/clpt.2010.81.

Alm, E (2003). 'Biological networks'. In: *Current Opinion in Structural Biology* 13.2, pp. 193–202. DOI: 10.1016/s0959-440x(03)00031-9.

Alsmadi, Izzat and Ikdam Alhami (2015). 'Clustering and classification of email contents'. In: *Journal of King Saud University - Computer and Information Sciences* 27.1, pp. 46–57. DOI: 10.1016/j.jksuci.2014.03.014.

Anderson, Amy E, Arthur G Pratt, Mamdouh A K Sedhom, John Paul Doran, Christine Routledge, Ben Hargreaves, Philip M Brown, Kim-Anh Lê Cao, John D Isaacs and Ranjeny Thomas (2015). 'IL-6-driven STAT signalling in circulating CD4+ lymphocytes is a marker for early anticitrullinated peptide antibody-negative rheumatoid arthritis'. In: *Annals of the Rheumatic Diseases* 75.2, pp. 466–473. DOI: 10.1136/annrheumdis-2014-205850.

Androulakis, Ioannis P. (2016). 'Quantitative Systems Pharmacology: A Framework for Context'. In: *Current Pharmacology Reports* 2.3, pp. 152–160. DOI: 10.1007/s40495-016-0058-x.

Antoniak, Silvio et al. (2013). 'PAR-1 contributes to the innate immune response during viral infection'. In: *Journal of Clinical Investigation* 123.3, pp. 1310–1322. DOI: `10.1172/jci66125`.

Ara, T. et al. (2013). 'Critical Role of STAT3 in IL-6-Mediated Drug Resistance in Human Neuroblastoma'. In: *Cancer Research* 73.13, pp. 3852–3864. DOI: `10.1158/0008-5472.can-12-2353`.

Armstrong, J.K., R.B. Wenby, H.J. Meiselman and T.C. Fisher (2004). 'The Hydrodynamic Radii of Macromolecules and Their Effect on Red Blood Cell Aggregation'. In: *Biophysical Journal* 87.6, pp. 4259–4270. DOI: `10.1529/biophysj.104.047746`.

Ayling, Kieran, Lucy Fairclough, Paddy Tighe, Ian Todd, Vanessa Halliday, Jon Garibaldi, Simon Royal, Aljali Hamed, Heather Buchanan and Kavita Vedhara (2018). 'Positive mood on the day of influenza vaccination predicts vaccine effectiveness: A prospective observational cohort study'. In: *Brain, Behavior, and Immunity* 67, pp. 314–323. DOI: `10.1016/j.bbi.2017.09.008`.

Bakan, Ahmet, Neysa Nevins, Ami S. Lakdawala and Ivet Bahar (2012). 'Druggability Assessment of Allosteric Proteins by Dynamics Simulations in the Presence of Probe Molecules'. In: *Journal of Chemical Theory and Computation* 8.7, pp. 2435–2447. DOI: `10.1021/ct300117j`.

Baker, M., S. Denman-Johnson, B. S. Brook, I. Gaywood and M. R. Owen (2012). 'Mathematical modelling of cytokine-mediated inflammation in rheumatoid arthritis'. In: *Mathematical Medicine and Biology* 30.4, pp. 311–337. DOI: `10.1093/imammb/dqs026`.

Baker, Simon, Douwe Kiela and Anna Korhonen (2016). 'Robust Text Classification for Sparsely Labelled Data Using Multi-level Embeddings'. In: *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*. Osaka, Japan: The COLING 2016 Organizing Committee, pp. 2333–2343.

Balijepalli, Chandra and Olivia Oppong (2014). 'Measuring vulnerability of road network considering the extent of serviceability of critical road links in urban areas'. In: *Journal of Transport Geography* 39, pp. 145–155. DOI: `10.1016/j.jtrangeo.2014.06.025`.

Ban, T. A. (2006). 'The role of serendipity in drug discovery'. In: *Dialogues Clin Neurosci* 8.3, pp. 335–344.

Bao, M., O.A. Harari, A.M. Jahreis, J.F. Schmidt and X. Zhang (2012). *Subcutaneously administered anti-il-6 receptor antibody*. US Patent App. 13/290,366. URL: `https://www.google.com/patents/US20120301460`.

Bastian, Hilda, Paul Glasziou and Iain Chalmers (2010). 'Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up?' In: *PLoS Medicine* 7.9, e1000326. DOI: `10.1371/journal.pmed.1000326`.

Baumann, H., H. Isseroff, J. J. Latimer and G. P. Jahreis (1988). 'Phorbol ester modulates interleukin 6- and interleukin 1-regulated ex-

pression of acute phase plasma proteins in hepatoma cells'. In: *J. Biol. Chem.* 263.33, pp. 17390–17396.

Beal, S., L.B. Sheiner, A. Boeckmann and R.J Bauer (2009). *NONMEM User's Guides. (1989-2009)*. Icon Development Solutions, Ellicott City, MD, USA.

Bekhuis, Tanja (2006). 'Conceptual biology, hypothesis discovery, and text mining: Swanson's legacy'. In: *Biomedical Digital Libraries* 3.1. DOI: 10.1186/1742-5581-3-2.

Belshe, Robert B, Kristin L Nichol, Steven B Black, Henry Shinefield, Julie Cordova, Robert Walker, Colin Hessel, Iksung Cho and Paul M Mendelman (2004). 'Safety, efficacy, and effectiveness of live, attenuated, cold-adapted influenza vaccine in an indicated population aged 5–49 years'. In: *Clinical infectious diseases* 39.7, pp. 920–927.

Benam, Kambez H. et al. (2015). 'Engineered In Vitro Disease Models'. In: *Annual Review of Pathology: Mechanisms of Disease* 10.1, pp. 195–262. DOI: 10.1146/annurev-pathol-012414-040418.

Benoit, A. et al. (2015). 'Hemagglutination Inhibition Antibody Titers as a Correlate of Protection Against Seasonal A/H3N2 Influenza Disease'. In: *Open Forum Infect Dis* 2.2, ofv067.

Berger, S. I. and R. Iyengar (2009). 'Network analyses in systems pharmacology'. In: *Bioinformatics* 25.19, pp. 2466–2472. DOI: 10.1093/bioinformatics/btp465.

Beyer, W.E.P, A.M Palache, G Lüchters, J Nauta and A.D.M.E Osterhaus (2004). 'Seroprotection rate, mean fold increase, seroconversion rate: which parameter adequately expresses seroresponse to influenza vaccination?' In: *Virus Research* 103.1-2, pp. 125–132. DOI: 10.1016/j.virusres.2004.02.024.

Bhattaram, V. A., C. Bonapace, D. M. Chilukuri, J. Z. Duan, C. Garnett et al. (2007). 'Impact of pharmacometric reviews on new drug approval and labeling decisions–a survey of 31 new drug applications submitted between 2005 and 2006'. In: *Clin. Pharmacol. Ther.* 81.2, pp. 213–221. DOI: 10.1038/sj.clpt.6100051.

Bhattaram, Venkatesh A. et al. (2005). 'Impact of pharmacometrics on drug approval and labeling decisions: A survey of 42 new drug applications'. In: *The AAPS Journal* 7.3, E503–E512. DOI: 10.1208/aapsj070351.

Bild, A. H. (2002). 'Cytoplasmic transport of Stat3 by receptor-mediated endocytosis'. In: *The EMBO Journal* 21.13, pp. 3255–3263. DOI: 10.1093/emboj/cdf351.

Black, J. W. and P. Leff (1983). 'Operational Models of Pharmacological Agonism'. In: *Proceedings of the Royal Society B: Biological Sciences* 220.1219, pp. 141–162. DOI: 10.1098/rspb.1983.0093.

Blanchard, F., Y. Wang, E. Kinzie, L. Duplomb, A. Godard and H. Baumann (2001). 'Oncostatin M Regulates the Synthesis and Turnover of gp130, Leukemia Inhibitory Factor Receptor , and Oncostatin

M Receptor by Distinct Mechanisms'. In: *Journal of Biological Chemistry* 276.50, pp. 47038–47045. DOI: `10.1074/jbc.m107971200`.

Bongioanni, P., F. Lombardo, G. Moscato, S. Mosti and G. Meucci (2000). 'T-cell interleukin-6 receptor binding in interferon-beta-1b-treated multiple sclerosis patients'. In: *Eur. J. Neurol.* 7.6, pp. 647–653.

Bosch, T. M. et al. (2006). 'Pharmacogenetic Screening of CYP3A and ABCB1 in Relation to Population Pharmacokinetics of Docetaxel'. In: *Clinical Cancer Research* 12.19, pp. 5786–5793. DOI: `10.1158/1078-0432.ccr-05-2649`.

Boulanger, M. J. (2003). 'Hexameric Structure and Assembly of the Interleukin-6/IL-6 -Receptor/gp130 Complex'. In: *Science* 300.5628, pp. 2101–2104. DOI: `10.1126/science.1083901`.

Breedveld, F. C., M. H. Weisman, A. F. Kavanaugh, S. B. Cohen, K. Pavelka, R. van Vollenhoven, J. Sharp, J. L. Perez and G. T. Spencer-Green (2006). 'The PREMIER study: A multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment'. In: *Arthritis Rheum.* 54.1, pp. 26–37.

Breiman, Leo (2001). 'Random Forests'. In: *Machine Learning* 45.1, pp. 5–32. ISSN: 1573-0565. DOI: `10.1023/A:1010933404324`.

Breitling, Rainer (2010). 'What is systems biology?' In: *Frontiers in Physiology* 1. DOI: `10.3389/fphys.2010.00009`.

Briso, Eva M., Oliver Dienz and Mercedes Rincon (2008). 'Cutting Edge: Soluble IL-6R Is Produced by IL-6R Ectodomain Shedding in Activated CD4 T Cells'. In: *The Journal of Immunology* 180.11, pp. 7102–7106. DOI: `10.4049/jimmunol.180.11.7102`.

Buckwalter, J. et al., eds. (2007). *Osteoarthritis, Inflammation and Degradation: A Continuum*. IOS Press. Chap. Biomarkers of Matrix Fragments, Inflammation Markers, pp. 267–269.

CHMP Assessment report RoActemra. EMA/CHMP/606295/2013, 20 February 2014.

Çalışkan, Minal, Samuel W. Baker, Yoav Gilad and Carole Ober (2015). 'Host Genetic Variation Influences Gene Expression Response to Rhinovirus Infection'. In: *PLOS Genetics* 11.4. Ed. by Greg Gibson, e1005111. DOI: `10.1371/journal.pgen.1005111`.

Cantwell, Carrie A., Esta Sterneck and Peter F. Johnson (1998). 'Interleukin-6-Specific Activation of the C/EBPδ Gene in Hepatocytes Is Mediated by Stat3 and Sp1'. In: *Molecular and Cellular Biology* 18.4, pp. 2108–2117. DOI: `10.1128/mcb.18.4.2108`.

Casanovas, Guillem, Anashua Banerji, Flavia d'Alessio, Martina U. Muckenthaler and Stefan Legewie (2014). 'A Multi-Scale Model of Hepcidin Promoter Regulation Reveals Factors Controlling Sys-

temic Iron Homeostasis'. In: *PLoS Comput Biol* 10.1. Ed. by Martha L. Bulyk, e1003421. DOI: `10.1371/journal.pcbi.1003421`.

Chen, R.-H., M.-C. Chang, Y.-H. Su, Y.-T. Tsai and M.-L. Kuo (1999). 'Interleukin-6 Inhibits Transforming Growth Factor- -induced Apoptosis through the Phosphatidylinositol 3-Kinase/Akt and Signal Transducers and Activators of Transcription 3 Pathways'. In: *Journal of Biological Chemistry* 274.33, pp. 23013–23019. DOI: `10.1074/jbc.274.33.23013`.

Choy, E. H. S., D. J. A. Connolly, N. Rapson, S. Jeal, J. C. C. Brown, G. H. Kingsley, G. S. Panayi and J. M. Johnston (2000). 'Pharmacokinetic, pharmacodynamic and clinical effects of a humanized IgG1 anti-CD4 monoclonal antibody in the peripheral blood and synovial fluid of rheumatoid arthritis patients'. In: *Rheumatology* 39.10, pp. 1139–1146. DOI: `10.1093/rheumatology/39.10.1139`.

Chung, E. S. (2003). 'Randomized, Double-Blind, Placebo-Controlled, Pilot Trial of Infliximab, a Chimeric Monoclonal Antibody to Tumor Necrosis Factor- , in Patients With Moderate-to-Severe Heart Failure: Results of the Anti-TNF Therapy Against Congestive Heart failure (ATTACH) Trial'. In: *Circulation* 107.25, pp. 3133–3140. DOI: `10.1161/01.cir.0000077913.60364.d2`.

Cimica, Velasco, Hui-Chen Chen, Janaki K. Iyer and Nancy C. Reich (2011). 'Dynamics of the STAT3 Transcription Factor: Nuclear Import Dependent on Ran and Importin-$\beta 1$'. In: *PLoS ONE* 6.5. Ed. by Venugopalan Cheriyath, e20188. DOI: `10.1371/journal.pone.0020188`.

ClinicalTrials.gov; NCT01856309 (2013 - [cited 2018 March]). 'Long-term Safety and Efficacy of Sirukumab in Participants With RA Completing Studies CNTO136ARA3002 or CNTO136ARA3003 (SIRROUND-LTE)'. In: *ClinicalTrials.gov [Internet]*. NLM Identifier: NCT01856309, Available from: Bethesda (MD): National Library of Medicine (US). URL: `https://clinicaltrials.gov/ct2/show/NCT01856309`.

ClinicalTrials.gov; NCT02473289 (2015 - [cited 2018 March]). 'An Efficacy and Safety Study of Sirukumab in Participants With Major Depressive Disorder'. In: *ClinicalTrials.gov [Internet]*. NLM Identifier: NCT02473289, Available from: Bethesda (MD): National Library of Medicine (US). URL: `https://clinicaltrials.gov/ct2/show/NCT02473289`.

Cockcroft, D. W. and M. H. Gault (1976). 'Prediction of creatinine clearance from serum creatinine'. In: *Nephron* 16.1, pp. 31–41.

Cohen-Pfeffer, J. L., S. Gururangan, T. Lester, D. A. Lim, A. J. Shaywitz, M. Westphal and I. Slavc (2017). 'Intracerebroventricular Delivery as a Safe, Long-Term Route of Drug Administration'. In: *Pediatr. Neurol.* 67, pp. 23–35.

Cohen, Paul R (2015). 'DARPA's Big Mechanism program'. In: *Physical Biology* 12.4, p. 045008. DOI: `10.1088/1478-3975/12/4/045008`.

Conover, W. J. (1999). *Practical Nonparametric Statistics, 3rd*. Wiley. ISBN: 0471160687.

Cosgrove, J, J Butler, K Alden, M Read, V Kumar, L Cucurull-Sanchez, J Timmis and M Coles (2015). 'Agent-Based Modeling in Systems Pharmacology'. In: *CPT: Pharmacometrics & Systems Pharmacology* 4.11, pp. 615–629. DOI: 10.1002/psp4.12018.

Cross, Marita et al. (2014). 'The global burden of rheumatoid arthritis: estimates from the Global Burden of Disease 2010 study'. In: *Annals of the Rheumatic Diseases* 73.7, pp. 1316–1322. DOI: 10.1136/annrheumdis-2013-204627.

Csajka, Chantal and Davide Verotta (2006). 'Pharmacokinetic–Pharmacodynamic Modelling: History and Perspectives'. In: *Journal of Pharmacokinetics and Pharmacodynamics* 33.3, pp. 227–279. DOI: 10.1007/s10928-005-9002-0.

Csermely, P., V. Agoston and S. Pongor (2005). 'The efficiency of multi-target drugs: the network approach might help drug design'. In: *Trends Pharmacol. Sci.* 26.4, pp. 178–182.

Cucurull-Sanchez, Lourdes, Karen G. Spink and Sterghios Athanasios Moschos (2012). 'Relevance of systems pharmacology in drug discovery'. In: *Drug Discovery Today* 17.13-14, pp. 665–670. DOI: 10.1016/j.drudis.2012.01.015.

Cummings, Jeffrey L, Travis Morstorf and Kate Zhong (2014). 'Alzheimer's disease drug-development pipeline: few candidates, frequent failures'. In: *Alzheimer's Research & Therapy* 6.4, p. 37. DOI: 10.1186/alzrt269.

Cutolo, M., A. Bisso, A. Sulli, L. Felli, M. Briata, C. Pizzorni and B. Villaggio (2000). 'Antiproliferative and antiinflammatory effects of methotrexate on cultured differentiating myeloid monocytic cells (THP-1) but not on synovial macrophages from patients with rheumatoid arthritis'. In: *J. Rheumatol.* 27.11, pp. 2551–2557.

Dasgupta, Anirban, Petros Drineas, Boulos Harb, Vanja Josifovski and Michael W Mahoney (2007). 'Feature selection methods for text classification'. In: *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, pp. 230–239.

Davis, A. P., T. C. Wiegers, M. C. Rosenstein and C. J. Mattingly (2012). 'MEDIC: a practical disease vocabulary used at the Comparative Toxicogenomics Database'. In: *Database* 2012.0, bar065–bar065. DOI: 10.1093/database/bar065.

Davis, Allan Peter, Cynthia J. Grondin, Robin J. Johnson, Daniela Sciaky, Benjamin L. King, Roy McMorran, Jolene Wiegers, Thomas C. Wiegers and Carolyn J. Mattingly (2016). 'The Comparative Toxicogenomics Database: update 2017'. In: *Nucleic Acids Research* 45.D1, pp. D972–D978. DOI: 10.1093/nar/gkw838.

Deftereos, Spyros N., Christos Andronis, Ellen J. Friedla, Aris Persidis and Andreas Persidis (2011). 'Drug repurposing and ad-

verse event prediction using high-throughput literature analysis'. In: *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 3.3, pp. 323–334. DOI: `10.1002/wsbm.147`.

Demner-Fushman, Dina, James G. Mork, Sonya E. Shooshan and Alan R. Aronson (2010). 'UMLS content views appropriate for NLP processing of the biomedical literature vs. clinical text'. In: *Journal of Biomedical Informatics* 43.4, pp. 587–594. DOI: `10.1016/j.jbi.2010.02.005`.

Desgeorges, A., C. Gabay, P. Silacci, D. Novick, P. Roux-Lombard, G. Grau, J. M. Dayer, T. Vischer and P. A. Guerne (1997). 'Concentrations and origins of soluble interleukin 6 receptor-alpha in serum and synovial fluid'. In: *J. Rheumatol.* 24.8, pp. 1510–1516.

DiMasi, Joseph A. (November 18, 2014). 'Innovation in the Pharmaceutical Industry: New estimates of R&D costs'. In: *R&D Cost Study Briefing*. Available from: Tufts Center for the Study of Drug Development (TCSDD). URL: `http://csdd.tufts.edu/files/uploads/Tufts_CSDD_briefing_on_RD_cost_study_-_Nov_18,_2014..pdf`.

Diamant, Marcus, Klaus Rieneck, Nadir Mechti, Xue-Guang Zhang, Morten Svenson, Klaus Bendtzen and Bernard Klein (1997). 'Cloning and expression of an alternatively spliced mRNA encoding a soluble form of the human interleukin-6 signal transducer gp130 1'. In: *FEBS Letters* 412.2, pp. 379–384. DOI: `10.1016/s0014-5793(97)00750-3`.

Dienz, Oliver and Mercedes Rincon (2009). 'The effects of IL-6 on CD4 T cell responses'. In: *Clinical Immunology* 130.1, pp. 27–33. DOI: `10.1016/j.clim.2008.08.018`.

Dittrich, E., S. Rose-John, C. Gerhartz, J. Mullberg, T. Stoyan, K. Yasukawa, P. C. Heinrich and L. Graeve (1994). 'Identification of a region within the cytoplasmic domain of the interleukin-6 (IL-6) signal transducer gp130 important for ligand-induced endocytosis of the IL-6 receptor'. In: *J. Biol. Chem.* 269.29, pp. 19014–19020.

Dittrich, E., C. R. Haft, L. Muys, P. C. Heinrich and L. Graeve (1996). 'A Di-leucine Motif and an Upstream Serine in the Interleukin-6 (IL-6) Signal Transducer gp130 Mediate Ligand-induced Endocytosis and Down-regulation of the IL-6 Receptor'. In: *Journal of Biological Chemistry* 271.10, pp. 5487–5494. DOI: `10.1074/jbc.271.10.5487`.

Dulong, S., A. Ballesta, A. Okyar and F. Levi (2015). 'Identification of Circadian Determinants of Cancer Chronotherapy through In Vitro Chronopharmacology and Mathematical Modeling'. In: *Molecular Cancer Therapeutics* 14.9, pp. 2154–2164. DOI: `10.1158/1535-7163.mct-15-0129`.

Durant, Lydia et al. (2010). 'Diverse Targets of the Transcription Factor STAT3 Contribute to T Cell Pathogenicity and Homeostasis'. In: *Immunity* 32.5, pp. 605–615. DOI: `10.1016/j.immuni.2010.05.003`.

Dwivedi, G., L. Fitz, M. Hegen, S. W. Martin, J. Harrold et al. (2014). 'A multiscale model of interleukin-6-mediated immune regulation in Crohn's disease and its application in drug discovery and development'. In: *CPT Pharmacometrics Syst Pharmacol* 3, e89. DOI: `10.1038/psp.2013.64`.

Edginton, A. N., W. Schmitt and S. Willmann (2006). 'Application of physiology-based pharmacokinetic and pharmacodynamic modeling to individualized target-controlled propofol infusions'. In: *Adv Ther* 23.1, pp. 143–158.

Ericsson, A. C., M. J. Crim and C. L. Franklin (2013). 'A brief history of animal modeling'. In: *Mo Med* 110.3, pp. 201–205.

Ette, Ene I. and Paul J. Williams (2007). *Pharmacometrics: The Science of Quantitative Pharmacology*. Ed. by Ene I. Ette and Paul J. Williams. John Wiley & Sons, Inc.

FDA Briefing Document, Arthritis Advisory Committee Meeting (August 2017). *U.S Food and Drug Administration, FDA Briefing Document, Arthritis Advisory Committee Meeting*. Available from: URL: `https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM569150.pdf`.

Fan, Jianqing, Fang Han and Han Liu (2014). 'Challenges of Big Data analysis'. In: *National Science Review* 1.2, pp. 293–314. DOI: `10.1093/nsr/nwt032`.

Ferl, Gregory Z., Frank-Peter Theil and Harvey Wong (2016). 'Physiologically based pharmacokinetic models of small molecules and therapeutic antibodies: a mini-review on fundamental concepts and applications'. In: *Biopharmaceutics & Drug Disposition* 37.2, pp. 75–92. DOI: `10.1002/bdd.1994`.

Fleischmann, R., J. van Adelsberg, Y. Lin, G. D. Castelar-Pinheiro, J. Brzezicki, P. Hrycaj, N. M. Graham, H. van Hoogstraten, D. Bauer and G. R. Burmester (2017). 'Sarilumab and Nonbiologic Disease-Modifying Antirheumatic Drugs in Patients With Active Rheumatoid Arthritis and Inadequate Response or Intolerance to Tumor Necrosis Factor Inhibitors'. In: *Arthritis Rheumatol* 69.2, pp. 277–290.

Flipo, R. M., J. F. Maillefert, P. Chazerain, I. Idier, M. Coudert and J. Tebib (2017). 'Factors influencing the use of tocilizumab as monotherapy in patients with rheumatoid arthritis in a real-life setting: results at 1 year of the ACT-SOLO study'. In: *RMD Open* 3.1, e000340. DOI: `10.1136/rmdopen-2016-000340`.

Fontes, Jillian A., Noel R. Rose and Daniela Čiháková (2015). 'The varying faces of IL-6: From cardiac protection to cardiac failure'. In: *Cytokine* 74.1, pp. 62–68. DOI: `10.1016/j.cyto.2014.12.024`.

Foo, Jasmine, Juliann Chmielecki, William Pao and Franziska Michor (2012). 'Effects of Pharmacokinetic Processes and Varied Dosing Schedules on the Dynamics of Acquired Resistance to Erlotinib in

EGFR-Mutant Lung Cancer'. In: *Journal of Thoracic Oncology* 7.10, pp. 1583–1593. DOI: 10.1097/jto.0b013e31826146ee.

Forman, George (2003). 'An Extensive Empirical Study of Feature Selection Metrics for Text Classification'. In: *Journal of Machine Learning Research* 3, pp. 1289–1305.

Francis, Thomas, Jonas E. Salk, Harold E. Pearson and Philip N. Brown (1945). 'PROTECTIVE EFFECT OF VACCINATION AGAINST INDUCED INFLUENZA A 1'. In: *Journal of Clinical Investigation* 24.4, pp. 536–546. DOI: 10.1172/jci101633.

Frey, Nicolas, Susan Grange and Thasia Woodworth (2010). 'Population Pharmacokinetic Analysis of Tocilizumab in Patients With Rheumatoid Arthritis'. In: *The Journal of Clinical Pharmacology* 50.7, pp. 754–766. DOI: 10.1177/0091270009350623.

Fujimoto, K., H. Ida, Y. Hirota, M. Ishigai, J. Amano and Y. Tanaka (2015). 'Intracellular Dynamics and Fate of a Humanized Anti-Interleukin-6 Receptor Monoclonal Antibody, Tocilizumab'. In: *Molecular Pharmacology* 88.4, pp. 660–675. DOI: 10.1124/mol.115. 099184.

Gama, João (2000). 'A Linear-Bayes Classifier'. In: *Advances in Artificial Intelligence*. Ed. by Maria Carolina Monard and Jaime Simão Sichman. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 269–279.

Gambús, Pedro L. and Iñaki F. Trocóniz (2014). 'Pharmacokinetic-pharmacodynamic modelling in anaesthesia'. In: *British Journal of Clinical Pharmacology* 79.1, pp. 72–84. DOI: 10.1111/bcp.12286.

Gao, Wei, Jennifer McCormick, Mary Connolly, Emese Balogh, Douglas J Veale and Ursula Fearon (2014). 'Hypoxia and STAT3 signalling interactions regulate pro-inflammatory pathways in rheumatoid arthritis'. In: *Annals of the Rheumatic Diseases* 74.6, pp. 1275–1283. DOI: 10.1136/annrheumdis-2013-204105.

Gaohua, Lu, Khaled Abduljalil, Masoud Jamei, Trevor N. Johnson and Amin Rostami-Hodjegan (2012). 'A pregnancy physiologically based pharmacokinetic (p-PBPK) model for disposition of drugs metabolized by CYP1A2, CYP2D6 and CYP3A4'. In: *British Journal of Clinical Pharmacology* 74.5, pp. 873–885. DOI: 10.1111/j. 1365-2125.2012.04363.x.

Garbers, Christoph et al. (2011). 'Inhibition of Classic Signaling Is a Novel Function of Soluble Glycoprotein 130 (sgp130), Which Is Controlled by the Ratio of Interleukin 6 and Soluble Interleukin 6 Receptor'. In: *Journal of Biological Chemistry* 286.50, pp. 42959–42970. DOI: 10.1074/jbc.m111.295758.

Garbers, Christoph, Samadhi Aparicio-Siegmund and Stefan Rose-John (2015). 'The IL-6/gp130/STAT3 signaling axis: recent advances towards specific inhibition'. In: *Current Opinion in Immunology* 34, pp. 75–82. DOI: 10.1016/j.coi.2015.02.008.

Garner, J. P. (2014). 'The Significance of Meaning: Why Do Over 90% of Behavioral Neuroscience Results Fail to Translate to Humans, and What Can We Do to Fix It?' In: *ILAR Journal* 55.3, pp. 438–456. DOI: 10.1093/ilar/ilu047.

Gearing, DP et al. (1992). 'The IL-6 signal transducer, gp130: an oncostatin M receptor and affinity converter for the LIF receptor'. In: *Science* 255.5050, pp. 1434–1437. DOI: 10.1126/science.1542794.

Geerts, H., A. Spiros, P. Roberts and L. Alphs (2018). 'A quantitative systems pharmacology study on optimal scenarios for switching to paliperidone palmitate once-monthly'. In: *Schizophr. Res.*

Geerts, Hugo, Patrick Roberts, Athan Spiros and Steven Potkin (2015). 'Understanding responder neurobiology in schizophrenia using a quantitative systems pharmacology model: Application to iloperidone'. In: *Journal of Psychopharmacology* 29.4, pp. 372–382. DOI: 10.1177/0269881114568042.

Gesztelyi, Rudolf, Judit Zsuga, Adam Kemeny-Beke, Balazs Varga, Bela Juhasz and Arpad Tosaki (2012). 'The Hill equation and the origin of quantitative pharmacology'. In: *Archive for History of Exact Sciences* 66.4, pp. 427–438.

Gibiansky, Leonid and Nicolas Frey (2011). 'Linking interleukin-6 receptor blockade with tocilizumab and its hematological effects using a modeling approach'. In: *J Pharmacokinet Pharmacodyn* 39.1, pp. 5–16. DOI: 10.1007/s10928-011-9227-z.

Gillick, K., D. Pollpeter, P. Phalora, E.-Y. Kim, S. M. Wolinsky and M. H. Malim (2012). 'Suppression of HIV-1 Infection by APOBEC3 Proteins in Primary Human CD4+ T Cells Is Associated with Inhibition of Processive Reverse Transcription as Well as Excessive Cytidine Deamination'. In: *Journal of Virology* 87.3, pp. 1508–1517. DOI: 10.1128/jvi.02587-12.

Gkoutos, G. V., P. N. Schofield and R. Hoehndorf (2012). 'The Units Ontology: a tool for integrating units of measurement in science'. In: *Database* 2012.0, bas033–bas033. DOI: 10.1093/database/bas033.

Gonzalez, Graciela H., Tasnia Tahsin, Britton C. Goodale, Anna C. Greene and Casey S. Greene (2015). 'Recent Advances and Emerging Applications in Text and Data Mining for Biomedical Discovery'. In: *Briefings in Bioinformatics* 17.1, pp. 33–42. DOI: 10.1093/bib/bbv087.

Gough, Albert, Tong Ying Shun, D. Lansing Taylor and Mark Schurdak (2016). 'A metric and workflow for quality control in the analysis of heterogeneity in phenotypic profiles and screens'. In: *Methods* 96, pp. 12–26. DOI: 10.1016/j.ymeth.2015.10.007.

Graaf, Piet H. van der and Neil Benson (2011). 'Systems Pharmacology: Bridging Systems Biology and Pharmacokinetics-Pharmacodynamics (PKPD) in Drug Discovery and Development'. In: *Pharm Res* 28.7, pp. 1460–1464. DOI: 10.1007/s11095-011-0467-9.

Graeve, Lutz, Tatjana A. Korolenko, Ulrike Hemmann, Oliver Weier-
    gräber, Elke Dittrich and Peter C. Heinrich (1996). 'A complex
    of the soluble interleukin-6 receptor and interleukin-6 is intern-
    alized via the signal transducer gp130'. In: *FEBS Letters* 399.1-2,
    pp. 131–134. DOI: 10.1016/s0014-5793(96)01305-1.

Grigor, C., H. Capell, A. Stirling, A. D. McMahon, P. Lock, R. Vallance,
    W. Kincaid and D. Porter (2004). 'Effect of a treatment strategy
    of tight control for rheumatoid arthritis (the TICORA study): a
    single-blind randomised controlled trial'. In: *Lancet* 364.9430, pp. 263–
    269. DOI: 10.1016/S0140-6736(04)16676-2.

Grötzinger, J., T. Kernebeck, K.-J. Kallen and S. Rose-John (1999). 'IL-6
    Type Cytokine Receptor Complexes: Hexamer, Tetramer or Both?'
    In: *Biological Chemistry* 380.7-8. DOI: 10.1515/bc.1999.100.

Guengerich, F. P. (2011). 'Mechanisms of drug toxicity and relevance
    to pharmaceutical development'. In: *Drug Metab. Pharmacokinet.*
    26.1, pp. 3–14.

Guerne, P A, B L Zuraw, J H Vaughan, D A Carson and M Lotz (1989).
    'Synovium as a source of interleukin 6 in vitro. Contribution to
    local and systemic manifestations of arthritis.' In: *Journal of Clin-
    ical Investigation* 83.2, pp. 585–592. DOI: 10.1172/jci113921.

Hansson, E K, M A Amantea, P Westwood, P A Milligan, B E Houk,
    J French, M O Karlsson and L E Friberg (2013). 'PKPD Modeling
    of VEGF, sVEGFR-2, sVEGFR-3, and sKIT as Predictors of Tumor
    Dynamics and Overall Survival Following Sunitinib Treatment in
    GIST'. In: *CPT: Pharmacometrics & Systems Pharmacology* 2.11, e84.
    DOI: 10.1038/psp.2013.61.

Hardee, M. E. (2006). 'Erythropoietin Biology in Cancer'. In: *Clinical
    Cancer Research* 12.2, pp. 332–339. DOI: 10.1158/1078-0432.ccr-
    05-1771.

Harnisch, L, I Matthews, J Chard and MO Karlsson (2013). 'Drug
    and disease model resources: a consortium to create standards
    and tools to enhance model-based drug development'. In: *CPT
    Pharmacometrics Syst Pharmacol* 2, e34. ISSN: 2163-8306. DOI: 10.
    1038/psp.2013.10.

Harrison, R. K. (2016). 'Phase II and phase III failures: 2013-2015'. In:
    Nat Rev Drug Discov 15.12, pp. 817–818.

Hay, Michael, David W Thomas, John L Craighead, Celia Economides
    and Jesse Rosenthal (2014). 'Clinical development success rates
    for investigational drugs'. In: *Nat Biotechnol* 32.1, pp. 40–51. DOI:
    10.1038/nbt.2786.

Heinemann, Axel, Friedel Wischhusen, Klaus Püschel and Xavier Ro-
    giers (1999). 'Standard liver volume in the caucasian population'.
    In: *Liver Transplantation and Surgery* 5.5, pp. 366–368. DOI: 10.
    1002/lt.500050516.

Heinrich, Peter C., Iris Behrmann, Gerhard Müller-Newen, Fred Schaper
    and Lutz Graeve (1998). 'Interleukin-6-type cytokine signalling

through the gp130/Jak/STAT pathway'. In: *Biochem. J.* 334.2, pp. 297–314. DOI: 10.1042/bj3340297.

Heinrich, Peter C., Iris Behrmann, Serge Haan, Heike M. Hermanns, Gerhard Müller-Newen and Fred Schaper (2003). 'Principles of interleukin (IL)-6-type cytokine signalling and its regulation'. In: *Biochem. J.* 374.1, pp. 1–20. DOI: 10.1042/bj20030407.

Helmick, Charles G. et al. (2007). 'Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: Part I'. In: *Arthritis & Rheumatism* 58.1, pp. 15–25. DOI: 10.1002/art.23177.

Hench, P. S., E. C. Kendall, C. H. Slocumb and H. F. Polley (1949). 'The effect of a hormone of the adrenal cortex (17-hydroxy-11-dehydrocorticosterone: compound E) and of pituitary adrenocortical hormone in arthritis: preliminary report'. In: *Ann. Rheum. Dis.* 8.2, pp. 97–104.

Henkel, Janin, Daniela Gärtner, Christoph Dorn, Claus Hellerbrand, Nancy Schanze, Sheila R Elz and Gerhard P Püschel (2011). 'Oncostatin M produced in Kupffer cells in response to PGE2: possible contributor to hepatic insulin resistance and steatosis'. In: *Lab Invest* 91.7, pp. 1107–1117. DOI: 10.1038/labinvest.2011.47.

Hennigan, S. and A. Kavanaugh (2008). 'Interleukin-6 inhibitors in the treatment of rheumatoid arthritis'. In: *Ther Clin Risk Manag* 4.4, pp. 767–775.

Herrero-Zazo, María, Isabel Segura-Bedmar and Paloma Martínez (2013). 'Annotation Issues in Pharmacological Texts'. In: *Procedia - Social and Behavioral Sciences* 95, pp. 211–219. DOI: 10.1016/j.sbspro.2013.10.641.

Hibi, Masahiko, Masaaki Murakami, Mikiyoshi Saito, Toshio Hirano, Tetsuya Taga and Tadamitsu Kishimoto (1990). 'Molecular cloning and expression of an IL-6 signal transducer, gp130'. In: *Cell* 63.6, pp. 1149–1157. DOI: 10.1016/0092-8674(90)90411-7.

Hill, Archibald Vivian (1910). 'The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves'. In: *The Journal of Physiology* 40, pp. i–vii.

Hindmarsh, A. C. (1983). ODEPACK, a systematized collection of ODE solvers. In R. Stepleman, editor, Scientific Computing, Vol. 1 of IMACS Transactions on Sci- entific Computation, pages 55–64. IMACS / North-Holland, Amsterdam.

Hobson, D., R. L. Curry, A. S. Beare and A. Ward-Gardner (1972). 'The role of serum haemagglutination-inhibiting antibody in protection against challenge infection with influenza A2 and B viruses'. In: *J Hyg (Lond)* 70.4, pp. 767–777.

Hodgkin, A. L. and A. F. Huxley (1952). 'A quantitative description of membrane current and its application to conduction and excitation in nerve'. In: *J. Physiol. (Lond.)* 117.4, pp. 500–544.

Holford, N. H. and L. B. Sheiner (1981). 'Understanding the dose-effect relationship: clinical application of pharmacokinetic-pharmacodynamic models'. In: *Clin Pharmacokinet* 6.6, pp. 429–453.

— (1982). 'Kinetics of pharmacologic response'. In: *Pharmacol. Ther.* 16.2, pp. 143–166.

Hoops, S., S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes and U. Kummer (2006). 'COPASI–a COmplex PAthway SImulator'. In: *Bioinformatics* 22.24, pp. 3067–3074. DOI: `10.1093/bioinformatics/btl485`.

Hopkins, A. L. (2008). 'Network pharmacology: the next paradigm in drug discovery'. In: *Nat. Chem. Biol.* 4.11, pp. 682–690.

Horneff, G., G. R. Burmester, F. Emmrich and J. R. Kalden (1991). 'Treatment of rheumatoid arthritis with an anti-CD4 monoclonal antibody'. In: *Arthritis Rheum.* 34.2, pp. 129–140.

Huang, Y., S.A. Anderson, R.A. Forshee and H. Yang (2017). 'A modified dose-response model that describes the relationship between haemagglutination inhibition titre and protection against influenza infection'. In: *Journal of Applied Microbiology* 124.1, pp. 294–301. DOI: `10.1111/jam.13628`.

Hübner, Katrin, Sven Sahle and Ursula Kummer (2011). 'Applications and trends in systems biology in biochemistry'. In: *FEBS Journal* 278.16, pp. 2767–2857. DOI: `10.1111/j.1742-4658.2011.08217.x`.

Hucka, M., A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle et al. (2003). 'The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models'. In: *Bioinformatics* 19.4, pp. 524–531. DOI: `10.1093/bioinformatics/btg015`.

Hunter, Christopher A and Simon A Jones (2015). 'IL-6 as a keystone cytokine in health and disease'. In: *Nature Immunology* 16.5, pp. 448–457. DOI: `10.1038/ni.3153`.

Ingelman-Sundberg, Magnus (2004). 'Pharmacogenetics of cytochrome P450 and its applications in drug therapy: the past, present and future'. In: *Trends in Pharmacological Sciences* 25.4, pp. 193–200. DOI: `10.1016/j.tips.2004.02.007`.

Isomaki, P., I. Junttila, K.-L. Vidqvist, M. Korpela and O. Silvennoinen (2014). 'The activity of JAK-STAT pathways in rheumatoid arthritis: constitutive activation of STAT3 correlates with interleukin 6 levels'. In: *Rheumatology* 54.6, pp. 1103–1113. DOI: `10.1093/rheumatology/keu430`.

Iuliano, A Danielle et al. (2017). 'Estimates of global seasonal influenza-associated respiratory mortality: a modelling study'. In: *The Lancet*. DOI: `10.1016/s0140-6736(17)33293-2`.

Iyer, Srinivasan V, Rave Harpaz, Paea LePendu, Anna Bauer-Mehren and Nigam H Shah (2014). 'Mining clinical text for signals of adverse drug-drug interactions'. In: *Journal of the American Medical*

Informatics Association 21.2, pp. 353–362. DOI: `10.1136/amiajnl-2013-001612`.

Jackson, Clive G. (2014). 'Target Selection and Validation in Drug Discovery'. In: *Reducing Drug Attrition*. Ed. by James R. Empfield and Michael P Clark. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–72. ISBN: 978-3-662-43914-2. DOI: `10.1007/7355_2014_59`.

Jackson, E. L. and H. Lu (2016). 'Three-dimensional models for studying development and disease: moving on from organisms to organs-on-a-chip and organoids'. In: *Integrative Biology* 8.6, pp. 672–683. DOI: `10.1039/c6ib00039h`.

Jackson, Robert C., Giovanni Y. Di Veroli, Siang-Boon Koh, Ian Goldlust, Frances M. Richards and Duncan I. Jodrell (2017). 'Modelling of the cancer cell cycle as a tool for rational drug development: A systems pharmacology approach to cyclotherapy'. In: *PLOS Computational Biology* 13.5. Ed. by James Gallo, e1005529. DOI: `10.1371/journal.pcbi.1005529`.

Jacobs, C. A., M. P. Beckmann, K. Mohler, C. R. Maliszewski, W. C. Fanslow and D. H. Lynch (1993). 'Pharmacokinetic parameters and biodistribution of soluble cytokine receptors'. In: *Int Rev Exp Pathol* 34 Pt B, pp. 123–135.

Jaing, Crystal et al. (2017). 'Gene expression analysis of whole blood RNA from pigs infected with low and high pathogenic African swine fever viruses'. In: *Scientific Reports* 7.1. DOI: `10.1038/s41598-017-10186-4`.

Jalali, Saeideh, Amir Zarrinhaghighi, Saman Sadraei, Younes Ghasemi, Amirhossein Sakhteman and Pouya Faridi (2018). 'A Systems Pharmacology Study for Deciphering Anti Depression Activity of Nardostachys jatamansi'. In: *Current Drug Metabolism* 19. DOI: `10.2174/1389200219666180305151011`.

Jamei, Masoud, Steve Marciniak, Kairui Feng, Adrian Barnett, Geoffrey Tucker and Amin Rostami-Hodjegan (2009). 'The Simcyp ® Population-based ADME Simulator'. In: *Expert Opinion on Drug Metabolism & Toxicology* 5.2, pp. 211–223. DOI: `10.1517/17425250802691074`.

Jansky, L., P. Reymanova and J. Kopecky (2003). 'Dynamics of cytokine production in human peripheral blood mononuclear cells stimulated by LPS or infected by Borrelia'. In: *Physiol Res* 52.6, pp. 593–598.

Jensen, Lars Juhl, Jasmin Saric and Peer Bork (2006). 'Literature mining for the biologist: from information retrieval to biological discovery'. In: *Nature Reviews Genetics* 7.2, pp. 119–129. DOI: `10.1038/nrg1768`.

Jin, Chuanxin, Tinghuai Ma, Rongtao Hou, Meili Tang, Yuan Tian, Abdullah Al-Dhelaan and Mznah Al-Rodhaan (2015). 'Chi-square Statistics Feature Selection Based on Term Frequency and Distribution for Text Categorization'. In: *IETE Journal of Research* 61.4, pp. 351–362. DOI: `10.1080/03772063.2015.1021385`.

Joachims, Thorsten (1998). 'Text categorization with Support Vector Machines: Learning with many relevant features'. In: *Machine Learning: ECML-98*. Ed. by Claire Nédellec and Céline Rouveirol. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 137–142.

Johnson & Johnson (2017). 'Janssen receives complete response letter from U.S. FDA for sirukumab biologics license application'. In: *Johnson & Johnson*. URL: https://www.jnj.com/media-center/press-releases/janssen-receives-complete-response-letter-from-us-fda-for-sirukumab-biologics-license-application (visited on 27/03/2018).

Johnson, W. E., C. Li and A. Rabinovic (2007). 'Adjusting batch effects in microarray expression data using empirical Bayes methods'. In: *Biostatistics* 8.1, pp. 118–127. DOI: 10.1093/biostatistics/kxj037.

Jones, G. W. et al. (2010). 'Loss of CD4+ T Cell IL-6R Expression during Inflammation Underlines a Role for IL-6 Trans Signaling in the Local Maintenance of Th17 Cells'. In: *The Journal of Immunology* 184.4, pp. 2130–2139. DOI: 10.4049/jimmunol.0901528.

Jones, S. A. (2001). 'The soluble interleukin 6 receptor: mechanisms of production and implications in disease'. In: *The FASEB Journal* 15.1, pp. 43–58. DOI: 10.1096/fj.99-1003rev.

Jones, S. A., D. Novick, S. Horiuchi, N. Yamamoto, A. J. Szalai and G. M. Fuller (1999). 'C-reactive protein: a physiological activator of interleukin 6 receptor shedding'. In: *J. Exp. Med.* 189.3, pp. 599–604.

Jones, Simon A., Sankichi Horiuchi, Daniela Novick, Naoki Yamamoto and Gerald M. Fuller (1998). 'Shedding of the soluble IL-6 receptor is triggered by Ca2+ mobilization, while basal release is predominantly the product of differential mRNA splicing in THP-1 cells'. In: *European Journal of Immunology* 28.11, pp. 3514–3522. ISSN: 1521-4141. DOI: 10.1002/(SICI)1521-4141(199811)28:11<3514::AID-IMMU3514>3.0.CO;2-T.

Jones, Simon A., Jürgen Scheller and Stefan Rose-John (2011). 'Therapeutic strategies for the clinical blockade of IL-6/gp130 signaling'. In: *Journal of Clinical Investigation* 121.9, pp. 3375–3383. DOI: 10.1172/jci57158.

Jonge, Milly E. de, Alwin DR. Huitema, Jan HM. Schellens, Sjoerd Rodenhuis and Jos H. Beijnen (2005). 'Population pharmacokinetics of orally administered paclitaxel formulated in Cremophor EL'. In: *British Journal of Clinical Pharmacology* 59.3, pp. 325–334. DOI: 10.1111/j.1365-2125.2004.02325.x.

Jonsson, E. N. and M. O. Karlsson (1999). 'Xpose–an S-PLUS based population pharmacokinetic/pharmacodynamic model building aid for NONMEM'. In: *Comput Methods Programs Biomed* 58.1, pp. 51–64.

Jostock, Thomas, Jürgen Müllberg, Suat Özbek, Raja Atreya, Guido Blinn, Nicole Voltz, Martina Fischer, Markus F. Neurath and Stefan Rose-John (2001). 'Soluble gp130 is the natural inhibitor of soluble interleukin-6 receptor transsignaling responses'. In: *European Journal of Biochemistry* 268.1, pp. 160–167. DOI: 10.1046/j.1432-1327.2001.01867.x.

Jürchott, Karsten et al. (2016). 'Highly Predictive Model for a Protective Immune Response to the A(H1N1)pdm2009 Influenza Strain after Seasonal Vaccination'. In: *PLOS ONE* 11.3. Ed. by Sang-Moo Kang, e0150812. DOI: 10.1371/journal.pone.0150812.

Kamath, Amrita V. et al. (2011). 'Preclinical pharmacokinetics of MFGR1877A, a human monoclonal antibody to FGFR3, and prediction of its efficacious clinical dose for the treatment of t(4;14)-positive multiple myeloma'. In: *Cancer Chemotherapy and Pharmacology* 69.4, pp. 1071–1078. DOI: 10.1007/s00280-011-1807-5.

Kamei, Hiroko, Robert C. Jackson, Daniella Zheleva and Fordyce A. Davidson (2010). 'An integrated pharmacokinetic–pharmacodynamic model for an Aurora kinase inhibitor'. In: *Journal of Pharmacokinetics and Pharmacodynamics* 37.4, pp. 407–434. DOI: 10.1007/s10928-010-9166-0.

Kamimura, D., K. Ishihara and T. Hirano (2003). 'IL-6 signal transduction and its physiological roles: the signal orchestration model'. In: *Reviews of Physiology, Biochemistry and Pharmacology*. Springer Science *plus* Business Media, pp. 1–38. DOI: 10.1007/s10254-003-0012-2.

Kan, M. K. and G. B. Hopkins (1979). 'Measurement of liver volume by emission computed tomography'. In: *J. Nucl. Med.* 20.6, pp. 514–520.

Kannan, Senthil, Noor Dawany, Raj Kurupati, Louise C. Showe and Hildegund C.J. Ertl (2016). 'Age-related changes in the transcriptome of antibody-secreting cells'. In: *Oncotarget* 7.12. DOI: 10.18632/oncotarget.7958.

Kennedy, J. and R. Eberhart (1995). 'Particle swarm optimization'. In: *Proceedings of ICNN'95 - International Conference on Neural Networks*. IEEE. DOI: 10.1109/icnn.1995.488968.

Keystone, E. C., F. C. Breedveld, D. van der Heijde, R. Landewe, S. Florentinus, U. Arulmani, S. Liu, H. Kupper and A. Kavanaugh (2014). 'Longterm effect of delaying combination therapy with tumor necrosis factor inhibitor in patients with aggressive early rheumatoid arthritis: 10-year efficacy and safety of adalimumab from the randomized controlled PREMIER trial with open-label extension'. In: *J. Rheumatol.* 41.1, pp. 5–14.

Khoufache, Khaled et al. (2012). 'PAR1 contributes to influenza A virus pathogenicity in mice'. In: *Journal of Clinical Investigation* 123.1, pp. 206–214. DOI: 10.1172/jci61667.

Kim, Hansoo, Lyle Gurrin, Zanfina Ademi and Danny Liew (2013). 'Overview of methods for comparing the efficacies of drugs in the absence of head-to-head clinical trial data'. In: *British Journal of Clinical Pharmacology* 77.1, pp. 116–121. DOI: 10.1111/bcp.12150.

Kim, Hongkyun, Teresa S. Hawley, Robert G. Hawley and Heinz Baumann (1998). 'Protein Tyrosine Phosphatase 2 (SHP-2) Moderates Signaling by gp130 but Is Not Required for the Induction of Acute-Phase Plasma Protein Genes in Hepatic Cells'. In: *Molecular and Cellular Biology* 18.3, pp. 1525–1533. DOI: 10.1128/mcb.18.3.1525.

Kim, Yae-Jean, Calman Prussin, Brian Martin, Melissa A. Law, Thomas P. Haverty, Thomas B. Nutman and Amy D. Klion (2004). 'Rebound eosinophilia after treatment of hypereosinophilic syndrome and eosinophilic gastroenteritis with monoclonal anti–IL-5 antibody SCH55700'. In: *Journal of Allergy and Clinical Immunology* 114.6, pp. 1449–1455. DOI: 10.1016/j.jaci.2004.08.027.

Kitano, H. (2002). 'Systems biology: a brief overview'. In: *Science* 295.5560, pp. 1662–1664. DOI: 10.1126/science.1069492.

— (Mar. 2007). 'A robustness-based approach to systems-oriented drug design'. In: *Nat Rev Drug Discov* 6.3, pp. 202–210.

Knight-Schrijver, V.R., V. Chelliah, L. Cucurull-Sanchez and N. Le Novère (2016). 'The promises of quantitative systems pharmacology modelling for drug development'. In: *Computational and Structural Biotechnology Journal* 14, pp. 363–370. DOI: 10.1016/j.csbj.2016.09.002.

Kohavi, Ron and George H. John (1997). 'Wrappers for feature subset selection'. In: *Artificial Intelligence* 97.1-2, pp. 273–324. DOI: 10.1016/s0004-3702(97)00043-x.

Kohno, Nobuoki, Akihito Yokoyama, Tetsu Oyama, Yutaka Hirasawa, Kunio Hiwada, Yasuaki Okuda and Kiyoshi Takasugi (1998). 'Soluble interleukin-6 receptor in rheumatoid arthritis'. In: *Japanese Journal of Rheumatology* 8.2, pp. 131–138. DOI: 10.1007/bf03041336.

Kokebie, Rediet, Rohit Aggarwal, Sukhwinderjit Lidder, Arnavaz A Hakimiyan, David C Rueger, Joel A Block and Susan Chubinskaya (2011). 'The role of synovial fluid markers of catabolism and anabolism in osteoarthritis, rheumatoid arthritis and asymptomatic organ donors'. In: *Arthritis Res Ther* 13.2, R50. DOI: 10.1186/ar3293.

Kola, Ismail and John Landis (2004). 'Can the pharmaceutical industry reduce attrition rates?' In: *Nature reviews Drug discovery* 3.8, pp. 711–716. DOI: 10.1038/nrd1470.

Kolesar, J., R. C. Brundage, M. Pomplun, D. Alberti, K. Holen, A. Traynor, P. Ivy and G. Wilding (2011). 'Population pharmacokinetics of 3-aminopyridine-2-carboxaldehyde thiosemicarbazone (Triapine®) in cancer patients'. In: *Cancer Chemother. Pharmacol.* 67.2, pp. 393–400.

Koller-Strametz, J., R. Pacher, B. Frey, T. Kos, W. Woloszczuk and B. Stanek (1998). 'Circulating tumor necrosis factor-alpha levels in chronic heart failure: relation to its soluble receptor II, interleukin-6, and neurohumoral variables'. In: *J. Heart Lung Transplant.* 17.4, pp. 356–362.

Korcsmáros, Tamás, Máté S Szalay, Csaba Böde, István A Kovács and Péter Csermely (2007). 'How to design multi-target drugs'. In: *Expert Opinion on Drug Discovery* 2.6, pp. 799–808. DOI: 10.1517/17460441.2.6.799.

Kosinsky, Yuri, Simon J. Dovedi, Kirill Peskov, Veronika Voronova, Lulu Chu, Helen Tomkinson, Nidal Al-Huniti, Donald R. Stanski and Gabriel Helmlinger (2018). 'Radiation and PD-(L)1 treatment combinations: immune response and dose optimization via a predictive systems model'. In: *Journal for ImmunoTherapy of Cancer* 6.1. DOI: 10.1186/s40425-018-0327-9.

Kotake, Shigeru, Kazuto Sato, Kang Jung Kim, Naoyuki Takahashi, Nobuyuki Udagawa, Ichiro Nakamura, Akira Yamaguchi, Tadamitsu Kishimoto, Tatsuo Suda and Sadao Kashiwazaki (1996). 'Interleukin-6 and soluble interleukin-6 receptors in the synovial fluids from rheumatoid arthritis patients are responsible for osteoclast-like cell formation'. In: *Journal of Bone and Mineral Research* 11.1, pp. 88–95. ISSN: 1523-4681. DOI: 10.1002/jbmr.5650110113.

Krause, A., N. Scaletta, J.-D. Ji and L. B. Ivashkiv (2002). 'Rheumatoid Arthritis Synoviocyte Survival Is Dependent on Stat3'. In: *The Journal of Immunology* 169.11, pp. 6610–6616. DOI: 10.4049/jimmunol.169.11.6610.

Krippendorff, Ben-Fillippo, Katharina Kuester, Charlotte Kloft and Wilhelm Huisinga (2009). 'Nonlinear pharmacokinetics of therapeutic proteins resulting from receptor mediated endocytosis'. In: *Journal of Pharmacokinetics and Pharmacodynamics* 36.3, pp. 239–260. DOI: 10.1007/s10928-009-9120-1.

Kruskal, William H. and W. Allen Wallis (1952). 'Use of Ranks in One-Criterion Variance Analysis'. In: *Journal of the American Statistical Association* 47.260, pp. 583–621. DOI: 10.1080/01621459.1952.10483441.

Kuny, Chad V. and Christopher S. Sullivan (2016). 'Virus–Host Interactions and the ARTD/PARP Family of Enzymes'. In: *PLOS Pathogens* 12.3. Ed. by Katherine R. Spindler, e1005453. DOI: 10.1371/journal.ppat.1005453.

Kuta, A. E. (1986). 'C-reactive protein is produced by a small number of normal human peripheral blood lymphocytes'. In: *Journal of Experimental Medicine* 164.1, pp. 321–326. DOI: 10.1084/jem.164.1.321.

Kwon, A.-Hon, Yoichi Matsui, Sang Kil Ha-Kawa and Yasuo Kamiyama (2001). 'Functional hepatic volume measured by technetium-99m-galactosyl-human serum albumin liver scintigraphy: com-

parison between hepatocyte volume and liver volume by computed tomography'. In: *The American Journal of Gastroenterology* 96.2, pp. 541–546. DOI: 10.1111/j.1572-0241.2001.03556.x.

Lagathu, Claire, Jean-Philippe Bastard, Martine Auclair, Mustapha Maachi, Jacqueline Capeau and Martine Caron (2003). 'Chronic interleukin-6 (IL-6) treatment increased IL-6 secretion and induced insulin resistance in adipocyte: prevention by rosiglitazone'. In: *Biochemical and Biophysical Research Communications* 311.2, pp. 372–379. DOI: 10.1016/j.bbrc.2003.10.013.

Lally, Frank, Emily Smith, Andrew Filer, Michael A. Stone, John S. Shaw, Gerard B. Nash, Christopher D. Buckley and G. Ed Rainger (2005). 'A novel mechanism of neutrophil recruitment in a co-culture model of the rheumatoid synovium'. In: *Arthritis Rheum* 52.11, pp. 3460–3469. DOI: 10.1002/art.21394.

Lander, Eric S. et al. (2001). 'Initial sequencing and analysis of the human genome'. In: *Nature* 409.6822, pp. 860–921. DOI: 10.1038/35057062.

Lauffenburger, D.A. (2008). In: Quantitative and Systems Pharmacology Workshop Executive Summary, September 25-26. URL: https://www.nigms.nih.gov/News/reports/archivedreports2009-2007/Pages/PharmacologyConference20080925.aspx.

Lazebnik, Y. (2002). 'Can a biologist fix a radio?–Or, what I learned while studying apoptosis'. In: *Cancer Cell* 2.3, pp. 179–182.

Le Goff, B., F. Blanchard, J. M. Berthelot, D. Heymann and Y. Maugars (2010). 'Role for interleukin-6 in structural joint damage and systemic bone loss in rheumatoid arthritis'. In: *Joint Bone Spine* 77.3, pp. 201–205.

Le Novère, Nicolas et al. (2005). 'Minimum information requested in the annotation of biochemical models (MIRIAM)'. In: *Nature Biotechnology* 23.12, pp. 1509–1515. DOI: 10.1038/nbt1156.

Le Novère, Nicolas, Benjamin Bornstein, Alexander Broicher, Mélanie Courtot, Marco Donizelli et al. (2006). 'BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems'. In: *Nucleic Acids Research* 34.Database issue, pp. D689–D691. DOI: 10.1093/nar/gkj092.

Le Novère, Nicolas et al. (2009). 'The Systems Biology Graphical Notation'. In: *Nature Biotechnology* 27.8, pp. 735–741. DOI: 10.1038/nbt.1558.

Le, Quoc and Tomas Mikolov (2014). 'Distributed Representations of Sentences and Documents'. In: *Proceedings of the 31st International Conference on Machine Learning*. Ed. by Eric P. Xing and Tony Jebara. Vol. 32. Proceedings of Machine Learning Research 2. Bejing, China: PMLR, pp. 1188–1196.

Lee, Ho (2014). 'Genetically Engineered Mouse Models for Drug Development and Preclinical Trials'. In: *Biomolecules & Therapeutics* 22.4, pp. 267–274. DOI: 10.4062/biomolther.2014.074.

Lee, J. Y., C. E. Garnett, J. V. Gobburu, V. A. Bhattaram, S. Brar et al. (2011). 'Impact of pharmacometric analyses on new drug approval and labelling decisions: a review of 198 submissions between 2000 and 2008'. In: *Clin Pharmacokinet* 50.10, pp. 627–635. DOI: 10.2165/11593210-000000000-00000.

Lehmann, U. (2002). 'SHP2 and SOCS3 Contribute to Tyr-759-dependent Attenuation of Interleukin-6 Signaling through gp130'. In: *Journal of Biological Chemistry* 278.1, pp. 661–671. DOI: 10.1074/jbc.m210552200.

Leil, Tarek A. and Richard Bertz (2014). 'Quantitative Systems Pharmacology can reduce attrition and improve productivity in pharmaceutical research and development'. In: *Frontiers in Pharmacology* 5. DOI: 10.3389/fphar.2014.00247.

Lemmers, A. et al. (2009). 'An inhibitor of interleukin-6 trans-signalling, sgp130, contributes to impaired acute phase response in human chronic liver disease'. In: *Clinical & Experimental Immunology* 156.3, pp. 518–527. DOI: 10.1111/j.1365-2249.2009.03916.x.

Levesque, H. and O. Lafont (2000). '[Aspirin throughout the ages: a historical review]'. In: *Rev Med Interne* 21 Suppl 1, 8s–17s.

Li, Yongtao, Fan Ming, Huimin Huang, Kelei Guo, Huanchun Chen, Meilin Jin and Hongbo Zhou (2017). 'Proteome Response of Chicken Embryo Fibroblast Cells to Recombinant H5N1 Avian Influenza Viruses with Different Neuraminidase Stalk Lengths'. In: *Scientific Reports* 7, p. 40698. DOI: 10.1038/srep40698.

Liaw, Andy and Matthew Wiener (2002). 'Classification and Regression by randomForest'. In: *R News* 2.3, pp. 18–22. URL: http://CRAN.R-project.org/doc/Rnews/.

Lindqvist, Daniel, Shorena Janelidze, Peter Hagell, Sophie Erhardt, Martin Samuelsson, Lennart Minthon, Oskar Hansson, Maria Björkqvist, Lil Träskman-Bendz and Lena Brundin (2009). 'Interleukin-6 Is Elevated in the Cerebrospinal Fluid of Suicide Attempters and Related to Symptom Severity'. In: *Biological Psychiatry* 66.3, pp. 287–292. DOI: 10.1016/j.biopsych.2009.01.030.

Lipscomb, C. E. (2000). 'Medical Subject Headings (MeSH)'. In: *Bull Med Libr Assoc* 88.3, pp. 265–266.

Lloret-Villas, A, TM Varusai, N Juty, C Laibe, N Le NovÈre, H Hermjakob and V Chelliah (2017). 'The Impact of Mathematical Modeling in Understanding the Mechanisms Underlying Neurodegeneration: Evolving Dimensions and Future Directions'. In: *CPT: Pharmacometrics & Systems Pharmacology* 6.2, pp. 73–86. DOI: 10.1002/psp4.12155.

Lombardino, Joseph G. and John A. Lowe (2004). 'A guide to drug discovery: The role of the medicinal chemist in drug discovery —

then and now'. In: *Nature Reviews Drug Discovery* 3.10, pp. 853–862. DOI: 10.1038/nrd1523.

Lovato, P., C. Brender, J. Agnholt, J. Kelsen, K. Kaltoft, A. Svejgaard, K. W. Eriksen, A. Woetmann and N. Odum (2003). 'Constitutive STAT3 Activation in Intestinal T Cells from Patients with Crohn's Disease'. In: *Journal of Biological Chemistry* 278.19, pp. 16777–16781. DOI: 10.1074/jbc.m207999200.

Love, Michael I, Wolfgang Huber and Simon Anders (2014). 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2'. In: *Genome Biology* 15.12. DOI: 10.1186/s13059-014-0550-8.

Lu, James, Katrin Hübner, M. Nazeem Nanjee, Eliot A. Brinton and Norman A. Mazer (2014). 'An In-Silico Model of Lipoprotein Metabolism and Kinetics for the Evaluation of Targets and Biomarkers in the Reverse Cholesterol Transport Pathway'. In: *PLoS Computational Biology* 10.3. Ed. by Daniel A. Beard, e1003509. DOI: 10.1371/journal.pcbi.1003509.

Lust, John A., Kathleen A. Donovan, Michael P. Kline, Philip R. Greipp, Robert A. Kyle and Nita J. Maihle (1992). 'Isolation of an mRNA encoding a soluble form of the human interleukin-6 receptor'. In: *Cytokine* 4.2, pp. 96–100. DOI: 10.1016/1043-4666(92)90043-q.

MIWA, MAKOTO, RUNE SÆTRE, JIN-DONG KIM and JUN'ICHI TSUJII (2010). 'EVENT EXTRACTION WITH COMPLEX EVENT CLASSIFICATION USING RICH FEATURES'. In: *Journal of Bioinformatics and Computational Biology* 08.01, pp. 131–146. DOI: 10.1142/s0219720010004586.

Macintyre, S. S., I. Kushner and D. Samols (1985). 'Secretion of C-reactive protein becomes more efficient during the course of the acute phase response'. In: *J. Biol. Chem.* 260.7, pp. 4169–4173.

Mager, D. E. and W. J. Jusko (2001). 'General pharmacokinetic model for drugs exhibiting target-mediated drug disposition'. In: *J Pharmacokinet Pharmacodyn* 28.6, pp. 507–532.

Mager, Donald E. and Holly H.C. Kimko (2016). *Systems Pharmacology and Pharmacodynamics (AAPS Advances in the Pharmaceutical Sciences Series)*. Springer. ISBN: 978-3-319-44534-2.

Mahale, Parag, Sheeba K. Thomas, Andreas Kyvernitakis and Harrys A. Torres (2015). 'Management of Multiple Myeloma Complicated by Hepatitis C Virus Reactivation: The Role of New Antiviral Therapy'. In: *Open Forum Infectious Diseases* 3.1, ofv211. DOI: 10.1093/ofid/ofv211.

Mangoni, A. A. and S. H. D. Jackson (2003). 'Age-related changes in pharmacokinetics and pharmacodynamics: basic principles and practical applications'. In: *British Journal of Clinical Pharmacology* 57.1, pp. 6–14. DOI: 10.1046/j.1365-2125.2003.02007.x.

Manicourt, Daniel-Henri, Rafak Triki, Kanji Fukuda, Jean-Pierre Devogelaer, Charles Nagant De Deuxchaisnes and Eugene J.-M. A.

Thonar (1993). 'Levels of circulating tumor necrosis factor α and interleukin-6 in patients with rheumatoid arthritis. relationship to serum levels of hyaluronan and antigenic keratan sulfate'. In: *Arthritis & Rheumatism* 36.4, pp. 490–499. DOI: 10.1002/art.1780360409.

Manning, Christopher D., Prabhakar Raghavan and Hinrich Schütze (2008). *Introduction to Information Retrieval*. Cambridge University Press. ISBN: 0521865719.

Mascio, M. Di et al. (2009). 'Noninvasive in vivo imaging of CD4 cells in simian-human immunodeficiency virus (SHIV)-infected nonhuman primates'. In: *Blood* 114.2, pp. 328–337. DOI: 10.1182/blood-2008-12-192203.

Matthews, Vance et al. (2003). 'Cellular Cholesterol Depletion Triggers Shedding of the Human Interleukin-6 Receptor by ADAM10 and ADAM17 (TACE)'. In: *Journal of Biological Chemistry* 278.40, pp. 38829–38839. DOI: 10.1074/jbc.m210584200.

McCallum, Andrew and Kamal Nigam (1998). 'A Comparison of Event Models for Naive Bayes Text Classification'. In: *Learning for Text Categorization: Papers from the 1998 AAAI Workshop*, pp. 41–48.

McInnes, Iain B. and Georg Schett (2007). 'Cytokines in the pathogenesis of rheumatoid arthritis'. In: *Nat Rev Immunol* 7.6, pp. 429–442. DOI: 10.1038/nri2094.

McLoughlin, R. M., B. J. Jenkins, D. Grail, A. S. Williams, C. A. Fielding, C. R. Parker, M. Ernst, N. Topley and S. A. Jones (2005). 'IL-6 trans-signaling via STAT3 directs T cell infiltration in acute inflammation'. In: *Proceedings of the National Academy of Sciences* 102.27, pp. 9589–9594. DOI: 10.1073/pnas.0501794102.

McNeil, Michael M. et al. (2016). 'Risk of anaphylaxis after vaccination in children and adults'. In: *Journal of Allergy and Clinical Immunology* 137.3, pp. 868–878. DOI: 10.1016/j.jaci.2015.07.048.

Mélet, J., D. Mulleman, P. Goupille, B. Ribourtout, H. Watier and G. Thibault (2013). 'Rituximab-Induced T Cell Depletion in Patients With Rheumatoid Arthritis: Association With Clinical Response'. In: *Arthritis & Rheumatism* 65.11, pp. 2783–2790. DOI: 10.1002/art.38107.

Memoli, Bruno et al. (2010). 'A translational approach to micro-inflammation in end-stage renal disease: molecular effects of low levels of interleukin-6'. In: *Clin. Sci.* 119.4, pp. 163–174. DOI: 10.1042/cs20090634.

Mihara, Masahiko, Keiko Kasutani, Makoto Okazaki, Akito Nakamura, Shigeto Kawai, Masamichi Sugimoto, Yoshihiro Matsumoto and Yoshiyuki Ohsugi (2005). 'Tocilizumab inhibits signal transduction mediated by both mIL-6R and sIL-6R, but not by the receptors of other members of IL-6 cytokine family'. In: *International Immunopharmacology* 5.12, pp. 1731–1740. DOI: 10.1016/j.intimp.2005.05.010.

Mihara, Masahiko, Yoshiyuki Ohsugi and Kishimoto (2011). 'Tocilizumab, a humanized anti-interleukin-6 receptor antibody, for treat-

ment of rheumatoid arthritis'. In: *Open Access Rheumatology: Research and Reviews*, p. 19. DOI: 10.2147/oarrr.s17118.

Mikitsh, J. L. and A. M. Chacko (2014). 'Pathways for small molecule delivery to the central nervous system across the blood-brain barrier'. In: *Perspect Medicin Chem* 6, pp. 11–24.

Milazzo, Laura, Mara Biasin, Nadia Gatti, Luca Piacentini, Fosca Niero, Barbara Zanone Poma, Massimo Galli, Mauro Moroni, Mario Clerici and Agostino Riva (2006). 'Thalidomide in the Treatment of Chronic Hepatitis C Unresponsive to Alfa-Interferon and Ribavirin'. In: *The American Journal of Gastroenterology* 101.2, pp. 399–402. DOI: 10.1111/j.1572-0241.2006.00350.x.

Minichiello, Emeline, Luca Semerano and Marie-Christophe Boissier (2016). 'Time trends in the incidence, prevalence, and severity of rheumatoid arthritis: A systematic literature review'. In: *Joint Bone Spine* 83.6, pp. 625–630. DOI: 10.1016/j.jbspin.2016.07.007.

Mitchell, Mark, Baurzhan Muftakhidinov, Tobias Winchen, Zbigniew Jędrzejewski-Szmek and The Gitter Badger (Aug. 2016). *engauge-digitizer: Support for smaller monitors*. DOI: 10.5281/zenodo.61108.

Monto, Arnold S., Suzanne E. Ohmit, Joshua G. Petrie, Emileigh Johnson, Rachel Truscon, Esther Teich, Judy Rotthoff, Matthew Boulton and John C. Victor (2009). 'Comparative Efficacy of Inactivated and Live Attenuated Influenza Vaccines'. In: *New England Journal of Medicine* 361.13, pp. 1260–1267. DOI: 10.1056/nejmoa0808652.

Moradi, Babak, Philipp Schnatzer, Sébastien Hagmann, Nils Rosshirt, Tobias Gotterbarm, Jan Kretzer, Marc Thomsen, Hanns-Martin Lorenz, Felix Zeifang and Theresa Tretter (2014). 'CD4+CD25+/highCD127low/-regulatory T cells are enriched in rheumatoid arthritis and osteoarthritis joints—analysis of frequency and phenotype in synovial membrane, synovial fluid and peripheral blood'. In: *Arthritis Res Ther* 16.2, R97. DOI: 10.1186/ar4545.

Morcos, Peter N., Xiaoping Zhang, Christine McIntyre, Beate Bittner, Lucy Rowell and Zubair Hussain (2013). 'Pharmacokinetics and pharmacodynamics of single subcutaneous doses of tocilizumab administered with or without rHuPH20'. In: *Int. Journal of Clinical Pharmacology and Therapeutics* 51.07, pp. 537–548. DOI: 10.5414/cp201847.

Morieri, Mario Luca, Angelina Passaro and Giovanni Zuliani (2017). 'Interleukin-6 "Trans-Signaling" and Ischemic Vascular Disease: The Important Role of Soluble gp130'. In: *Mediators of Inflammation* 2017, pp. 1–6. DOI: 10.1155/2017/1396398.

Morrow, J. K., L. Tian and S. Zhang (2010). 'Molecular networks in drug discovery'. In: *Crit Rev Biomed Eng* 38.2, pp. 143–156.

Moya, C., J. Hahn, P. Cheng, A. Jayaraman and Z. Huang (2011). 'Investigation of IL-6 and IL-10 signalling via mathematical model-

ling'. In: *IET Systems Biology* 5.1, pp. 15–26. DOI: `10.1049/iet-syb.2009.0060`.

Mullberg, J., E. Dittrich, L. Graeve, C. Gerhartz, K. Yasukawa, T. Taga, T. Kishimoto, P. C. Heinrich and S. Rose-John (1993). 'Differential shedding of the two subunits of the interleukin-6 receptor'. In: *FEBS Lett.* 332.1-2, pp. 174–178.

Musana, K. A., S. H. Yale, J. J. Mazza and K. D. Reed (2004). 'Practical considerations to influenza vaccination'. In: *Clin Med Res* 2.4, pp. 256–259.

Musante, CJ, S Ramanujan, BJ Schmidt, OG Ghobrial, J Lu and AC Heatherington (2016). 'Quantitative Systems Pharmacology: A Case for Disease Models'. In: *Clinical Pharmacology & Therapeutics* 101.1, pp. 24–27. DOI: `10.1002/cpt.528`.

Nakaya, Helder I et al. (2011). 'Systems biology of vaccination for seasonal influenza in humans'. In: *Nature Immunology* 12.8, pp. 786–795. DOI: `10.1038/ni.2067`.

Narazaki, M., K. Yasukawa, T. Saito, Y. Ohsugi, H. Fukui, Y. Koishihara, G. D. Yancopoulos, T. Taga and T. Kishimoto (1993). 'Soluble forms of the interleukin-6 signal-transducing receptor component gp130 in human serum possessing a potential to inhibit signals through membrane-anchored gp130'. In: *Blood* 82.4, pp. 1120–1126.

Nayak, Jennifer L., Theresa F. Fitzgerald, Katherine A. Richards, Hongmei Yang, John J. Treanor and Andrea J. Sant (2012). 'CD4+ T-Cell Expansion Predicts Neutralizing Antibody Responses to Monovalent, Inactivated 2009 Pandemic Influenza A(H1N1) Virus Subtype H1N1 Vaccine'. In: *The Journal of Infectious Diseases* 207.2, pp. 297–305. DOI: `10.1093/infdis/jis684`.

Nesbitt, J. E. and G. M. Fuller (1992). 'Dynamics of interleukin-6 internalization and degradation in rat hepatocytes'. In: *J. Biol. Chem.* 267.9, pp. 5739–5742.

Niemand, C., A. Nimmesgern, S. Haan, P. Fischer, F. Schaper, R. Rossaint, P. C. Heinrich and G. Muller-Newen (2003). 'Activation of STAT3 by IL-6 and IL-10 in primary human macrophages is differentially modulated by suppressor of cytokine signaling 3'. In: *J. Immunol.* 170.6, pp. 3263–3272.

Nijsen, Marjoleen J.M.A. et al. (2018). 'Preclinical QSP Modeling in the Pharmaceutical Industry: An IQ Consortium Survey Examining the Current Landscape'. In: *CPT: Pharmacometrics & Systems Pharmacology*. DOI: `10.1002/psp4.12282`.

Nishimoto, N. et al. (2003). 'Toxicity, pharmacokinetics, and dosefinding study of repetitive treatment with the humanized antiinterleukin 6 receptor antibody MRA in rheumatoid arthritis. Phase I/II clinical study'. In: *J. Rheumatol.* 30.7, pp. 1426–1435.

Nishimoto, N., K. Terao, T. Mima, H. Nakahara, N. Takagi and T. Kakehi (2008). 'Mechanisms and pathologic significances in in-

crease in serum interleukin-6 (IL-6) and soluble IL-6 receptor after administration of an anti-IL-6 receptor antibody, tocilizumab, in patients with rheumatoid arthritis and Castleman disease'. In: *Blood* 112.10, pp. 3959–3964. DOI: 10.1182/blood-2008-05-155846.

Norris, Callie A. et al. (2014). 'Synthesis of IL-6 by Hepatocytes Is a Normal Response to Common Hepatic Stimuli'. In: *PLoS ONE* 9.4. Ed. by Laurent Rénia, e96053. DOI: 10.1371/journal.pone.0096053.

Nowell, M. A., P. J. Richards, S. Horiuchi, N. Yamamoto, S. Rose-John, N. Topley, A. S. Williams and S. A. Jones (2003). 'Soluble IL-6 receptor governs IL-6 activity in experimental arthritis: blockade of arthritis severity by soluble glycoprotein 130'. In: *J. Immunol.* 171.6, pp. 3202–3209.

Nygaard, Vegard, Einar Andreas Rødland and Eivind Hovig (2015). 'Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses'. In: *Biostatistics*, kxv027. DOI: 10.1093/biostatistics/kxv027.

O'Mara-Eves, Alison, James Thomas, John McNaught, Makoto Miwa and Sophia Ananiadou (2015). 'Using text mining for study identification in systematic reviews: a systematic review of current approaches'. In: *Systematic Reviews* 4.1. DOI: 10.1186/2046-4053-4-5.

Ogata, Atsushi et al. (2014). 'Phase III Study of the Efficacy and Safety of Subcutaneous Versus Intravenous Tocilizumab Monotherapy in Patients With Rheumatoid Arthritis'. In: *Arthritis Care & Research* 66.3, pp. 344–354. DOI: 10.1002/acr.22110.

Ohta, Shuji, Tomomi Tsuru, Kimio Terao, Seiji Mogi, Midori Suzaki, Eisuke Shono, Yoshimasa Ishida, Eriko Tarumi and Masato Imai (2013). 'Mechanism-based approach using a biomarker response to evaluate tocilizumab subcutaneous injection in patients with rheumatoid arthritis with an inadequate response to synthetic DMARDs (MATSURI study)'. In: *The Journal of Clinical Pharmacology* 54.1, pp. 109–119. DOI: 10.1002/jcph.185.

Okamoto, Hideyuki, Masahiro Yamamura, Yoshitaka Morita, Seishi Harada, Hirofumi Makino and Zensuke Ota (1997). 'The synovial expression and serum levels of interleukin-6, interleukin-11, leukemia inhibitory factor, and oncostatin m in rheumatoid arthritis'. In: *Arthritis & Rheumatism* 40.6, pp. 1096–1105. DOI: 10.1002/art.1780400614.

Oprisa, Ciprian, George Cabau and Adrian Colesa (2013). 'From Plagiarism to Malware Detection'. In: *2013 15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*. IEEE. DOI: 10.1109/synasc.2013.37.

Osterholm, Michael T, Nicholas S Kelley, Alfred Sommer and Edward A Belongia (2012). 'Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis'. In: *The Lancet Infectious Diseases* 12.1, pp. 36–44. DOI: 10.1016/s1473-3099(11)70295-x.

Pagani, Giuliano Andrea and Marco Aiello (2013). 'The Power Grid as a complex network: A survey'. In: *Physica A: Statistical Mechanics and its Applications* 392.11, pp. 2688–2700. DOI: 10.1016/j.physa.2013.01.023.

Page, K. R., E. Mezzalana, A. J. MacDonald, S. Zamuner, G. De Nicolao and A. van Maurik (2015). 'Temporal Pharmacokinetic/Pharmacodynamic Interaction between Human CD3 Antigen-Targeted Monoclonal Antibody Otelixizumab and CD3 Binding and Expression in Human Peripheral Blood Mononuclear Cell Static Culture'. In: *Journal of Pharmacology and Experimental Therapeutics* 355.2, pp. 199–205. DOI: 10.1124/jpet.115.224899.

Palache, Abraham, Valerie Oriol-Mathieu, Atika Abelin and Tamara Music (2014). 'Seasonal influenza vaccine dose distribution in 157 countries (2004–2011)'. In: *Vaccine* 32.48, pp. 6369–6376. DOI: 10.1016/j.vaccine.2014.07.012.

Paul, William E. (2008). *Fundamental Immunology*. , pp 142, ISBN: 0781765196. Lippincott Williams & Wilkins. ISBN: 0781765196.

Pepys, Mark B. and Gideon M. Hirschfield (2003). 'C-reactive protein: a critical update'. In: *Journal of Clinical Investigation* 111.12, pp. 1805–1812. DOI: 10.1172/jci200318921.

Pérez-Ruixo, Juan José et al. (2009). 'Pharmacokinetics and Pharmacodynamics of the Erythropoietin Mimetibody™ Construct CNTO 528 in Healthy Subjects'. In: *Clinical Pharmacokinetics* 48.9, pp. 601–613. DOI: 10.2165/11317190-000000000-00000.

Petrie, Joshua G., Suzanne E. Ohmit, Emileigh Johnson, Rachel T. Cross and Arnold S. Monto (2011). 'Efficacy Studies of Influenza Vaccines: Effect of End Points Used and Characteristics of Vaccine Failures'. In: *The Journal of Infectious Diseases* 203.9, pp. 1309–1315. DOI: 10.1093/infdis/jir015.

Petzold, Linda (1983). 'Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations'. In: *SIAM Journal on Scientific and Statistical Computing* 4.1, pp. 136–148. DOI: 10.1137/0904010.

Pichardo-Almarza, Cesar and Vanessa Diaz-Zuccarini (2017a). 'From PK/PD to QSP: Understanding the Dynamic Effect of Cholesterol-Lowering Drugs on Atherosclerosis Progression and Stratified Medicine'. In: *Current Pharmaceutical Design* 22.46, pp. 6903–6910. DOI: 10.2174/1381612822666160905095402.

— (2017b). 'Understanding the Effect of Statins and Patient Adherence in Atherosclerosis via a Quantitative Systems Pharmacology

Model Using a Novel, Hybrid, and Multi-Scale Approach'. In: *Frontiers in Pharmacology* 8. DOI: 10.3389/fphar.2017.00635.

Pignatti, P. et al. (2003). 'High circulating levels of biologically inactive IL-6/SIL-6 receptor complexes in systemic juvenile idiopathic arthritis: evidence for serum factors interfering with the binding to gp130'. In: *Clinical and Experimental Immunology* 131.2, pp. 355–363. DOI: 10.1046/j.1365-2249.2003.02052.x.

Pop, L. A., J. F. van den Broek, A. G. Visser and A. J. van der Kogel (1996). 'Constraints in the use of repair half times and mathematical modelling for the clinical application of HDR and PDR treatment schedules as an alternative for LDR brachytherapy'. In: *Radiother Oncol* 38.2, pp. 153–162.

Porter, M.F. (1980). 'An algorithm for suffix stripping'. In: *Program* 14.3, pp. 130–137. DOI: 10.1108/eb046814.

Qi, Yun feng et al. (2013). 'Elucidating the crosstalk mechanism between IFN-gamma and IL-6 via mathematical modelling'. In: *BMC Bioinformatics* 14.1, p. 41. DOI: 10.1186/1471-2105-14-41.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

R.W, Moskowitz, Altman R.D., Hochberg M.C. et al. (2007). *Osteoarthritis: diagnosis and medical/surgical management. 4th ed.* Philadelphia (PA): Lippincott Williams & Wilkins., p. 204.

Raia, V. et al. (2011). 'Dynamic mathematical modeling of IL13-induced signaling in Hodgkin and primary mediastinal B-cell lymphoma allows prediction of therapeutic targets'. In: *Cancer Res.* 71.3, pp. 693–704.

Ramji, D.P., A. Vitelli, F. Tronche, R. Cortese and G. Ciliberto (1993). 'The two C/EBP isoforms, IL6DBP/NFIL6 and CEBP6δ/NFIL63, are induced by IL6β to promote acute phase gene transcription via different mechanisms'. In: *Nucl Acids Res* 21.2, pp. 289–294. DOI: 10.1093/nar/21.2.289.

Rasch, Elizabeth K., Rosemarie Hirsch, Ryne Paulose-Ram and Marc C. Hochberg (2003). 'Prevalence of rheumatoid arthritis in persons 60 years of age and older in the United States: Effect of different methods of case classification'. In: *Arthritis & Rheumatism* 48.4, pp. 917–926. DOI: 10.1002/art.10897.

Rau, R. (2010). 'Efficacy of methotrexate in comparison to biologics in rheumatoid arthritis'. In: *Clin. Exp. Rheumatol.* 28.5 Suppl 61, pp. 58–64.

Reber, Adrian and Jacqueline Katz (2013). 'Immunological assessment of influenza vaccines and immune correlates of protection'. In: *Expert Review of Vaccines* 12.5, pp. 519–536. DOI: 10.1586/erv.13.35.

Reddy, V Pilla, K J Petersson, A A Suleiman, A Vermeulen, J H Proost and L E Friberg (2012). 'Pharmacokinetic–Pharmacodynamic Mod-

eling of Severity Levels of Extrapyramidal Side Effects With Markov Elements'. In: *CPT: Pharmacometrics & Systems Pharmacology* 1.9, e1. DOI: `10.1038/psp.2012.9`.

Revell, Andrew D, Luminiţa Ene, Dan Duiculescu, Dechao Wang, Mike Youle, Anton Pozniak, Julio Montaner and Brendan A Larder (2012). 'The use of computational models to predict response to HIV therapy for clinical cases in Romania'. In: *GERMS* 2.1, pp. 6–11. DOI: `10.11599/germs.2012.1007`.

Reyes-Aldasoro, Constantino Carlos (2017). 'The proportion of cancer-related entries in PubMed has increased considerably; is cancer truly "The Emperor of All Maladies"?' In: *PLOS ONE* 12.3. Ed. by Giuseppe Novelli, e0173671. DOI: `10.1371/journal.pone.0173671`.

Rhee, Frits van et al. (2010). 'Siltuximab, a Novel Anti–Interleukin-6 Monoclonal Antibody, for Castleman's Disease'. In: *Journal of Clinical Oncology* 28.23, pp. 3701–3708. DOI: `10.1200/jco.2009.27.2377`.

Richards, Peter J. et al. (2006). 'Functional characterization of a soluble gp130 isoform and its therapeutic capacity in an experimental model of inflammatory arthritis'. In: *Arthritis Rheum* 54.5, pp. 1662–1672. DOI: `10.1002/art.21818`.

Rijsbergen, C. J. Van (1979). *Information Retrieval*. 2nd. Newton, MA, USA: Butterworth-Heinemann. ISBN: 0408709294.

Rippe, B. and B. Haraldsson (1994). 'Transport of macromolecules across microvascular walls: the two-pore theory'. In: *Physiol. Rev.* 74.1, pp. 163–219.

Ritchie, Michael, Lioudmila Tchistiakova and Nathan Scott (2013). 'Implications of receptor-mediated endocytosis and intracellular trafficking dynamics in the development of antibody drug conjugates'. In: *mAbs* 5.1, pp. 13–21. DOI: `10.4161/mabs.22854`.

RoActemra: EPAR - Product Information. European public assessment report (EPAR), Annex I, Summary Of Product Characteristics, Committee for Medicinal Products for Human Use (CHMP), EMEA/H/C/000955, 29 September 2017.

Robak, T., A. Gladalska, H. Stepień and E. Robak (1998). 'Serum levels of interleukin-6 type cytokines and soluble interleukin-6 receptor in patients with rheumatoid arthritis'. In: *Mediators of Inflammation* 7.5, pp. 347–353. DOI: `10.1080/09629359890875`.

Robertson, Stephen (2004). 'Understanding inverse document frequency: On theoretical arguments for IDF'. In: *Journal of Documentation* 60, p. 2004.

Rose-John, Stefan (2012). 'IL-6 Trans-Signaling via the Soluble IL-6 Receptor: Importance for the Pro-Inflammatory Activities of IL-6'. In: *International Journal of Biological Sciences* 8.9, pp. 1237–1247. DOI: `10.7150/ijbs.4989`.

Rosenbaum, J T, R Cugnini, D C Tara, S Hefeneider and J C Ansel (1992). 'Production and modulation of interleukin 6 synthesis by synoviocytes derived from patients with arthritic disease.' In: *Annals of the Rheumatic Diseases* 51.2, pp. 198–202. DOI: `10.1136/ard.51.2.198`.

Sadreev, Ildar I., Michael Z. Q. Chen, Gavin I. Welsh, Yoshinori Umezawa, Nikolay V. Kotov and Najl V. Valeyev (2014). 'A Systems Model of Phosphorylation for Inflammatory Signaling Events'. In: *PLoS ONE* 9.10. Ed. by Gautam Sethi, e110913. DOI: `10.1371/journal.pone.0110913`.

Sams-Dodd, F (2005). 'Target-based drug discovery: is something wrong?' In: *Drug Discovery Today* 10.2, pp. 139–147. DOI: `10.1016/s1359-6446(04)03316-1`.

Sánchez-Maroño, Noelia, Amparo Alonso-Betanzos and María Tombilla-Sanromán (2007). 'Filter Methods for Feature Selection – A Comparative Study'. In: *Intelligent Data Engineering and Automated Learning - IDEAL 2007*. Ed. by Hujun Yin, Peter Tino, Emilio Corchado, Will Byrne and Xin Yao. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 178–187. ISBN: 978-3-540-77226-2.

Santos, Gaël Dos, Elisabeth Neumeier and Rafik Bekkat-Berkani (2015). 'Influenza: Can we cope better with the unpredictable?' In: *Human Vaccines & Immunotherapeutics* 12.3, pp. 699–708. DOI: `10.1080/21645515.2015.1086047`.

Saric, J., L. J. Jensen, R. Ouzounova, I. Rojas and P. Bork (2005). 'Extraction of regulatory gene/protein networks from Medline'. In: *Bioinformatics* 22.6, pp. 645–650. DOI: `10.1093/bioinformatics/bti597`.

Scheifele, David W., Bernard Duval, Margaret L. Russell, Richard Warrington, Gaston DeSerres, Danuta M. Skowronski, Marc Dionne, James Kellner, Dele Davies and Judy MacDonald (2003). 'Ocular and Respiratory Symptoms Attributable to Inactivated Split Influenza Vaccine: Evidence from a Controlled Trial Involving Adults'. In: *Clinical Infectious Diseases* 36.7, pp. 850–857. DOI: `10.1086/368189`.

Scheller, Jürgen, Athena Chalaris, Dirk Schmidt-Arras and Stefan Rose-John (2011). 'The pro- and anti-inflammatory properties of the cytokine interleukin-6'. In: *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1813.5, pp. 878–888. DOI: `10.1016/j.bbamcr.2011.01.034`.

Scherrmann, J. M. (2002). 'Drug delivery to brain via the blood-brain barrier'. In: *Vascul. Pharmacol.* 38.6, pp. 349–354.

Schmidt-Arras, Dirk and Stefan Rose-John (2016). 'IL-6 pathway in the liver: From physiopathology to therapy'. In: *Journal of Hepatology* 64.6, pp. 1403–1415. DOI: `10.1016/j.jhep.2016.02.004`.

Schoeberl, B. et al. (2009). 'Therapeutically Targeting ErbB3: A Key Node in Ligand-Induced Activation of the ErbB Receptor-PI3K

Axis'. In: *Science Signaling* 2.77, ra31–ra31. DOI: `10.1126/scisignal.2000352`.

Schroers, A. (2005). 'Dynamics of the gp130 cytokine complex: A model for assembly on the cellular membrane'. In: *Protein Science* 14.3, pp. 783–790. DOI: `10.1110/ps.041117105`.

Schuck, Edgar et al. (2015). 'Preclinical Pharmacokinetic/Pharmacodynamic Modeling and Simulation in the Pharmaceutical Industry: An IQ Consortium Survey Examining the Current Landscape'. In: *The AAPS Journal* 17.2, pp. 462–473. DOI: `10.1208/s12248-014-9716-2`.

Schumacher, Neele et al. (2015). 'Shedding of Endogenous Interleukin-6 Receptor (IL-6R) Is Governed by A Disintegrin and Metalloproteinase (ADAM) Proteases while a Full-length IL-6R Isoform Localizes to Circulating Microvesicles'. In: *Journal of Biological Chemistry* 290.43, pp. 26059–26071. DOI: `10.1074/jbc.m115.649509`.

Schuster, Björn, Werner Meinert, Stefan Rose-John and Karl-Josef Kallen (2003). 'The human interleukin-6 (IL-6) receptor exists as a preformed dimer in the plasma membrane'. In: *FEBS Letters* 538.1-3, pp. 113–116. DOI: `10.1016/s0014-5793(03)00154-6`.

Seidman, Jessica C., Stephanie A. Richard, Cécile Viboud and Mark A. Miller (2012). 'Quantitative review of antibody response to inactivated seasonal influenza vaccines'. In: *Influenza and Other Respiratory Viruses* 6.1, pp. 52–62. DOI: `10.1111/j.1750-2659.2011.00268.x`.

Sernadela, P. and J. L. Oliveira (2017). 'A semantic-based workflow for biomedical literature annotation'. In: *Database (Oxford)* 2017.

Shadick, Nancy A. (2006). 'C-Reactive Protein in the Prediction of Rheumatoid Arthritis in Women'. In: *Arch Intern Med* 166.22, p. 2490. DOI: `10.1001/archinte.166.22.2490`.

Shanks, Niall, Ray Greek and Jean Greek (2009). 'Are animal models predictive for humans?' In: *Philosophy, Ethics, and Humanities in Medicine* 4.1, p. 2. DOI: `10.1186/1747-5341-4-2`.

Sharma, J., H. Lv and J. M. Gallo (2013). 'Intratumoral Modeling of Gefitinib Pharmacokinetics and Pharmacodynamics in an Orthotopic Mouse Model of Glioblastoma'. In: *Cancer Research* 73.16, pp. 5242–5252. DOI: `10.1158/0008-5472.can-13-0690`.

Sheiner, L. B. and J.-L. Steimer (2000). 'Pharmacokinetic/Pharmacodynamic Modeling in Drug Development'. In: *Annual Review of Pharmacology and Toxicology* 40.1, pp. 67–95. DOI: `10.1146/annurev.pharmtox.40.1.67`.

Sheiner, L. B., D. R. Stanski, S. Vozeh, R. D. Miller and J. Ham (1979). 'Simultaneous modeling of pharmacokinetics and pharmacodynamics: application to d-tubocurarine'. In: *Clin. Pharmacol. Ther.* 25.3, pp. 358–371.

Shinde, C. G., M. P. Venkatesh, T. M. Kumar and H. G. Shivakumar (2014). 'Methotrexate: a gold standard for treatment of rheumatoid arthritis'. In: *J Pain Palliat Care Pharmacother* 28.4, pp. 351–358.

Shrivastava, Amit Kumar, Harsh Vardhan Singh, Arun Raizada and Sanjeev Kumar Singh (2015). 'C-reactive protein, inflammation and coronary heart disease'. In: *The Egyptian Heart Journal* 67.2, pp. 89–97. DOI: 10.1016/j.ehj.2014.11.005.

"Siebert, S., A. Tsoukas, J. Robertson and I. McInnes (2015). 'Cytokines as therapeutic targets in rheumatoid arthritis and other inflammatory diseases'. In: *Pharmacol. Rev.* 67.2, pp. 280–309. DOI: 10.1124/pr.114.009639.

Silman, Alan J and Jacqueline E Pearson (2002). 'Epidemiology and genetics of rheumatoid arthritis'. In: *Arthritis Research* 4.Suppl 3, S265. DOI: 10.1186/ar578. URL: https://doi.org/10.1186/ar578.

Simard, Carl, Marc Cloutier and Sonia Néron (2014). 'Rapid determination of il-6 specific activity by flow cytometry'. In: *Journal of Immunological Methods* 415, pp. 63 –65. ISSN: 0022-1759. DOI: 10.1016/j.jim.2014.09.005.

Simkin, Peter A (2014). 'The human knee: A window on the microvasculature'. In: *Tissue Barriers* 3.1-2, e970465. DOI: 10.4161/21688362.2014.970465.

Simkin, Peter A., John E. Bassett and Eun-Mi Koh (1995). 'Synovial perfusion in the human knee:A methodologic analysis'. In: *Seminars in Arthritis and Rheumatism* 25.1, pp. 56–66. DOI: 10.1016/s0049-0172(95)80018-2.

Singh, Abhay, Arul Jayaraman and Juergen Hahn (2006). 'Modeling regulatory mechanisms in IL-6 signal transduction in hepatocytes'. In: *Biotechnology and Bioengineering* 95.5, pp. 850–862. DOI: 10.1002/bit.21026.

Singh, I, EE Nagiec, JM Thompson, W Krzyzanski and P Singh (2015). 'A Systems Pharmacology Model of Erythropoiesis in Mice Induced by Small Molecule Inhibitor of Prolyl Hydroxylase Enzymes'. In: *CPT: Pharmacometrics & Systems Pharmacology* 4.2, pp. 106–115. DOI: 10.1002/psp4.12.

Singh, Vivek Kumar, Nisha Tiwari and Shekhar Garg (2011). 'Document Clustering Using K-Means, Heuristic K-Means and Fuzzy C-Means'. In: *2011 International Conference on Computational Intelligence and Communication Networks*. IEEE. DOI: 10.1109/cicn.2011.62.

Skowronski, Danuta M., Barbara Strauss, Gaston De Serres, Diane MacDonald, Stephen A. Marion, Monika Naus, David M. Patrick and Perry Kendall (2003). 'Oculo-respiratory Syndrome: A New Influenza Vaccine–Associated Adverse Event?' In: *Clinical Infectious Diseases* 36.6, pp. 705–713. DOI: 10.1086/367667.

Smith, L. P., F. T. Bergmann, D. Chandran and H. M. Sauro (2009). 'Antimony: a modular model definition language'. In: *Bioinformatics* 25.18, pp. 2452–2454. DOI: 10.1093/bioinformatics/btp401.

Smith, Wilson, C.H. Andrewes and P.P. Laidlaw (1933). 'A VIRUS OBTAINED FROM INFLUENZA PATIENTS'. In: *The Lancet* 222.5732, pp. 66–68. DOI: 10.1016/s0140-6736(00)78541-2.

Smolen, Josef S, Michael E Weinblatt, Shihong Sheng, Yanli Zhuang and Benjamin Hsu (2014). 'Sirukumab, a human anti-interleukin-6 monoclonal antibody: a randomised, 2-part (proof-of-concept and dose-finding), phase II study in patients with active rheumatoid arthritis despite methotrexate therapy'. In: *Annals of the Rheumatic Diseases* 73.9, pp. 1616–1625. DOI: 10.1136/annrheumdis-2013-205137.

Song, Yong-Ling and Su-Shing Chen (2009). 'Text mining biomedical literature for constructing gene regulatory networks'. In: *Interdisciplinary Sciences: Computational Life Sciences* 1.3, pp. 179–186. ISSN: 1867-1462. DOI: 10.1007/s12539-009-0028-7.

Sonne, O., O. Davidsen, B. K. M?ller and C. Munck Petersen (1990). 'Cellular targets and receptors for interleukin-6. I. In vivo and in vitro uptake of IL-6 in liver and hepatocytes'. In: *Eur. J. Clin. Invest.* 20.4, pp. 366–376.

Sorger, Peter K, Sandra RB Allerheiligen, Darrell R Abernethy, Russ B Altman, Kim LR Brouwer et al. (2011). 'Quantitative and systems pharmacology in the post-genomic era: new approaches to discovering drugs and understanding therapeutic mechanisms'. In: *An NIH white paper by the QSP workshop group*. NIH Bethesda, pp. 1–48.

Spano, Angelo, Loredana Postiglione, Paola Sabatini, I. Soriente, M.G. Sangiolo, V. Bruner and Raffaele Scarpa (2011). 'Circulating Levels of IL6, sIL6-r, sgp130 and Gamma-IFN in Patients with Hepatitic C Virus Related Arthritis (HCVrA) and Rheumatoid Arthritis (RA). [abstract].' In: *Arthritis & Rheumatology* 63.S10.

Srirangan, S. and E. H. Choy (2010). 'The role of Interleukin 6 in the pathophysiology of rheumatoid arthritis'. In: *Therapeutic Advances in Musculoskeletal Disease* 2.5, pp. 247–256. DOI: 10.1177/1759720x10378372.

Stahl, E. A. et al. (2010). 'Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci'. In: *Nat. Genet.* 42.6, pp. 508–514.

Standing, Joseph F. (2016). 'Understanding and applying pharmacometric modelling and simulation in clinical practice and research'. In: *British Journal of Clinical Pharmacology* 83.2, pp. 247–254. DOI: 10.1111/bcp.13119.

Starr, R. et al. (1997). 'A family of cytokine-inducible inhibitors of signalling'. In: *Nature* 387.6636, pp. 917–921. DOI: 10.1038/43206.

Stavrou, Spyridon and Susan R. Ross (2015). 'APOBEC3 Proteins in Viral Immunity'. In: *The Journal of Immunology* 195.10, pp. 4565–4570. DOI: 10.4049/jimmunol.1501504.

Stratton, Kathleen, Andrew Ford, Erin Rusch and Ellen Wright Clayton (2011). 'Committee to Review Adverse Effects of Vaccines, Institute of Medicine; Adverse Effects of Vaccines: Evidence and Causality'. In: ed. by Kathleen Stratton, Andrew Ford, Erin Rusch and Ellen Wright Clayton.

Sun, Yu et al. (2017). 'The effects of interleukin-6 neutralizing antibodies on symptoms of depressed mood and anhedonia in patients with rheumatoid arthritis and multicentric Castleman's disease'. In: *Brain, Behavior, and Immunity* 66, pp. 156–164. DOI: 10.1016/j.bbi.2017.06.014.

Swanson, D. R. (1990). 'Medical literature as a potential source of new knowledge'. In: *Bull Med Libr Assoc* 78.1, pp. 29–37.

Swat, M. J., S. Moodie, S. M. Wimalaratne, N. R. Kristensen, M. Lavielle et al. (2015). 'Pharmacometrics Markup Language (PharmML): Opening New Perspectives for Model Exchange in Drug Development'. In: *CPT Pharmacometrics Syst Pharmacol* 4.6, pp. 316–319. DOI: 10.1002/psp4.57.

Szepietowski, Jacek C. et al. (2013). 'Phase 1, randomized, double-blind, placebo-controlled, multiple intravenous, dose-ascending study of sirukumab in cutaneous or systemic lupus erythematosus'. In: *Arthritis & Rheumatism*, n/a–n/a. DOI: 10.1002/art.38091.

Tabish, S. A. (2007). 'Is Diabetes Becoming the Biggest Epidemic of the Twenty-first Century?' In: *Int J Health Sci (Qassim)* 1.2, pp. V–VIII.

Tabrizi, Mohammad A., Chih-Ming L. Tseng and Lorin K. Roskos (2006). 'Elimination mechanisms of therapeutic monoclonal antibodies'. In: *Drug Discovery Today* 11.1-2, pp. 81–88. DOI: 10.1016/s1359-6446(05)03638-x.

Taga, Tetsuya, Masahiko Hibi, Yuuichi Hirata, Katsuhiko Yamasaki, Kiyoshi Yasukawa, Tadashi Matsuda, Toshio Hirano and Tadamitsu Kishimoto (1989). 'Interleukin-6 triggers the association of its receptor with a possible signal transducer, gp130'. In: *Cell* 58.3, pp. 573–581. DOI: 10.1016/0092-8674(89)90438-8.

Takeda, Kiyoshi, Tsuneyasu Kaisho, Nobuaki Yoshida, Junji Takeda, Tadamitsu Kishimoto and Shizuo Akira (1998). 'Stat3 Activation Is Responsible for IL-6-Dependent T Cell Proliferation Through Preventing Apoptosis: Generation and Characterization of T Cell-Specific Stat3-Deficient Mice'. In: *The Journal of Immunology* 161.9, pp. 4652–4660.

Takeuchi, Tsutomu, Yoshiya Tanaka, Hisashi Yamanaka, Masayoshi Harigai, Toshikazu Nakano, Koshiro Akagi, Yoshifumi Ukyo and Benjamin Hsu (2018). 'Efficacy and safety of sirukumab in Japanese patients with moderate to severe rheumatoid arthritis inad-

equately controlled by disease modifying anti-rheumatic drugs: Subgroup analysis of a phase 3 study'. In: *Modern Rheumatology*, pp. 1–9. DOI: 10.1080/14397595.2018.1428929.

Talbot, H. Keipp B., Romina Libster and Kathryn Edwards (2012). 'Influenza vaccination for older adults'. In: *Human Vaccines & Immunotherapeutics* 8.1, pp. 96–101. DOI: 10.4161/hv.8.1.18129.

Tan, Shuguang et al. (2017). 'Hemagglutinin-specific CD4 + T-cell responses following 2009-pH1N1 inactivated split-vaccine inoculation in humans'. In: *Vaccine* 35.42, pp. 5644–5652. DOI: 10.1016/j.vaccine.2017.08.061.

Tanaka, Masao, Masaaki Kishimura, Shoichi Ozaki, Fumio Osakada, Hidetaka Hashimoto, Mitsuo Okubo, Masao Murakami and Kazuwa Nakao (2000). 'Cloning of novel soluble gp130 and detection of its neutralizing autoantibodies in rheumatoid arthritis'. In: *Journal of Clinical Investigation* 106.1, pp. 137–144. DOI: 10.1172/jci7479.

Tari, Luis B. and Jagruti H. Patel (2014). 'Systematic Drug Repurposing Through Text Mining'. In: *Biomedical Literature Mining*. Ed. by Vinod D. Kumar and Hannah Jane Tipney. New York, NY: Springer New York, pp. 253–267. ISBN: 978-1-4939-0709-0. DOI: 10.1007/978-1-4939-0709-0_14.

Taylor, Nick Paul (25th Sept. 2017a). 'FDA rejects Johnson & Johnson's rheumatoid arthritis drug sirukumab on safety grounds, demands more data.' In: *FierceBiotech*. URL: https://www.fiercebiotech.com/biotech/fda-rejects-johnson-johnson-s-rheumatoid-arthritis-drug-sirukumab-safety-grounds-demands (visited on 26/02/2018).

— (17th Oct. 2017b). 'Johnson & Johnson drops sirukumab after FDA blow, cans phase 3 AML trial'. In: *FierceBiotech*. URL: https://www.fiercebiotech.com/biotech/johnson-johnson-drops-sirukumab-after-fda-blow-cans-phase-3-aml-trial (visited on 26/02/2018).

Tenhumberg, Stephanie, Björn Schuster, Lixin Zhu, Marina Kovaleva, Jürgen Scheller, Karl-Josef Kallen and Stefan Rose-John (2006). 'gp130 dimerization in the absence of ligand: Preformed cytokine receptor complexes'. In: *Biochemical and Biophysical Research Communications* 346.3, pp. 649–657. DOI: 10.1016/j.bbrc.2006.05.173.

Teorell, Torsten (1937a). 'Kinetics of distribution of substances administered to the body, I: The extravascular modes of administration'. In: *Archives internationales de pharmacodynamie et de therapie* 57, pp. 205–225.

— (1937b). 'Kinetics of distribution of substances administered to the body, II: The intravascular modes of administration'. In: *Archives internationales de pharmacodynamie et de therapie* 57, pp. 226–240.

Terasaki, Kaori, Sydney I. Ramirez and Shinji Makino (2016). 'Mechanistic Insight into the Host Transcription Inhibition Function of Rift

Valley Fever Virus NSs and Its Importance in Virulence'. In: *PLOS Neglected Tropical Diseases* 10.10. Ed. by Maya Williams, e0005047. DOI: 10.1371/journal.pntd.0005047.

Thiel, Christoph, Henrik Cordes, Isabel Conde, José Vicente Castell, Lars Mathias Blank and Lars Kuepfer (2016). 'Model-based contextualization of in vitro toxicity data quantitatively predicts in vivo drug response in patients'. In: *Archives of Toxicology* 91.2, pp. 865–883. DOI: 10.1007/s00204-016-1723-x.

Thiel, Stefan, Heike Dahmen, Astrid Martens, Gerhard Müller-Newen, Fred Schaper, Peter C. Heinrich and Lutz Graeve (1998). 'Constitutive internalization and association with adaptor protein-2 of the interleukin-6 signal transducer gp130'. In: *FEBS Letters* 441.2, pp. 231–234. DOI: 10.1016/s0014-5793(98)01559-2.

Thiel, Stefan, Ulrike Sommer, Marcin Kortylewski, Claude Haan, Iris Behrmann, Peter C Heinrich and Lutz Graeve (2000). 'Termination of IL-6-induced STAT activation is independent of receptor internalization but requires de novo protein synthesis'. In: *FEBS Letters* 470.1, pp. 15–19. DOI: 10.1016/s0014-5793(00)01276-x.

Tian, Jianjun Paul, Avner Friedman, Jin Wang and E. Antonio Chiocca (2008). 'Modeling the effects of resection, radiation and chemotherapy in glioblastoma'. In: *Journal of Neuro-Oncology* 91.3, pp. 287–293. DOI: 10.1007/s11060-008-9710-6.

Tornøe, Christoffer W., Henrik Agersø, Thomas Senderovitz, Henrik A. Nielsen, Henrik Madsen, Mats O. Karlsson and E. Niclas Jonsson (2007). 'Population pharmacokinetic/pharmacodynamic (PK/PD) modelling of the hypothalamic?pituitary?gonadal axis following treatment with GnRH analogues'. In: *British Journal of Clinical Pharmacology* 63.6, pp. 648–664. DOI: 10.1111/j.1365-2125.2006.02820.x.

Traynard, Pauline, Luis Tobalina, Federica Eduati, Laurence Calzone and Julio Saez-Rodriguez (2017). 'Logic Modeling in Quantitative Systems Pharmacology'. In: *CPT: Pharmacometrics & Systems Pharmacology* 6.8, pp. 499–511. DOI: 10.1002/psp4.12225.

Trstenjak, Bruno, Sasa Mikac and Dzenana Donko (2014). 'KNN with TF-IDF based Framework for Text Categorization'. In: *Procedia Engineering* 69, pp. 1356–1364. DOI: 10.1016/j.proeng.2014.03.129.

Trucchi, C., C. Paganino, A. Orsi, D. De Florentiis and F. Ansaldi (2015). 'Influenza vaccination in the elderly: why are the overall benefits still hotly debated?' In: *J Prev Med Hyg* 56.1, pp. 37–43.

Truman, R. W., G. J. Ebenezer, M. T. Pena, R. Sharma, G. Balamayooran, T. H. Gillingwater, D. M. Scollard, J. C. McArthur and A. Rambukkana (2014). 'The Armadillo as a Model for Peripheral Neuropathy in Leprosy'. In: *ILAR Journal* 54.3, pp. 304–314. DOI: 10.1093/ilar/ilt050.

Tsuchida, Anika I, Michiel Beekhuizen, Marijn Rutgers, Gerjo JVM van Osch, Joris EJ Bekkers, Arjan GJ Bot, Bernd Geurts, Wouter JA Dhert, Daniel BF Saris and Laura B Creemers (2012). 'Interleukin-6 is elevated in synovial fluid of patients with focal cartilage defects and stimulates cartilage matrix production in an in vitro regeneration model'. In: *Arthritis Res Ther* 14.6, R262. DOI: 10.1186/ar4107.

Tudor, C. O., K. E. Ross, G. Li, K. Vijay-Shanker, C. H. Wu and C. N. Arighi (2015). 'Construction of phosphorylation interaction networks by text mining of full-length articles using the eFIP system'. In: *Database* 2015.0, bav020–bav020. DOI: 10.1093/database/bav020.

U.S Food and Drug Administration, BLA: 125276. Biologic License Application review: BLA: 125276, January 2010. Available from: URL: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2010/125276s000PharmR.pdf.

Uhlen, M. et al. (2015). 'Tissue-based map of the human proteome'. In: *Science* 347.6220, pp. 1260419–1260419. DOI: 10.1126/science.1260419.

Ushijima, Ryosuke, Naoko Sakaguchi, Arihiro Kano, Atsushi Maruyama, Yoichi Miyamoto, Toshihiro Sekimoto, Yoshihiro Yoneda, Kenji Ogino and Taro Tachibana (2005). 'Extracellular signal-dependent nuclear import of STAT3 is mediated by various importin αs'. In: *Biochemical and Biophysical Research Communications* 330.3, pp. 880–886. DOI: 10.1016/j.bbrc.2005.03.063.

Vaidyanathan, Aparajitha, Lynne Sawers, Anne-Louise Gannon, Probir Chakravarty, Alison L Scott, Susan E Bray, Michelle J Ferguson and Gillian Smith (2016). 'ABCB1 (MDR1) induction defines a common resistance mechanism in paclitaxel- and olaparib-resistant ovarian cancer cells'. In: *British Journal of Cancer* 115.4, pp. 431–441. DOI: 10.1038/bjc.2016.203.

Vane, J. R. (1971). 'Inhibition of Prostaglandin Synthesis as a Mechanism of Action for Aspirin-like Drugs'. In: *Nature New Biology* 231.25, pp. 232–235. DOI: 10.1038/newbio231232a0.

Vermeire, S (2006). 'Laboratory markers in IBD: useful, magic, or unnecessary toys?' In: *Gut* 55.3, pp. 426–431. DOI: 10.1136/gut.2005.069476.

Villiers, Pierre J.T. De, A. Duncan Steele, Louis A. Hiemstra, Ruth Rappaport, Andrew J. Dunning, William C. Gruber and Bruce D. Forrest (2009). 'Efficacy and safety of a live attenuated influenza vaccine in adults 60 years of age and older'. In: *Vaccine* 28.1, pp. 228–234. DOI: 10.1016/j.vaccine.2009.09.092.

Vora, Suchi and Hui Yang (2017). 'A comprehensive study of eleven feature selection algorithms and their impact on text classification'. In: *2017 Computing Conference*. IEEE. DOI: 10.1109/sai.2017.8252136.

Vratsanos, George (August 2, 2017). *Sirukumab, Presentation to the Presentation to the Arthritis Advisory Committee*. Meeting of the Arthritis Advisory Committee, FDA, FDA White Oak Campus, Silver Spring, Maryland. Available from: URL: https://www.fda.gov/downloads/ AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/ UCM570357.pdf.

Wagner, John G (1981). 'History of pharmacokinetics'. In: *Pharmacology & therapeutics* 12.3, pp. 537–562.

Wallis, W. J. and P. A. Simkin (1983). 'Antirheumatic drug concentrations in human synovial fluid and synovial tissue. Observations on extravascular pharmacokinetics'. In: *Clin Pharmacokinet* 8.6, pp. 496–522.

Wang, Y. and G. M. Fuller (1994). 'Phosphorylation and internalization of gp130 occur after IL-6 activation of Jak2 kinase in hepatocytes.' In: *Molecular Biology of the Cell* 5.7, pp. 819–828. DOI: 10.1091/mbc.5.7.819.

Waring, Michael J. et al. (2015). 'An analysis of the attrition of drug candidates from four major pharmaceutical companies'. In: *Nature Reviews Drug Discovery* 14.7, pp. 475–486. DOI: 10.1038/nrd4609.

Weber, J., J. C. Yang, S. L. Topalian, D. R. Parkinson, D. S. Schwartzentruber, S. E. Ettinghausen, H. Gunn, A. Mixon, H. Kim and D. Cole (1993). 'Phase I trial of subcutaneous interleukin-6 in patients with advanced malignancies'. In: *J. Clin. Oncol.* 11.3, pp. 499–506.

Weeber, M. (2003). 'Generating Hypotheses by Discovering Implicit Associations in the Literature: A Case Report of a Search for New Potential Therapeutic Uses for Thalidomide'. In: *Journal of the American Medical Informatics Association* 10.3, pp. 252–259. DOI: 10.1197/jamia.m1158.

Wei, Chih-Hsuan, Hung-Yu Kao and Zhiyong Lu (2013). 'PubTator: a web-based text mining tool for assisting biocuration'. In: *Nucleic Acids Research* 41.W1, W518–W522. DOI: 10.1093/nar/gkt441.

Weiergraber, Oliver et al. (1995). 'Soluble Human Interleukin-6 Receptor. Expression in Insect Cells, Purification and Characterization'. In: *Eur J Biochem* 234.2, pp. 661–669. DOI: 10.1111/j.1432-1033.1995.661_b.x.

Wessels, J. A., T. W. Huizinga and H. J. Guchelaar (2008). 'Recent insights in the pharmacological actions of methotrexate in the treatment of rheumatoid arthritis'. In: *Rheumatology (Oxford)* 47.3, pp. 249–255. DOI: 10.1093/rheumatology/kem279.

Wilhelm, Abraham J., Peer de Graaf, Agnes I. Veldkamp, Jeroen J. W. M. Janssen, Peter C. Huijgens and Eleonora L. Swart (2012). 'Population pharmacokinetics of ciclosporin in haematopoietic allogeneic stem cell transplantation with emphasis on limited sampling strategy'. In: *British Journal of Clinical Pharmacology* 73.4, pp. 553–563. DOI: 10.1111/j.1365-2125.2011.04116.x.

Willmann, S., A. N. Edginton, K. Coboeken, G. Ahr and J. Lippert (2009). 'Risk to the breast-fed neonate from codeine treatment to the mother: a quantitative mechanistic modeling study'. In: *Clin. Pharmacol. Ther.* 86.6, pp. 634–643.

Wolking, Stefan, Elke Schaeffeler, Holger Lerche, Matthias Schwab and Anne T. Nies (2015). 'Impact of Genetic Polymorphisms of ABCB1 (MDR1, P-Glycoprotein) on Drug Disposition and Potential Clinical Implications: Update of the Literature'. In: *Clinical Pharmacokinetics* 54.7, pp. 709–735. DOI: 10.1007/s40262-015-0267-1.

Woo, Sukyung, Wojciech Krzyzanski and William J. Jusko (2007). 'Target-mediated pharmacokinetic and pharmacodynamic model of recombinant human erythropoietin (rHuEPO)'. In: *Journal of Pharmacokinetics and Pharmacodynamics* 34.6, pp. 849–868. DOI: 10.1007/s10928-007-9074-0.

Wrammert, Jens et al. (2008). 'Rapid cloning of high-affinity human monoclonal antibodies against influenza virus'. In: *Nature* 453.7195, pp. 667–671. DOI: 10.1038/nature06890.

Wu, Jing, Monica Saksena, Vincent Soriano, Eugenia Vispo and Nitin K Saksena (2015). 'Differential regulation of cytotoxicity pathway discriminating between HIV, HCV mono- and co-infection identified by transcriptome profiling of PBMCs'. In: *Virology Journal* 12.1, p. 4. DOI: 10.1186/s12985-014-0236-6.

Würthwein, Gudrun et al. (2013). 'Population Pharmacokinetics of Escalating Doses of Caspofungin in a Phase II Study of Patients with Invasive Aspergillosis'. In: *Antimicrobial Agents and Chemotherapy* 57.4, pp. 1664–1671. DOI: 10.1128/aac.01912-12.

Xu, Z., E. Bouman-Thio, C. Comisar, B. Frederick, B. Van Hartingsveldt, J. C. Marini, H. M. Davis and H. Zhou (2011). 'Pharmacokinetics, pharmacodynamics and safety of a human anti-IL-6 monoclonal antibody (sirukumab) in healthy subjects in a first-in-human study'. In: *Br J Clin Pharmacol* 72.2, pp. 270–281.

Yamaoka, Kiyoshi, Terumichi Nakagawa and Toyozo Uno (1978). 'Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations'. In: *Journal of Pharmacokinetics and Biopharmaceutics* 6.2, pp. 165–175. DOI: 10.1007/bf01117450.

Yan, Xiaoyu, Donald E. Mager and Wojciech Krzyzanski (2009). 'Selection between Michaelis–Menten and target-mediated drug disposition pharmacokinetic models'. In: *Journal of Pharmacokinetics and Pharmacodynamics* 37.1, pp. 25–47. DOI: 10.1007/s10928-009-9142-8.

Yang, William E. et al. (2016). 'Transcriptomic Analysis of the Host Response and Innate Resilience to EnterotoxigenicEscherichia coliInfection in Humans'. In: *Journal of Infectious Diseases* 213.9, pp. 1495–1504. DOI: 10.1093/infdis/jiv593.

Yang, Yiming and Xin Liu (1999). 'A re-examination of text categorization methods'. In: *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '99*. ACM Press. DOI: 10.1145/312624.312647.

Yang, Yiming and Jan O. Pedersen (1997). 'A Comparative Study on Feature Selection in Text Categorization'. In: Morgan Kaufmann Publishers, pp. 412–420.

Yano, Y., S. L. Beal and L. B. Sheiner (2001). 'Evaluating pharmacokinetic/pharmacodynamic models using the posterior predictive check'. In: *J Pharmacokinet Pharmacodyn* 28.2, pp. 171–192.

Yates, James W. T., Phillippa Dudley, Jane Cheng, Celina D'Cruz and Barry R. Davies (2015). 'Validation of a predictive modeling approach to demonstrate the relative efficacy of three different schedules of the AKT inhibitor AZD5363'. In: *Cancer Chemotherapy and Pharmacology* 76.2, pp. 343–356. DOI: 10.1007/s00280-015-2795-7.

Ye, Hua, Jing Zhang, Jun Wang, Yanyan Gao, Yan Du, Chun Li, Minghua Deng, Jianping Guo and Zhanguo Li (2015). 'CD4 T-cell transcriptome analysis reveals aberrant regulation of STAT3 and Wnt signaling pathways in rheumatoid arthritis: evidence from a case–control study'. In: *Arthritis Research & Therapy* 17.1. DOI: 10.1186/s13075-015-0590-9.

Yoo, Hee-Doo, Hea-Young Cho, Sang-No Lee, Hwa Yoon and Yong-Bok Lee (2012). 'Population pharmacokinetic analysis of risperidone and 9-hydroxyrisperidone with genetic polymorphisms of CYP2D6 and ABCB1'. In: *Journal of Pharmacokinetics and Pharmacodynamics* 39.4, pp. 329–341. DOI: 10.1007/s10928-012-9253-5.

Yu, B. (2008). 'An evaluation of text classification methods for literary study'. In: *Literary and Linguistic Computing* 23.3, pp. 327–343. DOI: 10.1093/llc/fqn015.

Zacour, Mary, Brian J. Ward, Angela Brewer, Patrick Tang, Guy Boivin, Yan Li, Michelle Warhuus, Shelly A. McNeil, Jason J. LeBlanc and Todd F. Hatchette (2016). 'Standardization of Hemagglutination Inhibition Assay for Influenza Serology Allows for High Reproducibility between Laboratories'. In: *Clinical and Vaccine Immunology* 23.3. Ed. by R. L. Hodinka, pp. 236–242. DOI: 10.1128/cvi.00613-15.

Zanger, Ulrich M. and Matthias Schwab (2013). 'Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation'. In: *Pharmacology & Therapeutics* 138.1, pp. 103–141. DOI: 10.1016/j.pharmthera.2012.12.007.

Zararsız, Gökmen, Dincer Goksuluk, Selcuk Korkmaz, Vahap Eldem, Gozde Erturk Zararsiz, Izzet Parug Duru and Ahmet Ozturk (2017). 'A comprehensive simulation study on classification of RNA-Seq

data'. In: *PLOS ONE* 12.8. Ed. by Christian Schönbach, e0182507. DOI: 10.1371/journal.pone.0182507.

Zhang, D., M. Sun, D. Samols and I. Kushner (1996). 'STAT3 participates in transcriptional activation of the C-reactive protein gene by interleukin-6'. In: *J. Biol. Chem.* 271.16, pp. 9503–9509. DOI: 10.1074/jbc.271.16.9503.

Zhang, Jingwei et al. (2011). 'Cellular pharmacokinetic mechanisms of adriamycin resistance and its modulation by 20(S)-ginsenoside Rh2 in MCF-7/Adr cells'. In: *British Journal of Pharmacology* 165.1, pp. 120–134. DOI: 10.1111/j.1476-5381.2011.01505.x.

Zhang, Liang et al. (2012). 'Inhibition of pyrimidine synthesis reverses viral virulence factor-mediated block of mRNA nuclear export'. In: *The Journal of Cell Biology* 196.3, pp. 315–326. DOI: 10.1083/jcb.201107058.

Zhang, Xiaoping, Angela Georgy and Lucy Rowell (2013a). 'Pharmacokinetics and pharmacodynamics of tocilizumab, a humanized anti-interleukin-6 receptor monoclonal antibody, following single-dose administration by subcutaneous and intravenous routes to healthy subjects'. In: *Int. Journal of Clinical Pharmacology and Therapeutics* 51.06, pp. 443–455. DOI: 10.5414/cp201819.

Zhang, Xiaoping, Ya-Chi Chen, Scott Fettner, Lucy Rowell, Tatiana Gott, Paul Grimsey and Adam Unsworth (2013b). 'Pharmacokinetics and pharmacodynamics of tocilizumab after subcutaneous administration in patients with rheumatoid arthritis'. In: *Int. Journal of Clinical Pharmacology and Therapeutics* 51.08, pp. 620–630. DOI: 10.5414/cp201904.

Zhao, Shan and Ravi Iyengar (2012). 'Systems Pharmacology: Network Analysis to Identify Multiscale Mechanisms of Drug Action'. In: *Annual Review of Pharmacology and Toxicology* 52.1, pp. 505–521. DOI: 10.1146/annurev-pharmtox-010611-134520.

Zhou, Xiaofei, Yue Hu and Li Guo (2014). 'Text Categorization based on Clustering Feature Selection'. In: *Procedia Computer Science* 31, pp. 398–405. DOI: 10.1016/j.procs.2014.05.283.

Zhuang, Y., D. E. de Vries, S. J. Marciniak, H. Liu, H. Zhou, H. M. Davis, F. Leon, D. Raible and Z. Xu (2016). 'Absolute Bioavailability and Pharmacokinetic Comparability of Sirukumab Following Subcutaneous Administration by a Prefilled Syringe or an Autoinjector'. In: *Clinical Pharmacology in Drug Development*. DOI: 10.1002/cpdd.328.

Ziegler, Louise, Ashwini Gajulapuri, Paolo Frumento, Alice Bonomi, Håkan Wallén, Ulf de Faire, Stefan Rose-John and Bruna Gigante (2018). 'Interleukin 6 trans-signalling and risk of future cardiovascular events'. In: *Cardiovascular Research* 115.1, pp. 213–221. DOI: 10.1093/cvr/cvy191.

Zohlnhöfer, Dietlind, Lutz Graeve, Stefan Rose-John, Heidi Schooltink, Elke Dittrich and Peter Claus Heinrich (1992). 'The hepatic interleukin-

6 receptor Down-regulation of the interleukin-6 binding subunit (gp80) by its ligand'. In: *FEBS Letters* 306.2-3, pp. 219–222. DOI: 10.1016/0014-5793(92)81004-6.

Zola, H. and L. Flego (1992). 'Expression of interleukin-6 receptor on blood lymphocytes without in vitro activation'. In: *Immunology* 76.2, pp. 338–340.