Mechanistic modelling to determine the limits of clinical and pre-clinical cancer immunotherapies



Liam V. Brown
Lady Margaret Hall
University of Oxford

A thesis submitted in partial fulfilment of the requirements for
Doctor of Philosophy

Michaelmas 2018

Acknowledgements

Attribution: Eamonn Gaffney, Mark Coles, Jonathan Wagg and I contributed to the research of every chapter. Tim Elliott provided advice and materials for Chapters 1 and 2. He, Rachel Darley and Andy van Hatern carried out the experimental work required for Chapter 2. Likewise, Ann Ager provided advice and materials for Chapters 3 and 4, and assisted with the writing of a manuscript for Chapter 3.

Thanks: are due to many people. Without them, my work would not have looked the same, and I can't continue to repeat that in every line. But, some special words are owed to some: To my supervisors, who've worked hard and patiently with me for these three years and of course provided invaluable insight. Thanks especially to Jon, who has been incredibly friendly and hospitable over the years, particularly in Switzerland. Thank you for everything that you did for me. I really hope that we can continue to work together in the future!

To my collaborators, without whom I would have been lost in questions of what the biology *could* look like.

To my research group, and to the staff and facilitators within the maths department.

To those that provided me money to do my research and to live: the EPSRC, Clarendon and Roche.

To Lady Margaret Hall, which provided me a home – in a literal sense and a welfare sense – and a place where I met many wonderful people and learnt many things, from rowing to the fact that a fish knife is indeed not a broken knife.

To my friends and family, for keeping me sane.

Abstract

Immune-mediated clearance of a tumour is predicated on successful activation, migration and engagement of cytotoxic $CD8^+$ T-cells. There are numerous factors that may impede these processes and that vary considerably between patients. In this thesis, I describe the use of mechanistic models to quantify and identify mechanisms that limit immune responses and immunotherapeutics against tumours. I describe the use of a model of T-cell activation in the lymph node to quantify the impact of patient and vaccine-specific variables on short peptide vaccination success. The model is used to simulate a virtual clinical trial, with predicted patient responses consistent with IMA901, a clinical trial of a short peptide vaccination in renal cell carcinoma patients. This leads to the conclusion that the limited efficacy of IMA901 could be due to the short peptide off-rates, and to suggestions that could have improved IMA901's results.

I then describe an ODE model of T-cell trafficking through the bloodstream to quantify the maximum delivery rate of bioengineered T-cells to healthy tissue and tumours in different organs and species, and how these rates scale between species. Predicted absolute delivery rates of T-cells in mice are found to be much larger than equivalent rates in humans. This could explain why pre-clinical success of bioengineered T-cells in solid tumours has not translated to the clinic as it has for haematological cancers. Finally, this model is extended with PDEs to quantify the persistence of T-cells within organs and fit to lymphocyte localisation data in the rat. The advantages over the use of ODEs for describing T-cell localisation are discussed along with potential and planned future work.

I close by discussing results in an immuno-oncological context, and the extent to which mouse, mathematical and mechanistic models are representative or useful in human medicine.

Contents

0 Introduction and literature review **0**

0.0 Motivation and objectives 0

0.1 Biological review 2

0.1.0 The adaptive immune system 2

0.1.1 Antigen presentation 5

0.1.2 Lymphoid tissue structure 7

0.1.2.0 Lymph nodes 8

0.1.2.1 Peyer's patches 9

0.1.2.2 The spleen 10

0.1.3 T- and B-cell trafficking and homing 10

0.1.4 T-cell cytotoxicity 12

0.1.5 The cancer-immunity cycle 12

0.1.6 Immuno-oncology and immunotherapy 14

0.2 Mathematical and computational modelling in immuno-oncology 17

0.2.0 Models of tumour kinetics 18

0.2.1 Models of the lymph node 21

0.2.2 Pharmacokinetics and pharmacodynamics 25

0.2.3 Models of T-cell trafficking 28

0.2.4 Age-structured models 31

0.3 Summary of DPhil 31

1 An *in silico* model of cytotoxic T-lymphocyte activation in the lymph node following short peptide vaccination **35**

1.0 Introduction 35

1.1 Methods 36

1.1.0 Model overview 36

1.1.1 Vaccination site 38

1.1.2 Agent-based lymph node model 39

1.1.3 Probability of T-cell activation 40
1.1.4 T-cell dynamics 43
1.1.5 Dendritic cell dynamics 43
1.1.6 Parameter selection 45
1.1.7 Sensitivity analyses 45
1.2 Results 47
1.2.0 Peptide dynamics 47
1.2.0.0 Perturbation analysis 47
1.2.0.1 Time course 48
1.2.1 CTL activation probability against numbers of T-cells or DCs 49
1.2.2 CTL activation probability against antigen off-rate 50
1.2.3 Comparison to previous work 50
1.2.4 Motion of dendritic cells 52
1.2.5 Stochastic entry of dendritic cells 54
1.2.6 Analysis of the sensitivity of CTL activation probability 56
1.3 Discussion and conclusions 57
1.3.0 Precise dynamics of single peptide binding in the dermis are not an important factor for T-cell activation probability 57
1.3.1 Model observations are not sensitive to dendritic cell motion 60
1.3.2 The role of peptide off-rate and DC/T-cell numbers in CTL activation 61
1.3.3 Differences between this model and prior work 62
1.3.4 T-cell movement and range 63
1.3.5 The sensitive parameters in our model support expectations 64
1.3.6 Conclusions and future work 64
1.A Appendix: Sensitivity analysis techniques 66
1.A.0 Random Forest 66
1.A.1 Variance-based sensitivity analysis 66
1.A.2 Extended FAST 67

2 What determines efficacy of short-peptide vaccinations? Simulating mechanisms of clinical failure 69

2.0 Introduction 69
2.1 Methods 70
2.1.0 IMA901 trial procedure 70
2.1.1 Virtual clinical trial procedure 70
2.1.2 Measurement of peptide-MHC-I off-rates 72

2.1.2.0 Fluorescence of immobilised MHC-I . . . 72
2.1.2.1 B-cell MHC-I BFA decay . . . 73
2.1.3 Finding all parameter distributions that match results via Approximate Bayesian Computation . . . 74
2.1.4 Hastening model evaluation . . . 75
2.1.5 Fitting experimental dendritic cell migration data . . . 78
2.1.6 Sensitivity analysis . . . 78
2.2 Results . . . 79
2.2.0 Prediction of peptide response for one virtual patient . . . 79
2.2.1 Simulated clinical trial results . . . 81
2.2.2 Measurement of the pMHC-I off-rates of the IMA901 peptides . . . 83
2.2.3 Simulated clinical trial results using pMHC-I off-rates estimated from BIMAS and NetMHC . . . 84
2.2.4 Alterations to clinical trial design . . . 85
2.2.5 Difference between observed results in phases I, II and III of IMA901 85
2.2.6 Simplified model accuracy . . . 87
2.2.7 Range of parameter values matching results of IMA901 . . . 89
2.2.8 Perturbation of local solutions to fit each phase . . . 90
2.2.9 Approximate Bayesian Computation result accuracy . . . 91
2.2.10 Approximate Bayesian Computation result reproducibility . . . 92
2.2.11 Model sensitivity analysis . . . 93
2.3 Discussion . . . 95
2.3.0 Peptide–MHC-I off-rates . . . 95
2.3.1 Parameter sensitivity and robustness . . . 99
2.3.2 Difference between the phases of IMA-901 . . . 101
2.3.3 Reliability of Approximate Bayesian Computation . . . 102
2.3.4 Model consequences for IMA901 . . . 104
2.4 Conclusions . . . 105
2.A Unsuccessful methods to hasten model evaluation . . . 107
2.A.0 Emulation . . . 107
2.A.0.0 Analytical solution . . . 109

3 The circulatory system limits T-cell delivery to tumours: implications for comparative CAR T-cell dosing of mice and men 111
3.0 Introduction . . . 111
3.1 Methods . . . 113

3.1.0 Model definition . . . 113
3.1.0.0 Summary . . . 113
3.1.0.1 ODE system . . . 115
3.1.0.2 Tumour perfusion . . . 117
3.1.0.3 Initial conditions . . . 119
3.1.0.4 Extravasation proportion values . . . 119
3.1.1 Parameter selection from literature . . . 120
3.2 Results . . . 121
3.2.0 ODE results versus time . . . 121
3.2.1 CAR T-cell delivery to organs in humans, mice and rats . . . 122
3.2.2 Tumour perfusion compared to healthy tissue . . . 124
3.2.3 CAR T-cell delivery to human tumours . . . 125
3.2.4 Delivery rates of other immune populations in humans . . . 126
3.2.5 Tumour control . . . 126
3.2.6 Sensitivity analysis . . . 127
3.3 Discussion . . . 129
3.3.0 Organ-specific delivery rates and their variation . . . 129
3.3.1 Species-specific delivery rates and dosage scaling . . . 130
3.3.2 Refinement of delivery rate predictions . . . 133
3.4 Conclusions . . . 134
3.A Tumour growth model . . . 135

4 Quantification of temporal lymphocyte localisation 137
4.0 Introduction . . . 137
4.1 Methods . . . 138
4.1.0 Lymphocyte localisation data . . . 138
4.1.1 Methods to fit model to data . . . 139
4.1.2 PDE model definition . . . 141
4.1.3 Organ exit propensity γ . . . 143
4.1.4 Numerical algorithms to solve PDEs . . . 143
4.1.5 Method of characteristics . . . 145
4.1.6 Combined ODE and characteristic solver . . . 146
4.2 Results . . . 149
4.2.0 Sensitivity of steady state of ODE system . . . 149
4.2.1 ODE fit to localisation data . . . 150
4.2.2 Motivation for PDEs . . . 152

4.2.3 Advection equation 154
4.2.4 PDE solution to single organ system 156
4.2.5 Method of characteristics solution 158
4.2.5.0 Simple system 158
4.2.5.1 Full system 159
4.2.6 PDE fit to localisation data 159
4.3 Discussion 162
4.4 Future work 169

5 Thesis discussion **171**
5.0 Research questions and hypotheses 171
5.1 Research summary 172
5.2 Context 174
5.3 Future work 178

Bibliography **180**

List of abbreviations

ABA	Allogen-stimulated Brachial and Axillary lymph node cells
ABC	Approximate Bayesian Computation
ALL	Acute Lymphoblastic Leukaemia
APC	Antigen Presenting Cell
BC	Boundary Condition
BFA	Brefeldin A
BS	Blood supply
CAIX	Carbonic Anhydrase nine
CAR	Chimeric Antigen Receptor
CD4/CD8	Cluster of Differentiation 4/8
CDF	Cumulative Distribution Function
CEA	Carcinoembryonic Antigen
CFA	Complete Freund's Adjuvant-stimulated brachial/axillary lymph node cells
CTL	Cytotoxic T-Lymphocyte
CTLA	Cytotoxic T-Lymphocyte(-Associated protein)
CYP	Cytochromes P450
DC	Dendritic Cell
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
(e)FAST	(extended) Fourier Amplitude Sensitivity Testing
FACS	Fluorescence-Activated Cell Sorting
FITC	Fluorescein Isothiocyanate [an antibody]
FLPSDC*FPSV	[A fluorescent standard peptide]
FRC	Fibroblastic Reticular Cell
G(M)-CSF	Granulocyte-(Macrophage) Colony-Stimulating Factor
GLSC	Graphics Library for Scientific Computing
HBV	Hepatitis-B virus
HEV	High Endothelial Venule
HLA	Human Leukocyte Antigen
ICRP	International Commission on Radiological Protection
IFN	Interferon
IL	Interleukin
LI	Large Intestine
LN	Lymph Node
MATLAB	Matrix Laboratory; programming language
MB	Megabyte
MFI	Mean Fluorescence Intensity

MFP	Mean Free Path
MHC	Major Histocompatibility Complex
MLN	Mesenteric Lymph Node cells
MSC	Mesenchymal Stem Cells
NK	Natural Killer
ODE	Ordinary Differential Equation
OS	Overall Survival
PBPK	Physiologically-Based Pharmacokinetic(s)
PBS	Phosphate-Buffered Saline
PC	Pulmonary Circuit
PD	Pharmacodynamic
PD-(L)1	Programmed Death (Ligand) 1
PDE	Partial Differential Equation
PKPD	Pharmacokinetics and Pharmacodynamics
PMC	Population Monte Carlo
(p)MHC(-I)	(peptide-) Major Histocompatibility Complex (one)
RCC	Renal Cell Carcinoma
RNA	Ribonucleic Acid
SED	Subepithelial Dome
SI	Small Intestine
SMC	Sequential Monte Carlo
TAMRA	Tetramethyl rhodamine
TAP	Transporter associated with Antigen Processing
TBO	Tumour-Bearing Organ
TCR	T-Cell Receptor
TDA	Thymus-Dependent Area
TDL	Thoracic Duct Lymphoblasts
TDLn	Thoracic Duct Lymphocytes – naïve cells
TNF	Tumour Necrosis Factor
TUK	Tumour Kinetic
US	United States
UV	Ultraviolet
VDJ	Variable, Diversity and Joining [segments]

Chapter 0

Introduction and literature review

Note on use of pronouns: *I have used 'I', 'me', 'my' and 'mine' to describe the writing of this thesis. Where 'we', 'us', 'our' and 'ours' appears, it refers to all those authors who collaborated on the research, both as supervisors or from further afield. These names are listed in the Acknowledgements.*

0.0 Motivation and objectives

Spontaneous regression of cancer is rare and attributed to immune mechanisms. It is hypothesised that the immune system destroys potential cancer cells throughout life via a process known as immunosurveillance [1], and evasion of the immune system is one of the so-called 'hallmarks' of cancer [2]. There are many mechanisms by which cancer can escape or out-compete the immune system and many therapies have been designed to combat this. Tumour cells can down-regulate surface receptors or molecules that can be recognised by the immune system (specifically, by 'killer' T-cells), which may be combated by the *in vitro* preparation and introduction of bioengineered T-cells that bind directly to molecules on tumour cells or have synthetic T-cell receptors. Tumours may inhibit the function of T-cells by upregulating the ligands of the Programmed Death receptor on T-cells (PD-1) or CTLA-4, against which anti-PD-(L)1 or anti-CTLA4 therapies can lead to clinical efficacy. The tumour may secrete other factors that recruit tumour-promoting or immune-suppressing cells, but a 'cancer vaccination' may be administered to raise an uninhibited T-cell response outside the tumour microenvironment, or a patient's tumour invading lymphocytes may be extracted and multiplied *ex vivo*. However, therapies such as these are often only effective in a minority of patients or a minority of tumour types, and many also drive toxicities. For some years, the clinic has been advancing ahead of scientific and especially quantitative understanding, making prediction of efficacy or adverse reactions difficult. For example, there have been many clinical studies of 'cancer vaccinations' that attempt to raise immune

responses against cancer. Of 167 completed phase II and III trials of cancer vaccines present in databases such as *Pubmed* by November 2015, only 14 had met their primary endpoints, according to an internal report in Roche. The reasons behind these failures, however, are not clear. Another example is the multitude of clinical studies of bioengineered T-cells that are designed to target cancer, which are clinically successful for leukaemias and lymphomas, but whose success for solid tumours in mice has failed to transfer to humans [3, 4]. Once more, there have been multiple hypotheses put forward to explain this but the problem has been neither solved nor extensively quantified. In this DPhil, we have aimed to place quantitative bounds on the expected efficacy of various treatments and modalities in cancer immunotherapy. We have produced simple, mechanistic models in order to extract biological meaning and hypotheses without our conclusions relying strongly on our choice of parameter values or the complexities of our models. In line with these aims, we seek to find the probabilities of successful outcomes, 'thresholds' for responses to occur or (maximal) limits on efficacies, rather than exact quantitation of the tumour or immune systems.

The following chapters of this thesis will detail the models that I have produced with my supervisors and collaborators. I will describe a model of T-cell activation in the lymph node, the use of this model to study the factors limiting the efficacy of a cancer vaccine, and two models of T-cell trafficking across the body, which I use to quantify factors limiting the efficacy of cellular therapies, the localisation of T-cells in different organs and their propensities for extravasating into or persisting within different organs. In this introduction, I will detail the biology underlying these models by describing,

- the adaptive immune system, including T-cells, which is responsible for the ability of the vertebrate immune system to react to and target almost any pathogen,
- the process by which foreign molecules (*e.g.* from viruses or bacteria) are presented to the immune system, leading to adaptive immune system activation and the eradication of invading pathogens,
- how lymphoid tissue such as the lymph node is specialised to facilitate this process,
- how T-cells traffic to and extravasate into different tissues to carry out their functions,
- how T-cells target and kill their target cells,
- the *cancer-immunity cycle*: a proposed mechanism [5] through which the immune system interacts with and eliminates tumours, and

- how immuno-oncology and immunotherapies attempt to prevent the breakdown of this cycle, focusing in particular on two immunotherapies: 'cancer vaccines' and bio-engineered T-cells.

Finally, I will briefly summarise previous mathematical and computational models of these biological areas: tumour kinetics, lymph node dynamics, pharmacokinetics and pharmacodynamics (*i.e.* how drugs and other molecules traffic across the body) and trafficking of T-cells and other cellular therapies.

0.1 Biological review

0.1.0 The adaptive immune system[0]

The vertebrate immune system has two branches, the *innate* and *adaptive* immune systems. The innate immune system makes rapid but static responses to invaders – it is characterised by such things as pattern recognition receptors that identify molecules common in viruses and bacteria but absent in the host animal. Its members include invader-engulfing macrophages and destructive cytokine-producing neutrophils, the complement system, which produces 'membrane attack complexes' on the surface of every cell to destroy unrecognised cells, and natural killer cells, which bind to target cells and destroy them via cell signalling or the delivery of destructive enzymes. Members of the innate system can produce cytokines that promote inflammation, upregulate the proliferation or production of themselves or other members of the innate immune system, or promote the recruitment of other members of the innate or adaptive immune systems to the site. The innate system has been present within animals for hundreds of millions of years, and has co-evolved in an environment with pathogens. Some pathogens have evolved to hide from the innate immune system, by presenting no common features for it to recognise or hiding within other cells (*e.g.* viruses). It is at this point that the adaptive immune system is required. It consists of two major classes of cells: T-cells, which mature in the thymus, and B-cells, which mature in the bone marrow. The receptors on each B and T-cell are (nearly) unique; each has a high affinity towards a different molecule, its *cognate antigen*. Together, T-cells are said to be able to recognise any protein, and B-cells any organic molecule, in the world. Though they have similar aims and look similar under a light microscope, the cell types have quite different lifespans and behaviours. B-cells produce antibodies with a variety of functions. The *base unit* of an antibody contains two heavy chains which make up the 'constant' region and two light chains that make up the antigen binding region, *i.e.* the part that can bind to invaders. Antibodies can 'decorate' invaders, preventing viruses

[0]Much of the basic biology described in Section 0.1 follows discussion by textbooks [6, 7].

from entering cells or otherwise impeding their function, can clump bacteria and viruses together, can facilitate the rapid binding and activation of the complement system against otherwise invisible (to the innate system) threats, can facilitate the binding of natural killer cells to and killing of target cells, or act as signalling molecules on mast cells to cause them to 'degranulate' (release destructive chemicals) and destroy large invaders that have penetrated the body's outer barriers. 'Killer' T-cells are able to bind directly to infected cells and induce cell death akin to the way natural killer cells do. Like the antibodies produced by B-cells, each T-cell is specific for a particular antigen, which it recognises through T-cell receptors on its membrane. Both B-cell antibodies and T-cells can be divided further into subtypes. There are five antibody classes with distinct constant regions: IgD, IgG, IgM, IgA and IgE (though IgG can be further subdivided). Each constant region has different binding properties, causing the antibodies to bind together in a pair or a pentamer, or to bind to natural killer cells to facilitate their function, or to complement proteins, or to mast cells. B-cells change which class of antibodies they produce in response to signals produced by the innate immune system or by a class of T-cells called 'helper' T-cells. There are three major classes of T-cells, classified according to their surface receptors, which are in turn used to infer their function. Killer T-cells are more properly known as cytotoxic T-cells or as $CD8^+$ T-cells, as they express CD8 on their cell membranes. It is these cells that are able to bind to infected host cells and kill them. Other T-cells express not CD8 but CD4, and these are either helper or regulatory cells. Helper cells (which also consist of multiple subclasses) express cytokines that are able to influence the activity of other immune cells according to the nature of the immune challenge, for example to make B-cells switch to anti-parasite IgE antibodies. Regulatory cells are able to suppress the activity of immune cells, particularly cytotoxic T-cells, to prevent the immune system from becoming too active in regions such as the intestines, which would otherwise be permanently inflamed due to the presence of commensal bacteria. A final 5% of T-cells express neither CD4 nor CD8, and the 'T-cell receptors' with which they bind to other cells are made not of the usual α and β proteins, but γ and δ proteins. They are mainly found in epithelial tissues (*e.g.* skin epidermis) and the gastrointestinal tract; they are involved with tissue homeostasis and in first-response immune defence [8]. They remain a subject of current research, but will not be the focus of this work.

T-cells and B-cells must be able to respond to any dangerous pathogen if they are to prevent the death of the host animal, and are hence able to respond to a great breadth of organic molecules. T-cells, however, are limited to the recognition of peptides presented by specialised receptors known as Major Histocompatibility Complexes. $CD8^+$ T-cells recognise peptides of length 8-11 amino acids on MHC-I receptors, and $CD4^+$ T-cells recognise

peptides of length 11-30 amino acids presented by MHC-II receptors. Were vertebrate DNA to code for every possible T- and B-cell receptor (where the B-cell receptor is simply its antibody embedded into the membrane), however, this would require an overwhelmingly large amount of information. Instead, the diversity is achieved by randomisation of the sections of DNA used to construct these receptors. In B-cells, the relevant DNA codes for gene segments known as V (variable), D (diversity) and J (joining) segments, followed by genes corresponding to the constant regions M, D, G, E and A. A random set of V, D and J segments are chosen, and if the segments form a complete and stable complex, then they become the B-cell receptor, and all other V, D and J segments are rendered inaccessible along with the genes on the other chromosome. The constant region determines the corresponding type of antibodies that the B-cell produces, and may be changed according to the B-cell's class as described above. When a B-cell becomes activated and begins to proliferate, the V, D and J genes are subject to a much higher rate of mutation than normal vertebrate DNA. The rate of B-cell proliferation is proportional to the rate of signalling (and thus affinity) of the B-cell receptors, and thus this process results in fine-tuning of the B-cell receptors towards higher affinities, known as somatic hypermutation. These two processes together generate a massive diversity of B-cell receptors that produces the specificity and breadth required for recognition of pathogenic molecules. In (most) T-cells, it is α and β proteins that are determined by VJ and VDJ genes respectively [9]. T-cells do not undergo somatic hypermutation, as they need recognise only proteins bound to special receptors (see Section 0.1.1) and thus perhaps do not need this additional process. Both T- and B-cells, however, are subjected to selection for self-tolerance; any cells that recognise 'self-antigen' are killed.

The fact that every T- and B- cell is unique and essentially responds to one molecule (or a very narrow band of similar molecules) raises the question of the numbers of cells required. The body may be invaded by any pathogen at any time, necessitating the presence of at least one of each 'type' of T- and B- cell in the body. However, during an infection, there may be many copies of an invader and large numbers of clones of particular T- and B-cells may be required. Ordinarily, un-activated (naïve) B- and T-cells are present in low numbers and are neither highly cytotoxic nor rapid producers of antibodies. When a pathogen invades the body, members of the innate immune system respond immediately to the threat due to pattern recognition receptors. Some of these cells are 'antigen presenting cells', that take antigen from the infection site to present to T-cells, and B-cells are presented antigen with a similar mechanism. At this point, B- and T-cells begin to proliferate to form the far greater numbers required to combat infection, and differentiate into either effector or memory cells. The effector cells become much more cytotoxic or produce many more antibodies, and

greatly upregulate the numbers of T- or B-cell receptors. Once the infection passes, the effector cells die and leave only the memory cells. These are long-lasting cells which are able to rapidly proliferate into even greater numbers of effector (and memory) cells, should their cognate antigen be presented to them once more. Some of them are central memory cells that traverse the bloodstream and lymphatic organs (see Section 0.1.2), whereas others are peripheral cells that lie in wait for antigen in outer tissues. The dual requirements of diversity and specificity of immune response place opposite selection pressures on the immune system. If T- and B-cells were too specific for their epitopes, then the probability of a given antigen being presented to lymphocytes cognate to it would become too low to raise an immune response before the animal succumbs to infection. Too little, and lymphocytes may not bind strongly enough to a given antigen to eradicate the invader, or may begin to recognise too many self-peptides as foreign. If the T- and B-cell repertoires were too diverse, then the probability of lymphocytes finding their cognate antigen would again become too low. Too limited, and some pathogens may escape the notice of the immune system. In humans, approximately 1 in 100,000 to 1 in 1,000,000 T-cells can respond to a given antigen [10], a ratio that is approximately conserved between different species, scaling with the log of animal mass [11].

0.1.1 Antigen presentation

Though this explains how enough effector cells are synthesised, it does not explain how antigen can be presented to the rare T- and B-cells that recognise it in a timely fashion, even if antigen presenting cells lie beneath every external surface of the body. There are several antigen presenting cells (APCs): dendritic cells, macrophages and B-cells. Each of these cells can present major histocompatibility complexes (MHC) of classes I and II. MHC-I typically contain peptides from the inside of the cell and are in fact expressed by nearly all cells in the body, whereas MHC-II are unique to APCs and present peptides found outside the cell. Within the APC, newly minted MHC-II receptors are enclosed in a vesicle. External proteins sampled from the surrounding tissues are brought within the cell enclosed within another vesicle, and the two are brought together and merged. Enzymes within the former vesicle then cleave the sampled protein to form peptides between 13 and 25 amino acids long. This binds to the MHC-II molecule and the new peptide-MHC (pMHC) complex is transported to the cell membrane for presentation. A similar process causes intracellular proteins to be presented on pMHC-I complexes. There are proteasomes within the cytoplasm of all cells that will cleave old, damaged or defective proteins. Inevitably, some of these peptides will be between around 8-11 amino acids long, which is the perfect length to fit onto MHC-I. If such peptides encounter transporters named TAP, they are

transported within the endoplasmic reticulum and – should the ends of the amino acid have suitable chemical properties (such as basic or hydrophobic amino acids), then they will be loaded onto MHC-I. The new peptide–MHC-I complex is contained within a vesicle and transported to the surface of the APC. In antigen presenting cells, cytoplasmic proteasomes selectively cut near basic or hydrophobic amino acids, which increases the probability that peptides are suitable for MHC-I molecules.

Immediately after an infection, APCs begin to upregulate MHC-I and MHC-II. Macrophages and B-cells are able to present to T-cells that happen to be at the site of infection, to prevent them from becoming anergised (exhausted). Dendritic cells (DCs), however, leave the tissue after activation by their pattern recognition receptors. They move through lymphatic vessels until they reach the T-cell area (paracortex) of a lymph node, and stay there for a few days before they die [12–17]. Antigen that have been opsonised by complement or antibodies similarly drain through the lymphatics and bind to follicular DCs within the B-cell areas (follicles) of the lymph nodes, by both active and passive processes. For T-cells to become activated, DCs need to present pMHC complexes to them. DCs, as their name implies, have long dendrites that they can use to sweep the area up to 20µm away [13, 17–20], allowing them to contact as many as 5000 T-cells per hour [13, 15, 21] as they move slowly through the paracortex. Many of these T-cells will not have affinity to the pMHC complexes on the DCs, and they will break contact with the DC after a few minutes [12, 13, 15, 20, 22]. Others will bind more strongly and form a so-called immunological synapse. The ligation of T-cell receptors (TCRs) causes receptors to move closer together (towards the contact area), and the ligation of a threshold number causes T-cell activation. In this case, the synapse will last for several hours [12, 13, 15, 20]; 'swarms' of T-cells have been observed bound to cognate antigen-presenting DCs [12, 13, 15, 20]. If antigen affinity is lower, T-cells can be activated through integration of many signals from these short interactions, referred to as 'kinapses' [23, 24]. In comparison to DCs, T-cells move much more quickly through the lymph node, with an average speed of approximately 10µm/min ± 20µm/min [18, 25, 26]. At long time scales, the movement resembles a random walk, but it is a more ordered 'run-and-tumble' at shorter timescales [13, 20, 22, 25, 27, 28]. T-cells must polarise before they can move and so will persist in one direction (or with small turning angles) for 20-30µm before they collide with something and stop. Then, they must repolarise before moving off again [13, 20, 22, 27, 29]. B-cells are activated by a similar process to T-cells, but they depend on antigen presented by follicular DCs. T-cells and B-cells require both a threshold number of cognate antigen to ligate their receptors and a costimulatory signal to become fully activated, otherwise they will anergise after a short period of effector activity and/or proliferation. T-cells' CD28 receptor must be ligated by B7 on an APC or a B-cell.

B-cells' CD40 must be ligated by CD40L on $CD4^+$ helper T-cells. Activated $CD4^+$ T-cells in the paracortex release chemokines that attract $CD8^+$ T-cells to the site of the DC-$CD4^+$ synapse [12, 13, 22, 25], as although $CD4^+$ T-cells bind to MHC-II, the DC is very likely to be presenting antigen specific to the same pathogen on their MHC-I receptors that the $CD8^+$ T-cells can bind to, too. Afterwards, $CD4^+$ T-cells move towards the perimeter of lymphoid follicles to provide costimulation to B-cells as APCs and B-cells do for $CD8^+$ cells; B-cells move periodically from the centre of the follicle (the 'dark zone' with a high density of B-cells) to proliferate and the 'light zone' to meet these $CD4^+$ T-cells. This process is repeated periodically because B-cells undergoing somatic hypermutation require repeated stimulation of their receptors to continue proliferation, to ensure that antigen affinity is increasing rather than decreasing.

When cells are fully activated, they begin to proliferate within the secondary lymphoid tissue (lymph nodes, Peyer's patches and the spleen). T- and B-cells with the highest affinity undergo the most proliferation, whilst those with lower affinity for their antigen leave for the sites of infection to begin an early response [25, 30]. Antigen presentation is not limited to secondary lymphoid tissue, however. There are two other types of APCs: macrophages and B-cells. Both are usually present within the site of infection with T-cells, and are able to present the latter with antigen and ligate CD28 receptors on T-cells to signal that infection is still ongoing, and to prevent T-cells becoming exhausted. $CD4^+$ cells provide the same service to B-cells in turn.

0.1.2 Lymphoid tissue structure

'Primary' lymphoid tissue refers to the sites where B- and T-cells are created and trained to not react to self-antigen, the body's own molecules. 'Secondary' lymphoid tissue includes the lymph nodes, Peyer's patches (special lymph nodes lining the intestines) and the spleen, and it is here that the innate and adaptive immune systems meet to activate the latter. I have already briefly discussed secondary lymphoid tissue above, but I have not detailed how its structure is related to its function.

Throughout the body, fluid continually leaks out of blood vessels to bathe cells and provide them with lubrication, temperature regulation, and nutrients and chemicals required for their function. This fluid must return to the blood, and it does so through lymphatic vessels lined with one-way valves. Pathogen or products of pathogens that have made their way into the extracellular fluid will also be forced through these vessels, so they are lined with macrophages that can destroy would-be invaders. Lymphatic vessels lead in turn to lymph nodes. Dendritic cells enter lymph nodes through these lymphatic vessels, too, whilst T- and B-cells enter from these afferent lymphatic vessels or from the bloodstream. Thus, T-

and B-cells are able to sample all antigen that drains or is carried from the tissues. Peyer's patches perform a similar function near the small intestine, sampling antigen extracted from the intestine itself. The spleen does this job for the bloodstream, acting as a filtering system. In this section, I will briefly describe each kind of secondary lymph tissue in turn and show how they are specialised for their functions.

0.1.2.0 Lymph nodes

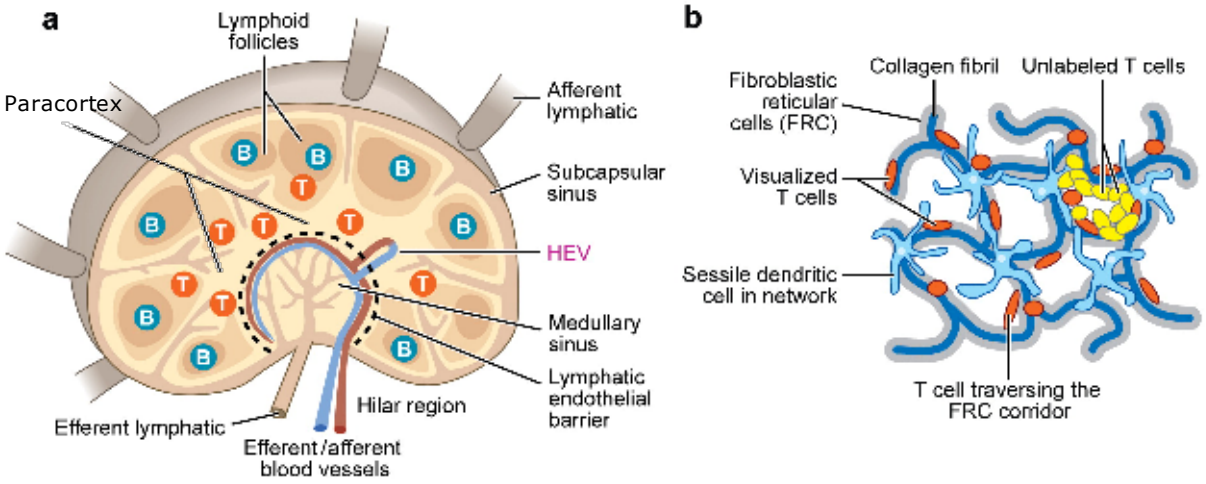


Figure 0.0. a) A schematic of lymph node structure. (B) and (T) indicate the respective areas occupied by B and T lymphocytes: lymphoid follicles and the paracortex. The afferent lymphatic vessels deposit dendritic cells, antigen and lymphocytes into the interstitial space surrounding the node, and the bloodstream deposits lymphocytes into the centre of the node. Cells exit via the efferent lymphatic vessel. **b)** The fibrous network in the paracortex, formed by fibroblastic reticular cells. This network may guide the motion of T-cells [13, 15, 20, 25, 27, 29, 31]. An analogous network exists in lymphoid follicles for B cells. Figure modified from Cahalan *et al* (2008) [13] within permissions.

The outer arc of the human lymph node is divided into sections by radial medullae (see Fig. 0.0) [7, 13]. The adjacent spaces are called lymphoid follicles, and the central region is the paracortex. The follicles and paracortex respectively house B- and T-cells. Dendritic cells (DCs) enter through the surrounding interstitial space after exiting afferent lymphatic vessels, whilst B- and T-cells extravasate from the high endothelial venules at the centre of the node, or the lymphatic vessels after exiting an earlier lymph node. All cells exit the lymph node via the efferent lymphatic vessel. Within each follicle is a network of follicular DCs and a B-cell stroma on which B-cells migrate. The paracortex is occupied by a similar network of T-reticular cells associated with DCs and T-cells. These networks express chemokines (CCL19 and CCL21 for T-cells and CXCL13 for B-cells) that influence the motion of lymphocytes to improve the efficiency of DC-lymphocyte contact [13, 15, 20, 25, 27, 29, 31], by raising their crawling speeds. These same chemokines recruit cells to the correct locations: T-cells and dendritic cells are guided towards the paracortex and B-cells

are guided towards B-cell follicles. As described above, however, $CD4^+$ cells and B-cells must meet each other for full activation, and the design of the lymph node facilitates this: T- and B-cells move radially outward after leaving the HEVs and no B-cell follicle is a large distance away from the paracortex. Lymph flow in Figure 0.0a runs from top to bottom, and will sweep out loose material or cells that are no longer subject to chemotaxis. In the case of T-cells, for example, they are drawn to the paracortex by interactions between lymphocyte CCR7 and lymph node stroma-derived CCL19. Under normal conditions, the rates of ingress and egress of cells in and out of the lymph node are equal, as T- and B-cells search for their cognate antigen. When T- or B-cells have found antigen that they have affinity for and begin to proliferate, the propensity for each cell to exit the lymph node is reduced and the lymph node swells dramatically in size.

0.1.2.1 Peyer's patches

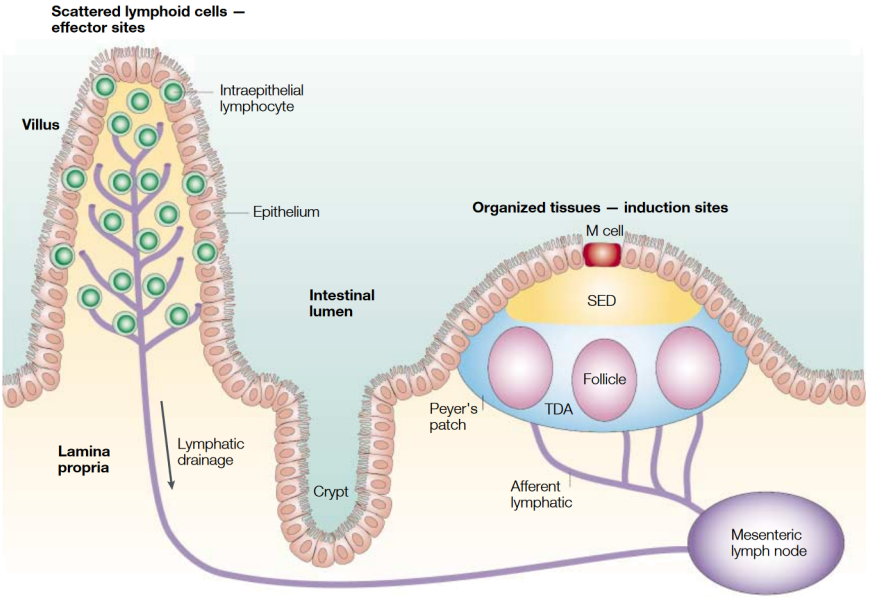


Figure 0.1. Schematic of a Peyer's patch, some lymphoid follicles and the draining lymph node. The M cell samples antigen from the intestinal lumen and passes them through the subepithelial dome (SED) to the Peyer's patch. Like a lymph node, there is a T-cell area (TDA, thymus-dependent area) and B-cell follicles. There are no afferent lymphatic vessels to the patch, but there are efferent lymphatic vessels. Figure reproduced from Mowat (2013) [32] within written restrictions.

Peyer's patches are very similar to lymph nodes but there are several key differences (see Figure 0.1). They have no afferent lymph vessels, and instead there are 'M-cells' adjacent to the lumen of the small intestine. These cells sample antigen from the intestine and release them into the subepithelial dome. Lymph from this tissue drains into the Peyer's patch, where antigen is presented to T- and B- cells after opsonisation or collection by dendritic cells. Like a lymph node, there are lymphoid follicles for B-cells and follicular dendritic cells,

and a T-cell area for T-cells and dendritic cells. There are also high endothelial venules to admit T- and B-cells, and efferent lymphatic vessels to allow lymph and cells to drain away to other lymph nodes of the lymphatic system.

0.1.2.2 The spleen

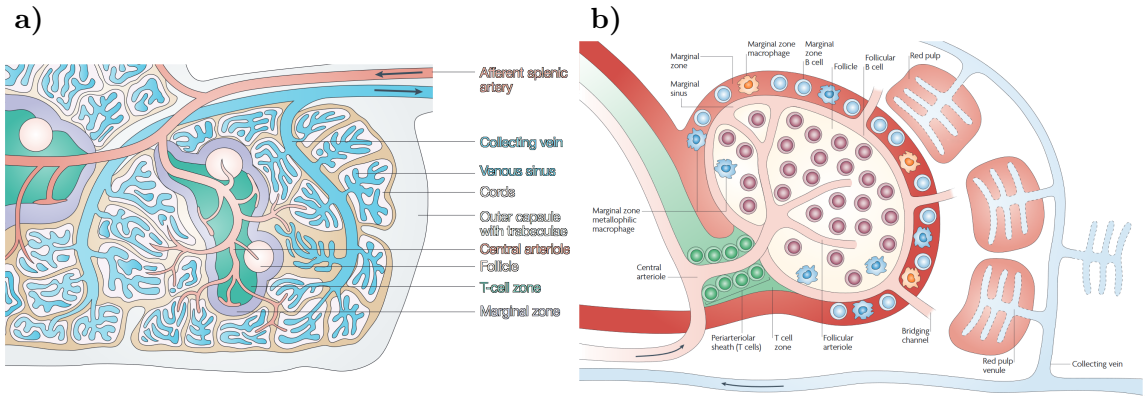


Figure 0.2. Schematics of the spleen. **a)** a 'zoomed-out' view of the tissue structure, showing the main afferent artery and efferent vein. The artery splits up into many arterioles, which pass through T-cell zones and deposit into follicles. The blood is then drained away by venous sinuses into collecting veins. **b)** a closer view of a single follicle, draining a single arteriole. T-cells are present around the arteriole and B-cells are present in the follicle and the surrounding marginal zone. Figures modified from **a)** Mebius and Kraal (2005) [33] and **b)** Shiv and Cariappa (2009) [34] within written restrictions.

Just as lymph vessels all pass through the lymph nodes, so that the immune system may screen the flow of all lymph, 5% of blood passes through the spleen every circulation, allowing the adaptive immune system to filter the contents of the blood. A schematic of the spleen is shown in Figure 0.2. Blood enters through the splenic artery and immediately splits up into multiple arterioles for each follicle. Blood passes through the arteriole and is deposited into a marginal sinus, before collection by the splenic vein. T-cells in the blood are pushed out into the periarteriolar sheath around the arteriole and B-cells into the follicle within the marginal sinuses. Within the marginal sinuses are (i) macrophages which capture pathogens and (ii) dendritic cells which collect bloodborne antigen. If the dendritic cells acquire antigen, they travel to the periarteriolar sheaf to activate T-cells, and antigen of course mixes freely with B-cells as blood enters the follicle. As in the lymph node, $CD4^+$ T-cells activated by the dendritic cells travel to the follicle to provide activating signals to B-cells. Also as in the lymph node, after some time in the spleen, T- and B-cells will leave, unless they become activated and begin to proliferate.

0.1.3 T- and B-cell trafficking and homing

T- and B-cells enter lymph nodes and other tissues by extravasating through the endothelium of venules. Lining the membranes of T- and B-cells are molecules known as selectins.

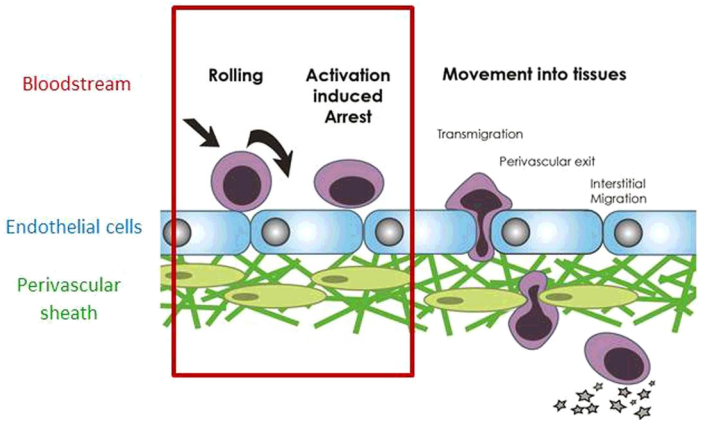


Figure 0.3. Schematic showing the process by which T-cells roll, stop and extravasate into tissues. Shown in the red box are the rolling and stopping stages; these are dependent on selectin and integrin ligands on T-cells and receptors in endothelial cells. Figure reproduced from Ager *et al* (2016) [35] within written restrictions.

As the cells move through venules in tissues, they may come across a partner molecule expressed by the endothelium. The interaction between these molecules causes them to slow down and begin to roll, and to express other molecules known as integrins that were pre-packaged within vesicles in the cytoplasm. Should these integrins also find their matching partner on the endothelial wall, then the cell comes to a complete halt and is able to transmigrate through the endothelial wall and into the tissue (Figure 0.3).

Naïve T- and B-cells of course localise mainly in the lymph nodes, Peyer's patches and the spleen, as they circulate looking for antigen. Activated T-cells have a higher propensity for outer tissues; for example, T-cells activated due to an infection near mesenteric organs will preferentially home to these organs. Activated B-cells do not need to engage cells in the infection site directly and so instead localise mainly in the lymph nodes and bone marrow, from where they produce antibodies to fight the infection. This specific homing is achieved through the specific selectins and integrins expressed by the T- and B-cells. For example, the propensity of naïve T-cells to localise in lymph nodes is due to their expression of L-selectin (CD62L), which binds with GLYCAM1 on the high endothelial venules of lymph nodes and Peyer's patches. E-selectin and P-selectin can be expressed by some (but not all) activated T-cells to allow them to extravasate into skin tissues. Other selectins and integrins allow T-cells to enter other tissues. The specific selectins and integrins up-regulated by an activated T-cell depend on molecules presented to it by the antigen presenting cell that activated it, as shown by previous studies [36] and supported by localisation data obtained by radio-labelling T-cells [37, 38].

0.1.4 T-cell cytotoxicity

Cytotoxic $CD8^+$ T-lymphocytes (CTLs) bind to target cells in a process similar to the way that they are activated, and kill in a process very similar to NK cells. They 'sample' many cells in the environment until they come upon the peptide–MHC-I (pMHC-I) complex that is cognate to their TCRs. As with dendritic cell binding, many TCRs and pMHC-I molecules are subsequently brought close together to form an immunological synapse. The synapse then triggers target cell death with one of two mechanisms. T-cells either ligate Fas protein on the target cell membrane, triggering the target cell to commit suicide, or they deliver cell-destroying enzymes such as granzyme B, by utilising perforin to either form a hole in the cell membrane, or to break open a vesicle after its delivery into the target cell. This process is completed without harm to bystander cells [24, 39]. *In vitro* studies have shown this process to be very fast [40], lasting a few minutes at most. *In vivo*, however, T-cells must search for target cells within the tissue before they can kill, which takes extra time. Halle *et al* [41] found that 2-16 virus-infected cells can be killed by each CTL per day, that not all contacts between a CTL and an infected cell lead to cell death, but that CTLs can work together to increase the probability of cell death. Vasconcelos *et al* [42] found that cell killing takes a few minutes (up to around 30), and that CTLs can kill cells simultaneously or serially, but that individual CTL activity comes in bursts.

There are also various mechanisms by which CTL killing can be suppressed. The more targets that CTLs kill, the more they upregulate Fas protein on their own membranes. This protein can then be ligated by their own Fas ligands or by those on other CTLs, resulting in their own deaths. After activation, T-cells begin to upregulate CTLA-4 in response to ligation of their TCRs, particularly if they are co-stimulated by B7-1 (but not B7-2) [43]. CTLA-4 binds to B7 on dendritic cells and macrophages with higher affinity than does CD28, the molecule that reinvigorates activated CTLs. Ligation of CTLA-4 reduces cytokine production and proliferation and causes TCR signalling to be attenuated, which reduces T-cell–APC contact time and hence inhibits their activation [44]. Programmed death receptor 1 (PD-1) is another receptor expressed by CTLs in response to repeated antigen stimulation [45]. As CTLs kill more cells, they express more PD1. The ligand to this receptor, PD-L1, is expressed on antigen presenting cells. Binding of PD-1 causes exhaustion of the T-cells to prevent cytolytic function. Exhausted T-cells stop producing chemokines (IL-2, TNF and IFN-γ), stop proliferating and lose the ability to kill cells [45].

0.1.5 The cancer-immunity cycle

The immune-system also has a key role in combating cancer, through many of the mechanisms used to fight pathogenic infections. It has been proposed that cancer and the immune

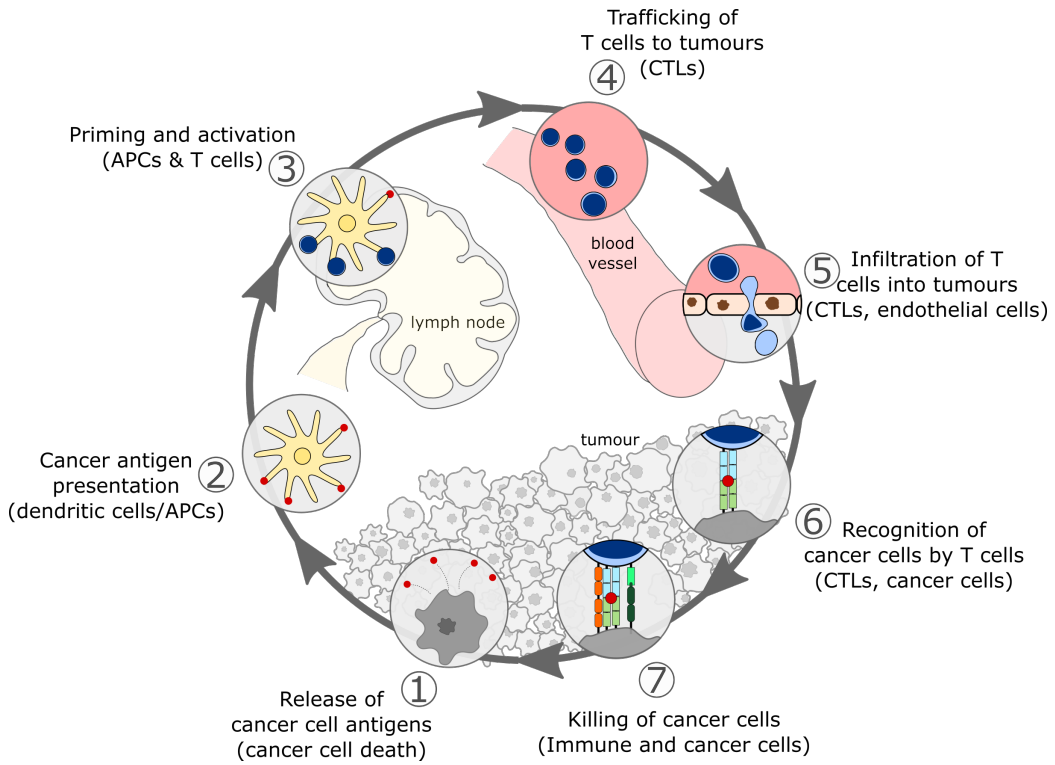


Figure 0.4. The cancer-immunity cycle [5], reproduced with permission. It comprises seven major steps beginning with release of cancer cell antigen and ending with the killing of cancer cells by the immune system. 'APC' is 'antigen presenting cell' and 'CTL' is cytotoxic T lymphocyte. Reproduced from a published figure [5].

system interact through the 'cancer-immunity cycle' [5], Figure 0.4. The cycle comprises of seven steps, beginning with **(Step 1)** the release of cancer antigen through an event such as cell death. **(Step 2)** These molecules are collected by dendritic cells, which migrate to lymphoid tissue such as lymph nodes if they recognise stress signals such as inflammatory markers, IFN-γ or pathogen-associated molecules. **(Step 3)** Once they reach lymphoid tissue, dendritic cells present pMHC-I to T-cells as normal and activate cognate $CD8^+$ cytotoxic T-cells. **(Step 4)** T-cells traffic across the bloodstream towards the site of the tumour and then **(Step 5)** extravasate into the tumour parenchyma, depending on the selectins and integrins on the surfaces of the T-cells and blood vessel endothelium. **(Step 6)** The cytotoxic T-cells then seek out cells displaying their cognate antigen, bind and **(Step 7)** kill the tumour cells. This induces the release of further antigen, which starts the cycle anew by either strengthening the previous response (if the same antigen is collected) or causing the activation of a new clonotype of T-cells (if a different antigen is collected), a process known as epitope spreading.

If a patient has a tumour, it follows that at least one of the steps in this cycle must

have halted or slowed sufficiently to become rate-limiting. There are processes that are known to affect every step. **(Step 1)** Tumour cells descend from healthy cells, so if they do not produce stress signals and have not acquired mutations that cause them to produce unusual peptides, there may in principle be nothing that can provoke a response from the immune system. Likewise, if only some tumour cells express a recognisable antigen, these cells would be killed and the cells that do not present it would survive and reproduce. Thus, epitope spreading and the multiple resulting 'waves' of attack are anticipated to be essential for tumour clearance. **(Step 2)** Some tumours are not inflamed or do not present any markers that may cause antigen-presenting cells to become activated, reducing the number of antigen that the cells can carry and their propensity for migrating to the lymph node. **(Step 3)** T-cells may not be presented a sufficient number of cognate antigen to be activated, or the numbers of interacting T-cells and DCs may be too small for there to be a significant likelihood of interaction. Furthermore, tumours may secrete factors that inhibit the activation of T-cells in tumour-draining lymph nodes. **(Step 4)** Tumour vasculature is often very poor; if T-cells cannot reach a tumour due to poor or non-existent vasculature, then they will not be able to have any effect. For example, vasculature that is tortuous and leaky lacks a strong interstitial pressure gradient from the vessel to the tissue, preventing the departure of T-cells from the vessel [46, 47]. A related issue exists for the tumours located in tissue that is poorly perfused (like the skin), where far fewer T-cells can reach the tumour per time than in a more perfused tissue (like the kidney). **(Step 5)** The ligands to T-cell selectins and integrins expressed by the endothelium of blood vessels are required for extravasation. If these are down-regulated or missing in the tumour vasculature or neighbouring tissues, then even if the vasculature is robust (not leaky or undeveloped), the T-cells will not extravasate. **(Step 6)** If only a subsection of tumour cells present recognisable antigen, then after T-cells have killed all of the cells presenting recognisable antigen, the cells without recognisable targets would survive and reproduce unharmed. **(Step 7)** Even if recognisable antigen are present, tumour cells may begin to express PD-L1, which will bind to PD-1 on T-cells and anergise them before they can eradicate the tumour.

0.1.6 Immuno-oncology and immunotherapy[1]

Immunotherapy is the use of interventions to (re-)invigorate the immune system and direct it against tumours. Many seek to remove problems such as those described for each step

[1]This subsection follows a ninety-page internal report commissioned by Roche in 2015 on cancer immunotherapies. It briefly reviews the science of each class of cancer immunotherapy, but focuses upon describing therapies currently on the market, in trials or that have failed. I outlined the same broad classes and cited some of the same references, but did not repeat text from the report.

above. One such example is 'checkpoint inhibition', the direction of antibodies against PD-(L)1 or CTLA-4, *i.e* to unblock steps 3 or 7 above. Drugs approved for use as checkpoint inhibitors include Ipilumab (anti-CTLA-4), Nivolumab and Pembrolizumab (both anti-PD-1). All of these have been approved for melanoma and each is approved or in trials for other indications such as lung, prostate, and head and neck cancers. Another class of therapies are immunostimulants. One study [48] produced a drug that strongly stimulates CD28 (receptor for B7) on T-cells, such that the T-cells become activated without the need for antigen (to unblock step 1 above). This, however, produced potent autoimmune toxicity in patients and the trial had to be abandoned. Other therapies introduce G-CSF, GM-CSF, TNF-α IFNs, IL-2, IL-12, IL-11, IL-15 or IL-18 to induce anti-tumour activity in the immune system. IL-2, for example, is approved for melanoma and renal cell carcinoma (RCC) because it increases $CD8^+$ and natural killer cell proliferation, and antibody production by B-cells [49]. IFN-α is also approved for melanoma, because it is anti-angiogenic, promotes activation of natural killer and dendritic cells, and inhibits cancer growth by promoting cell death [50, 51]. These therapies target steps 3 through 5. Other therapies make use of oncolytic viruses, with the aim to infect tumour cells with the virus and result in 'tumour' (viral) antigens being presented by the cells, targeting step 6. Viral DNA or RNA vectors can also be used for this purpose. Cytokine and viral therapies such as this aim more generally to make a 'cold' tumour (one that is uninflamed and has little immune activity) 'hot' (immunoactive). A 'cold' tumour may well have macrophages or dendritic cells present that can collect unusual tumour antigen, but with no typical 'signals' that an invasion is occurring – such as bacterial or virus-specific molecules that ligate pattern recognition receptors on the macrophage – then they will not hyperactivate and attack the cells nor migrate towards secondary lymphoid tissue, *i.e.* step 2 is blocked. Making the tumour 'hot' solves this issue, and the activation of antigen presenting cells may in turn then cause the activation of the adaptive immune system and drive tumour clearance. 'Cold' tumours may be furthered divided into those that are ignored by the immune system, where no tumour-specific immune cells are present at all (step 4 is blocked), and those that are 'immune deserts', where immune cells have arrived but are confined to the tumour periphery (step 5 is blocked). Finally, some therapies aim to improve tumour vasculature; anti-angiogenic therapies often result in 'normalisation' of vasculature. Normalisation refers to the pruning of errant vessels and the redevelopment of vasculature structure, which results in a stronger and more reliable interstitial pressure. This can facilitate extravasation and infiltration by T-cells (unblock step 5).

Some studies have attempted to produce cancer vaccinations, in which cancer antigen and immunostimulants are introduced simultaneously, to evoke an immune response without

having to make the local tumour microenvironment 'hot' (which may be infeasible). These antigen could be antigen that are common to many patients of the disease or be formed from the patient's own tumour cells, *e.g.* by irradiating cells *ex vivo* before reintroduction as a vaccine. An example of the former approach is the clinical trial of IMA901 [52, 53], in which 9 peptides common to renal cell carcinoma (kidney cancer) were introduced to patients with the disease. The trial was successful (associated with 'better disease control' [52] or longer overall survival) in phases I and II but not phase III. There have been many similar trials, but as stated in Section 0.0, only 14 of 167 phase II and III trials of cancer vaccines met their primary endpoints. The potential reasons behind such results, particularly for IMA901, motivated the work presented in Chapters 1 and 2 (see Section 0.3).

Another approach is to introduce not tumour antigen but tumour-targeting T-cells. Tumour invading lymphocytes (TILs) can be extracted from within a tumour, made to proliferate *ex vivo* and then reintroduced to the patient. The rationale behind this is that these TILs are able to target the tumour, but may be exhausted or simply not numerous enough. Increasing their numbers may tip the balance against the tumour. T-cell Bispecifics (TCBs) are antibodies that have been engineered to be specific to both CD3 on TCRs and to a tumour antigen, *e.g.* CD19 in some lymphomas and leukaemias. This allows potentially all T-cells to be drawn towards tumour cells and, once both parts of the antibody are ligated, their TCRs to be strongly engaged. This results in non-specific killing of tumour cells and T-cell proliferation and, critically, recruits any T-cell in an antigen-independent fashion. The downside is that the antibodies must be specific for an existing tumour antigen. Chimeric antigen receptor (CAR) T-cells are a similar technology. These are T-cells with bioengineered TCRs that are specific for a chosen tumour antigen and can bind independently of MHC presentation, which prevents tumour cells from hiding from the immune system by this route [54]. There have been multiple 'generations' of CAR T-cells; second and third-generation CARs included co-stimulatory molecules such as CD28 and 4-1BB to promote proliferation and inhibit exhaustion, whilst third- and fourth-generation CARs improved effector functions, even including enzymes that can degrade tumour extracellular matrix [54]. The cells are especially promising because of their improved capacity to proliferate and carry out effector functions over endogenous T-cells, and have proven efficacy in some lymphomas and leukaemias. For example, relapsed B-cell acute lymphoblastic leukaemia (B-ALL) patients have a median survival of 7 months [55, 56], but following CAR T-cell therapy, more than 80% achieved complete remission and 50-86% were alive 1 year later [55]. Those that do relapse typically do so due to loss of the target antigen, or rejection of CARs [55]. CAR T-cells have a dangerous side-effect profile, including cytokine release syndrome (where the release of cytokines by CAR T-cells over-activates the immune system

and leads to organ dysfunction and failure), CAR T-cell-related encephalopathy syndrome (with symptoms ranging from mild confusion to death) and tumour lysis syndrome (where tumour cells die so rapidly that the body's metabolism is affected by the influx of chemicals) [54, 55]. Current research is investigating the use of 'smart' T-cells that can be compelled to commit suicide or synthetically controlled [54], to alleviate these symptoms.

These therapies all have drawbacks and challenges associated with them. Many studies consider combinations of therapies, but these often carry a greater incidence rate of adverse events. For example, the efficacy of TILs is improved when they are introduced in combination with checkpoint inhibition, but this can then increase the rate at which TILs and other T-cells attack cells in inappropriate locations (on-target but off-tumour). A typical result for any therapy is that it is very effective in some patients, partially effective in other patients, and has no effect in still other patients. Invoking the cancer-immunity cycle, this may be because some of these patients have a blockade in the step that a given therapy targets, but other patients may have a blockade elsewhere. For antigen-specific therapies, it may simply be the case that some patients' tumours do not express the antigen being targeted. It is challenging to determine which of these processes could be the problem in any given patient, or to understand what limits the efficacy of a therapy in a patient population. In order to determine the limits and variability of the efficacy of potential therapies, it is important to quantitatively understand the underlying processes. To pursue the objectives of this thesis, we have produced mechanistic models of various processes critical for T-cell mediated tumour clearance.

0.2 Mathematical and computational modelling in immuno-oncology

Data-driven techniques such as machine learning have been used extensively in the biological sciences to generate predictions and statistical analyses of efficacy. However, the hypotheses and predictions they generate are generally abstractions, patterns, associations or correlations, and are less well-suited than mechanistic modelling to elucidate biological processes or mechanisms underlying a system. For example, statistical analysis of trial data may indicate that older populations are less likely to respond strongly to a 'cancer vaccination' and be able to predict response based on factors such as age, but a more mechanistic model would be required to elucidate the reason as to why age affects vaccine response, for example due to immune cell counts or affinities. In the following section, I will discuss examples of mechanistic modelling of tumour kinetics and lymph node biology that have yielded biological insight or impact, based upon a review with a wider scope than is necessary for this thesis, published in the Journal of Experimental and Clinical Immunology (Liam V. Brown

et al 2018 [57]). Tumour kinetic modelling includes several exemplary models which have generated biological insight in the field of oncology, and modelling of lymph node biology will be particularly important in Chapters 1 and 2. The models reviewed in these sections are summarised in Tables 0.2.0 and 0.2.1. This will be followed by a discussion of models of pharmacokinetics – how the concentration of drugs or therapeutics in the body changes over time – before I show how this methodology may be used to consider localisation of T-cells in different tissues in the body, which will be important for Chapters 3 and 4. I will end by considering age-structured PDEs, a methodology that will be utilised in Chapter 4.

0.2.0 Models of tumour kinetics

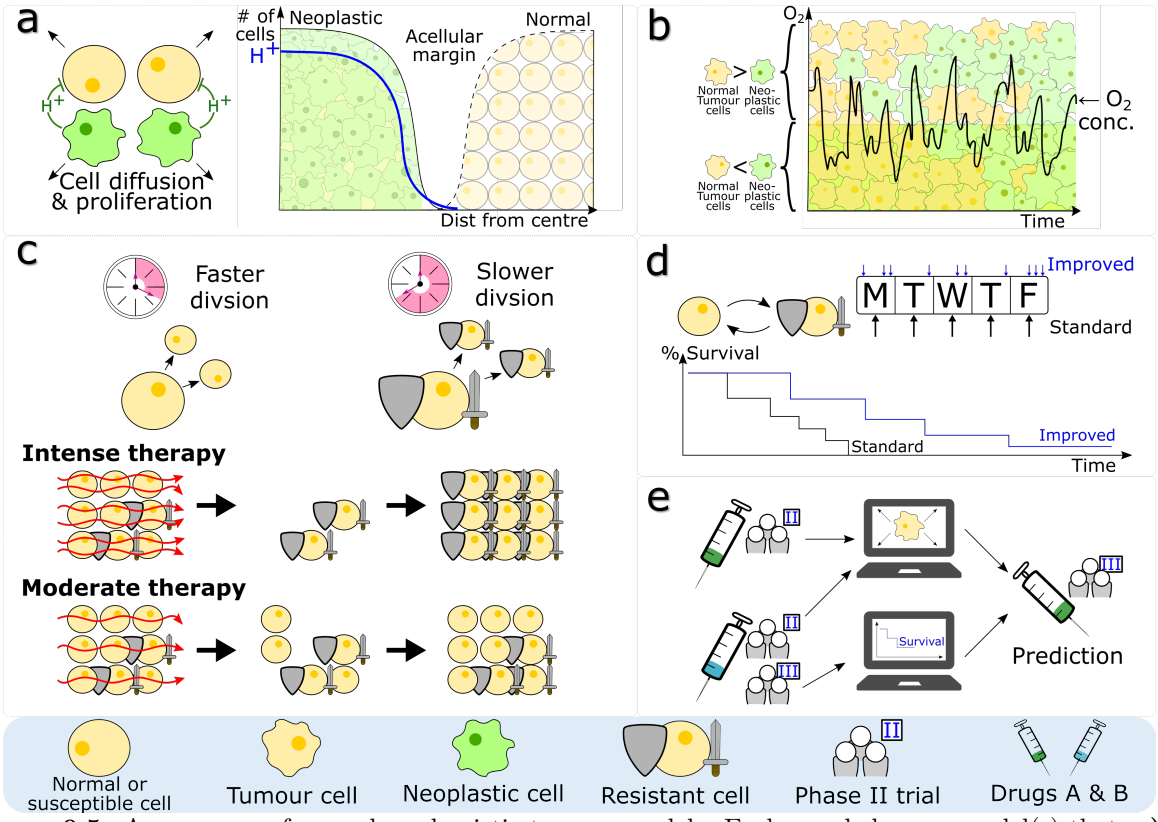


Figure 0.5. A summary of several mechanistic tumour models. Each panel shows a model(s) that, **a)** predicted various features of tumour invasion mediated by acid-producing cells [58]; **b)** provided an evolutionary basis for the Warburg effect, as neoplastic cells that out-compete other cells in low oxygen concentrations (green shaded region) have a fitness advantage in an environment with variable oxygen concentrations (shown by the black line) and take over the tumour population (background of plot) [59]; **c)** showed how moderate therapy improves patient survival over intensive therapy, that may select for resistant cells [60]; **d)** showed that a less intense, more frequent therapy schedule improves survival in a dynamic model of resistance [61]; **e)** predicted patient survival in a phase III clinical trial of a drug by parameterising tumour and survival models [62].

Mechanistic modelling has deepened understanding of cancer disease biology and responses to treatment (*e.g.* radiotherapy), and informed clinical trial designs (timing, dose,

Authors	Model aim	Model type	Figure
Gatenby *et al* [58]	Identify mechanism for cancer invasion	PDEs	0.5a
Gravenmier *et al* [59]	Identify evolutionary mechanism for Warburg effect	Discrete update equations	0.5b
Gatenby *et al* [60]	Determine how to treat cancer with radio/chemotherapy without selecting for resistant populations	Analytical and discrete update equations	0.5c
Leder *et al* [61]	Improve efficacy of therapeutic schedule by accounting for dynamic radioresistance of Glioblastomas	Integral equations	0.5d
Claret *et al* [62]	Predict anti-tumour response and overall survival in phase III studies using phase II data	ODEs and Kaplain-Meier survival analysis	0.5e
Norton *et al* [63]	Derive equations to describe growth curves of tumours with/without therapy	ODEs	
Claret *et al* [64]	Investigate predictive performance of tumour kinetic model parameters for overall survival of clinical patients	Multivariate equations, Kaplain-Meier and Cox regression analyses	

Table 0.0. A summary of the models presented in Section 0.2.0, showing both the aim and type of the reported models. The last column gives the label for any model presented graphically.

and the window between efficacy and toxicology). Using tumour kinetic modelling for illustration, I will present a few examples that highlight how mechanistic modelling can be used to generate insight where an experimental model could not have easily done so, as this is the approach that we have aimed for within our own modelling studies. Models have enhanced mechanistic insight into tumour progression and metastasis, and the role of microenvironmental pH in the invasion of surrounding tissue (Figure 0.5a). One such model [58] used a system of PDEs (three coupled reaction-diffusion equations of the densities of normal and neoplastic tissue, and H^+ ions) to predict the existence of a pH gradient from the tumour periphery into healthy tissue, which was hypothesised to induce an acellular gap around the tumour and the remodelling of tissue to facilitate tumour invasion. These hypotheses have been subsequently supported by experimental studies [58, 65]. A more recent model [59] used a pair of discrete update equations for competing tumour populations to propose a mechanism for the Warburg effect to arise even in oxygenated tumours, despite the apparent loss of fitness (Figure 0.5b). The two populations are Warburg and Pasteur populations, the former of which has a greater tendency to respire anaerobically than the latter. Discrete update equations are used instead of continuous differential equations to suit the stochastic

nature of the system: the tumour populations compete with Lotka-Volterra dynamics, but their carrying capacities depend on the balance of the respiration methods they use and their metabolic efficiencies, which in turn depend on the concentration of oxygen, which is updated with a random walk every timestep. The Warburg effect arises in the model as a 'bet-hedging' strategy in response to variable oxygen concentration. These examples illustrate how mechanistic hypotheses can be used to develop, explore and refine new concepts *in silico*, to inform subsequent definitive experimental studies.

Models have been used to improve human tumour therapy scheduling and to develop new concepts of how best to treat tumours. An example of such a concept is that increasing treatment strength does not necessarily improve patient survival, but rather counter-intuitively could potentially decrease patient survival. By assuming that cancer cells 'asymptotically approach a minimum information state necessary to maintain proliferation', one such model [60] defined the growth rate for each sub-population of cells as,

$$p_i(t) = C_i(t/t_0)^{\gamma_i \Phi} \cos^2\left(a_i\left(t\right) \ln\left(t/t_0\right)\right),$$

where $p(t) = \sum_{i=1}^{n} p_i(t)$ is a probability density function of observing the cancer in time $(t, t + \mathrm{d}t)$, i is an index for a sub-population of the cancer, C_i and t_0 are constants, $\Phi = 1.618...$ is the Fibonacci constant and $a_i(t)$ is the time-dependent therapeutic activity. In the absence of therapy, this is a power-law, whilst 'effective' therapy reduces p_i to zero over time. The authors presented analytical and computational results to predict that maintaining enough live radiosensitive cells is important to out-compete slower dividing but therapy resistant cells, thus improving patient survival. This hypothesis was subsequently confirmed in mice (Figure 0.5c). A similar conclusion was determined by modelling a heterogeneous tumour with a linear quadratic model (standard radiotherapy response equation) and dynamic radiation resistance [61] (Figure 0.5d). The authors modelled the number of cells in each dynamic state not with ODEs, but by assuming exponential growth of cells impacted by treatment (*i.e.* $N(t) \sim N_0 e^{\alpha d - \beta d^2}$, where d is the dose and α and β are positive constants) and dynamical switching of cells between sensitive and insensitive states. The total number of cells is formulated as an integral over time, considering growth, death and switching rates. The predicted improved schedule derived from this model, designed to avoid heavily irradiating a resistant population, yielded better survival compared to the standard radiotherapeutic schedule for both *in silico* and mouse models. This therapeutic strategy is currently being assessed in a clinical trial [66].

One class of models that have increasingly supported clinical trial designs are tumour kinetic (TUK) models. These models describe early (*e.g.* 8-week) patient-level changes in tumour burden over time and in response to treatment. The resulting models are used to

generate patient-level metrics such as estimated on-treatment tumour growth rates. Multivariate survival models may then be developed that link baseline patient characteristics and model derived metrics, such as tumour growth rates, to longer-term overall survival (OS) [62, 64]. Claret *et al* (2009) [62] used a TUK model to successfully predict the overall survival of patients receiving a drug in a phase III clinical trial, by developing a survival model from phase III trial data for a different drug and a TUK (ODE) model from phase II data for both drugs (Figure 0.5e). In this case, the ODE was simply $\frac{\mathrm{d}N(t)}{\mathrm{d}t} = rN(t) - dE(t)N(t)$, where N is the number of tumour cells, r is the growth rate, d is the death rate E is the drug exposure. Parameters were fit to log normal distributions and the survival time of patients as a function of parameters and covariates was fit with various survival models (*e.g.* Weibull). The fitted model was then used to predict the survival times in the phase III trial. TUK models have also been used to inform phase III clinical trial designs, or to inform clinical dosing regimens, as originally considered by Goldie and Coldman [67, 68]. For example, a system of two ODEs [63] was used to demonstrate that tumours that become small in response to therapy may become resistant, so would require dense, intensive therapy to prevent remission. This observation was "contrary to popular belief" at the time, but was later confirmed by clinical studies to improve survival. TUK models have also been recently applied to immunotherapy. Standard early oncology endpoints (objective response rate, progression free survival) are poor predictors of OS for immunotherapies. TUK metrics were recently explored as a potential surrogate endpoint for OS. Claret *et al* (2017, 2018) [64] used a tumour kinetic model and survival modelling similar to their previous work to explore the predictive performance of TUK metrics, using early data from clinical trials of the checkpoint inhibitor atezolizumab (anti-PDL1) in patients with non-small cell lung cancer. Their results were promising, suggesting that such approaches may be used to inform the design of immunotherapy clinical studies [64]. Mechanistic models in human cancer offer massive potential to impact on clinical development; modellers have started to create a plethora of models focused on non-immune processes [62, 69] which has recently been extended to model anti-tumour immune responses [64].

0.2.1 Models of the lymph node

In Chapters 1 and 2, I will describe the formation and utilisation of a model of T-cell activation in the lymph node. In this section, I will review previous mechanistic models of the lymph node that have complemented experimental medicine studies to drive a deeper mechanistic understanding than was possible by experimentation alone. One such model was used to explore whether chemokines could attract T-cells to antigen expressing dendritic cells [20] (Figure 0.6a). The authors developed a cellular automaton with naïve

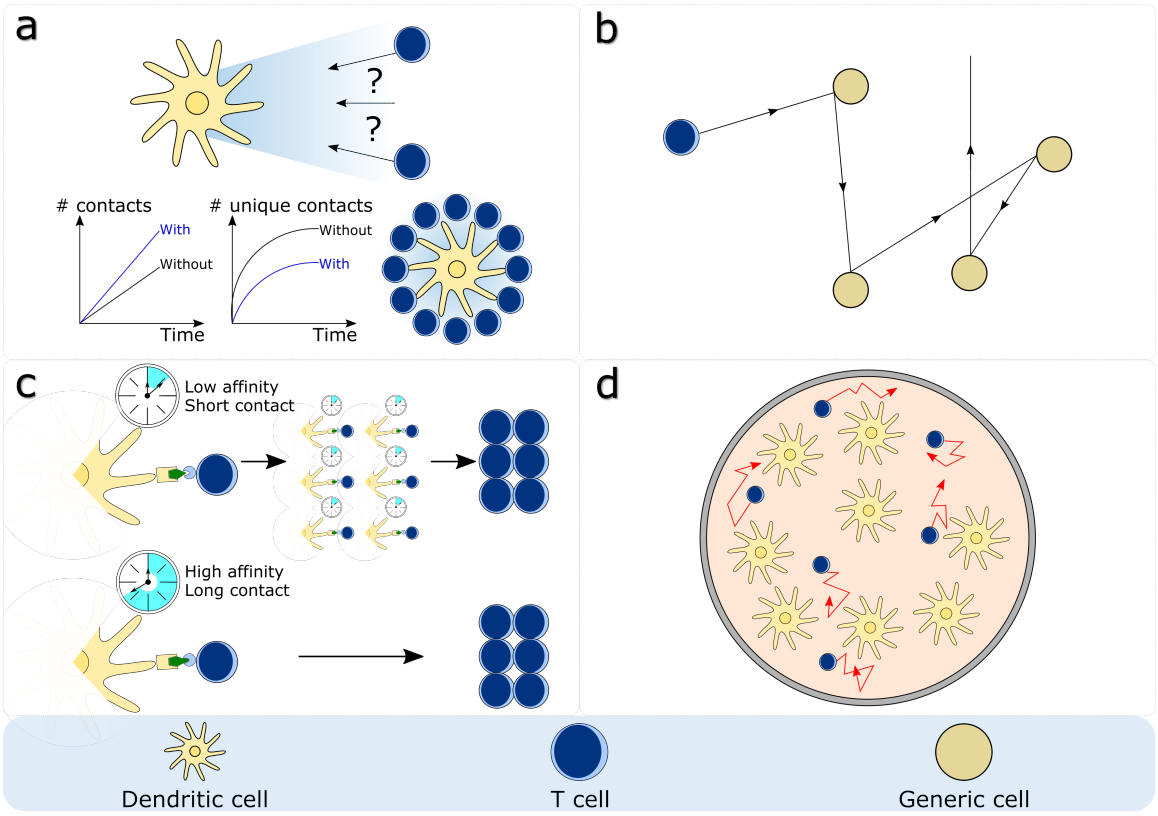


Figure 0.6. A summary of several mechanistic models of T-cell activation and dynamics. Each panel shows a model that, **a)** found that chemokines cannot attract naïve T-cells to antigen-bearing dendritic cells in an optimum search strategy [20]; **b)** showed that the dynamics of naïve T-cell movement can be explained entirely by interactions with their environment (as opposed to *e.g.* chemokines) [22]; **c)** investigated how naïve T-cells can integrate many low-affinity interactions with dendritic cells to activate [70]; **d)** considered the minimum number of dendritic cells required for T-cell response [19].

T-cells moving within the lymph node, based upon their own two-photon microscopy of T-cell/dendritic cell dynamics. Modelled T-cells choose a new direction every timestep. This was either random, weighted towards small angles (*i.e.* persistent direction of motion), or weighted towards nearby dendritic cells (chemotaxis). With small angles, the model reproduces observed short-term ordered motion and long-term random walks, but they found that an optimum search strategy cannot include chemotaxis to dendritic cells, as although it increases the number of T-cell–dendritic cell contacts per unit time, it sharply reduces the number of *unique* contacts due to crowding around dendritic cells. This leads to smaller and delayed numbers of activated T-cells.

A model has been used to determine whether the 'run-and-tumble' motion exhibited by T-cells (as opposed to diffusion) is due to an intrinsic, stochastic process, or whether it can be explained by the lymph node environment [22], and was developed in parallel with 4D multi-photon imaging experiments to validate model predictions (Figure 0.6b). The authors implemented the system as a Cellular Potts model of naïve T-cells and DCs.

Authors	Model aim	Model type	Figure
Riggs *et al* [20]	Reproduce T-cell–dendritic cell dynamics; determine whether chemokines are used to attract T-cells to dendritic cells	ABM (2D cellular automaton)	0.6a
Beltman *et al* [22]	Determine if run-and-tumble motion of T-cells is due to internal dynamics or their environment	ABM (3D Cellular Potts)	0.6b
Graw *et al* [31]	Determine effect of fibroblastic reticular cell network on T-cell movement and dynamics	ABM (3D cellular automaton)	
Bogle *et al* [71]	Reproduce cellular dynamics and densities consistent with two-photon microscopy	ABM (3D cellular automaton)	
Gong *et al* [72]	Predict level LN output in response to infections	ABM (3D cellular automaton)	
Moreau *et al* [70]	Determine dynamics of T-cell activation by 'kinapses' (many short, low-affinity contacts)	ABM (3D cellular automaton)	0.6c
Bouchnita *et al* [73]	Find suitable methodology for hybrid modelling of cellular, intracellular and extracellular chemokine dynamics	ABM (2D cellular automaton), ODEs and PDEs	
Celli *et al* [19]	Find minimum number of cells for immune response	ABM (3D off-lattice)	0.6d

Table 0.1. A summary of the models presented in Section 0.2.1, showing both the aim and type of the reported models. The last column gives the label for any model presented graphically. ABM: Agent-based model. LN: lymph node.

Modelled T-cells have a polarity aligned in the direction of motion. When a T-cell collides with a cell, it must realign this axis before moving again. With this requirement, the 'stop-and-go' motion can be explained entirely by collisions with other cells. The model produces a realistic T-cell motility (50-100$\mu m^2 min^{-1}$) and mean free movement period (around 2 minutes). The model produces streams of multiple T-cells following the same path without having to consider the lymph node's fibroblastic reticular cell (FRC) network: a second feature of T-cell motion that has been shown to be explainable by cell collisions rather than internal or chemotactic processes.

A later model by Graw and Regoes (2012) [31] also considered how the FRC network may influence immune cell dynamics with a 3D cellular automaton. FRCs were randomly positioned throughout the lattice and filaments of varying densities formed between them, and modelled naïve T-cells have a preference for following these fibres. The authors found that if the network is assumed not to affect cell velocities, then the fibre density had little

overall effect on the average time it takes for a T-cell to be activated, but that a disrupted or damaged network makes it *harder* for a T-cell to find DCs presenting its cognate antigen. Despite the small benefit reported, the authors note that the true network has a density that is maximally vulnerable to disruption. They suggested that there must thus be a different evolutionary advantage to the FRC network, for example, that it increases the rate at which DCs scan T-cells.

The models above aimed to address questions that had arisen from experimental work, to determine how lymph node biology and structure may or may not affect immune cell dynamics. Other studies have involved model-driven experimentation or attempts to predict the cellular output of the lymph node. Bogle and Dunbar (2010) [71] produced an on-lattice, agent based model of the lymph node in which simulated naïve T-cells contact DCs, become activated and proliferate. The model includes simulated inflammation of the LN and cell ingress and egress. The calculated rates of T-cell expansion and release were close to *in vivo* observations. A similar study [72] used a 3D cellular automaton and two-photon microscopy to attempt to predict the level of lymph node output in response to chronic and acute simulated infections, achieving close agreement. Another agent based model ([70], Figure 0.6c) with cells on a 3D lattice aimed to study how T-cell activation differs between synapses (strong, long-lasting interactions between T-cells and DCs) and kinapses (weaker and shorter interactions that only halt the T-cell briefly). By supposing that lower affinity interactions lead to kinapses and stronger ones lead to synapses, they found that observations were consistent with T-cells being able to integrate the last two days of interactions with DCs. They were also able to confirm the prediction that lower densities of dendritic cells lead to the elimination of kinapse-mediated T-cell activation. Other authors built a hybrid 2D off-lattice agent-based model [73], to study simulated concentrations of naïve $CD4^+$ and $CD8^+$ T-cells, dendritic cells, IL-2 and IFN-α. Cell movement is described by 2D diffusion, diffusion of extracellular cytokines is modelled by a set of reaction-diffusion PDEs, and cytokine-dependent intracellular signalling is modelled by ODEs. By modifying their diffusion constants, the spatial patterning of IFN-α and IL-2 can be modified and hence the spatio-temporal dynamics of T-cells and dendritic cells, leading to abnormally early terminal differentiation of T-cells. The authors argue that this may explain the exhaustion of T-cells observed during chronic viral infections.

The above models all investigated changing rates or levels of lymph node output. Another approach is to calculate the probability of any response occurring at all, as considered by Celli *et al* (2012) [19]. They used an off-lattice agent-based model of T-cells and two-photon microscopy to consider the minimum number of antigen-bearing dendritic cells and T-cells required for successful $CD4^+$ T-cell activation (Figure 0.6d). Simulated T-cells follow

a random walk with parameters set to match microscope observations. They used realistic mouse parameter values and quantified the probability of T-cell activation after 24 hours as a function of both dendritic and T-cell counts, and hence estimated the minimum required cell counts to drive activation. We have extended this model to quantify the 'trade-off' between the number of dendritic cells and the amount of antigen that they are carrying [74], which will be detailed in Chapter 1.

Of note is that all of the above models are agent-based, whether cellular automata or off-lattice models. There are multiple reasons for this. The first is that it is often the cell-cell interactions that are of interest, rather than population-level phenomena. The second is that many of these models were developed concurrently with microscopy studies, and an agent-based model is convenient for qualitative comparisons to these. Finally, the nature of immune interaction means that the cells of interest are rare within a population, and the assumption by ODEs that the population of these cells is large or well-mixed would thus be inappropriate. Though a PDE could model the spatial resolution, it is often desirable for each cell to have unique characteristics (such as level of activation) or for many different unique cells to be considered, meaning that a PDE solution could quickly involve a large number of equations, many of whose solutions would contain singularities (high concentrations of cells in a single voxel and zero elsewhere).

0.2.2 Pharmacokinetics and pharmacodynamics

Pharmacokinetics and Pharmacodynamics (PKPD), or Physiologically Based Pharmacokinetics (PBPK) is the modelling of human anatomy – usually the circulatory system – in order to predict drug concentrations across the body. It is important for the prediction of drug concentration profiles and effective dosages in Phase II/III trials, when extrapolating pre-clinical (animal) results, in paediatric trials, in patients with impaired organs, or in case of drug-drug interactions [75]. Chapters 3 and 4 concern the biology of T-cells trafficking through the bloodstream, and hence PBPK is a useful methodology to build on (we will find that it is useful for some, but not all, of our purposes). PBPK models have been used in drug development since 2000 and are readily accepted as supporting information by both the US Food and Drug Administration and the European Medicines Agency in regulatory submissions. These models generally take the form of multi-compartment ODEs, with compartments including generic 'blood vessels', vascular and interstitial compartments for each organ, and others as required. They can be parameterised for any species by considering the blood flow, blood volume and total volume of each organ, in addition to system parameters such as the cardiac output and total blood volume. The framework was first suggested in 1937 [76, 77] and was pioneered by Kischoff and Dedrick in the 1960s [78], but most of the

work has been in the last two decades, as computational resources have become sufficient for the calculations required. There have been systematic reviews of studies since 1980 [75] and particularly within the last few years [79]. An example of such an ODE system would take the form of mass action equations like the following mock system, which contains three processes:

1. Binding and unbinding of a ligand l to a receptor r to form a complex $[lr]$,
2. Movement of both ligand and receptor between the heart and organ(s) o,
3. Movement of both ligand and receptor between organ vasculature and interstitial space.

$$
\begin{aligned}
\frac{\mathrm{d}l_h}{\mathrm{d}t} &= -k_{\mathrm{on}} l_h r_h \quad +k_{\mathrm{off}}[lr]_h + \frac{B_o}{V_h}(l_o - l_h) \\
\frac{\mathrm{d}r_h}{\mathrm{d}t} &= -k_{\mathrm{on}} l_h r_h \quad +k_{\mathrm{off}}[lr]_h + \frac{B_o}{V_h}(r_o - r_h) \\
\frac{\mathrm{d}[lr]_h}{\mathrm{d}t} &= k_{\mathrm{on}} l_h r_h \quad -k_{\mathrm{off}}[lr]_h + \frac{B_o}{V_h}([lr]_o - [lr]_h) \\
\frac{\mathrm{d}l_o}{\mathrm{d}t} &= -k_{\mathrm{on}} l_o r_o \quad +k_{\mathrm{off}}[lr]_o - \frac{B_o}{V_o}(l_o - l_h) \quad -e_{l,o} l_o + \mu_{l,o}\tilde{l_o} \\
\frac{\mathrm{d}r_o}{\mathrm{d}t} &= -k_{\mathrm{on}} l_o r_o \quad +k_{\mathrm{off}}[lr]_o - \frac{B_o}{V_o}(r_o - r_h) \quad -e_{r,o} r_o + \mu_{r,o}\tilde{r}_o \\
\frac{\mathrm{d}[lr]_o}{\mathrm{d}t} &= k_{\mathrm{on}} l_o r_o \quad -k_{\mathrm{off}}[lr]_o - \frac{B_o}{V_o}([lr]_o - [lr]_h) \quad -e_{[lr],o}[lr]_o + \mu_{[lr],o}[\tilde{lr}]_o \\
\frac{\mathrm{d}\tilde{l_o}}{\mathrm{d}t} &= -k_{\mathrm{on}} \tilde{l_o}\tilde{r}_o \quad +k_{\mathrm{off}}[\tilde{lr}]_o \quad +e_{l,o} l_o - \mu_{l,o}\tilde{l_o} \\
\frac{\mathrm{d}\tilde{r}_o}{\mathrm{d}t} &= -k_{\mathrm{on}} \tilde{l_o}\tilde{r}_o \quad +k_{\mathrm{off}}[\tilde{lr}]_o \quad +e_{r,o} r_o - \mu_{r,o}\tilde{r}_o \\
\frac{\mathrm{d}[\tilde{lr}]_o}{\mathrm{d}t} &= k_{\mathrm{on}} \tilde{l_o}\tilde{r}_o \quad -k_{\mathrm{off}}[\tilde{lr}]_o \quad +e_{[lr],o}[lr]_o - \mu_{[lr],o}[\tilde{lr}]_o.
\end{aligned}
\tag{0.0}
$$

where l and r are ligand and receptor concentrations, $[lr]$ is a ligand-receptor complex, subscripts $\cdot_h$ and $\cdot_o$ indicate concentrations of each entity in the heart and some organ o's vasculature respectively, and $\tilde{\cdot}_o$ indicates the concentration of some entity in the interstitial space of organ o. B_o and V_o are the bloodflow and blood volume for organ o. k_{on} and k_{off} are on- and off-rates for the ligand from the receptor, $e_{l,o}$ is the extravasation rate of the ligand from the vasculature of organ o to interstitial space, $\mu_{l,o}$ is the return rate of the ligand from the interstitial space to the vasculature, and other e and μ terms are defined analogously. The system shown here is very simple, and others may have extra terms for processes such as receptor internalisation, multiple stages of binding or secretion, and may have extra compartments in the organ or blood vessels to consider, such as endosomal space. A recent example that presents their equations has been produced by Garg and Balthasar

[80]. Though ODE systems like these are ubiquitous when calculating dynamic behaviour of drug concentration, the resulting predictions are often also analysed with mixed-effect modelling of Cox proportional hazards and Kaplan-Meier survival estimates, in proprietary software packages such as NONMEM. To illustrate their usefulness, I will focus here on a few examples, following the perspective by Jones *et al* [77]. Direct citations for these examples were not given, likely due to the fact they relate to clinical compounds that were then in development.

The first example is the use of a PBPK model to test an alternate formulation of a drug, aiming to reduce its peak-to-trough ratio (that is, to make the concentration profile have shallower troughs and shorter peaks, without reducing overall drug exposure). The initial formulation was 'instant release' and was associated with neurological side effects. A human PBPK model indicated that a different capsule would restrict release in the stomach, and several formulations were simulated. The model predicted that the release could be slowed to as long as 10 hours with only an approximately 12% reduction in overall exposure to the drug $\left(\int_0^\infty \mathrm{d}t\ \text{concentration}(t)\right)$, with most absorption occurring in the large intestine. The predictions were tested in a non-human primate model and found to be within 'reasonable agreement'. The modified formulation was brought to clinic and results were also found to be in agreement with predictions.

The second example is the prediction of drug-drug interactions as a function of time and dosage. A candidate compound was known to have a higher unbound concentration in tissues than in plasma (blood), due to its ionic properties and pH differences. This fact affects predictions of drug clearance versus dose. It was further suspected that a particular 'cytochrome' (enzyme) named CYP degraded the drug in the liver, which was confirmed in experiments with human hepatocytes (liver cells). These factors together result in non-linear pharmacokinetics and potential drug-drug interactions. Further experiments were carried out that determined that 90% of the degradation of the drug was due to the enzyme CYP1A2. It is known that amine compounds such as fluvoxamine are taken up by this enzyme, too, and would inhibit clearance of the candidate compound. All of the experimental data on the candidate drug and fluvoxamine were incorporated into a PBPK model, which was able to match the accumulation of the compound with and without CYP1A2 inactivation and/or saturation. The simulations predicted the non-linear change in total drug exposure as a function of dose and dosing regimen, with or without inhibitor interactions. A clinical study of the compound given at various dosages and in tandem with fluvoxamine was designed, which confirmed the PBPK predictions. Modelling brought understanding of the pharmacokinetics as a function of both time and dosage, allowing the

compound to succeed where it would otherwise have been abandoned due to the unknown drug-drug interaction.

The third and final example is the extrapolation of adult dose-response to a paediatric (child) population. A model of the PKPD of a particular compound was built using preclinical data and adult human data after intra-vascular and oral administration. By identifying the enzymes responsible for drug clearance, the total rate of liver and renal clearances were estimated. A model was built to predict how each parameter changes as a function of age in a paediatric population. The resulting PBPK model was used to estimate concentration against time as a function of age, dosage, dosing regimen and delivery mechanism (size/type of tablet). The simulated results were a fair match to key features of the experimental data such as maximum concentration, half-life and steady state values, but the predicted overall absorption was less accurate. Nonetheless, the model provides predictive power for drug absorption and drug-drug interactions as a function of age. In particular, the model enables prediction of how much the absorption of a given drug may change in different age groups. The reason behind the discrepancy between the model and data is due to a large number of parameters for which paediatric values were unknown, such as the abundance of various enzymes and transporters. As understanding of these parameters increases, so can the predictive power of paediatric scaling in PKPD, which is invaluable in drug development, where clinical trials in children are difficult.

0.2.3 Models of T-cell trafficking

PBPK modelling is typically of the concentration profile of drugs or antibodies and these applications comprise the majority of published work. However, the methodology has also been applied to consider the distribution of cells or cellular therapies and these are the most relevant to the work presented in Chapters 3 and 4. Here, I will summarise some key examples that consider the trafficking of T-cells through the body with a PBPK model.

The earliest of such models was made by Zhu *et al* (1996) [81], who produced a system of 45 ODEs, describing 12 organs connected by blood and lymph flows. They parameterised the model using blood and lymph flow rates, vascular and interstitial volumes, and literature-estimated transmigration rates of cells in humans, mice and rats. They used the model to simulate and compare to data of naïve T-cells, TILs, non-activated NK cells and IL-2 activated NK cells in rats, humans and mice. The data came from a variety of literature sources, but usually contained multiple time points over a timescale between 24 hours and 9 days. The data generally only included a few organs and, despite the long time course, seemed to show that the localisation was at steady-state. The authors fitted the data using their ODE system by varying the attachment (to selectin) and arrest (due to integrin)

rates of cells in each organ with available data. They noted where T-cells spend most time (the lungs) in simulated results and hypothesised that this was due to high bloodflow and attachment rates. They thus suggested that T-cell delivery to tumours could be improved by introducing TILs in arteries supplying the tumour, to avoid initial entrapment of cells in the lung.

This work was later reutilised in 2002 [82] by the same research group to consider the recruitment of tumour-specific T-cells to tumour parenchyma. They introduced tumours to the legs of experimental mice and then injected either naïve or primed tumour-specific radiolabeled T-cells. They found that the distribution of both of these classes of T-cells were similar (as the resolution was likely not high enough to detect radiation in mouse lymph nodes, as observed in unpublished data from a collaborator), though primed T-cells were entrapped in the lung for longer. They found that liver and lung localisation is high, and even higher after a splenectomy is performed. They found that their previously published PBPK model produces a reasonable fit to the data, including after the splenectomy, but localisation in several organs (*e.g.* the spleen or heart) was over or underestimated in transient periods shortly after infusion. More importantly, they found that the model's predictions were consistent with the hypothesis that activated tumour-specific T-cells are no more likely to be recruited to the tumour than naïve ones. They suggested that TILs should be optimised for the adhesive molecules known to be present on tumours of specific types to remedy this.

Another extension in 2002 [83], once more by the same research group, built upon both this model and an additional model of antibody PBPK to consider the recruitment of 'bifunctional antibodies' (*i.e.* TCBs) to tumours in virtual mice. Model parameters such as anatomical data, rates of lymphocyte return from extravascular space, their depletion and generation in the body were all set as in 1996. The values of lymphocyte capture, release and arrest rates were assumed to be the same for all organs and based on data measured for naïve T-cells. No experimental data were collected in this study, and instead the localisation of T-cells and antibodies with and without the presence of each other was predicted. It was found that the biodistribution of T-cells is unaffected by antibodies (to within 1%), but T-cells affect the pharmacokinetics of the antibodies. The authors suggested that the only way for bifunctional antibodies to be effective in directing T-cells to tumours is to optimise them for receptors highly expressed in tumour vasculature to promote T-cell rolling and stopping, or to remodel the tumour vasculature to make this true, but they doubted that either would have a strong effect.

A more recent study by Poleszczuk *et al* (2016) [84] aimed to use a T-cell trafficking model to determine how the abscopal effect (the observation that radiotherapy against the

main tumour site can elicit a system-wide anti-tumour response) arises, and whether it is possible to target metastatic sites with radiotherapy to intentionally evoke the abscopal effect. They used a system of ODEs that includes the lungs, liver, gastrointestinal tract, breast and kidneys, with a tumour and metastatic sites in some/all of the lung, breast(s) and liver. The ODE system that they used to describe the number of effector cells in each organ i, E_i, was the following:

$$\frac{\mathrm{d}E_i}{\mathrm{d}t} = \lambda(E^* - E_i) - a(1-p)E_iC_i + \sum_{j=1}^{N} \omega_{ij}(\vec{C})\frac{fC_j}{g+C_j}E_j, \tag{0.1}$$

where E_i is the number of effector cells in organ i, λ is the effector cell decay rate, E^* is the steady state ("physiological") level of effector cells, a is the interaction rate of T-cells with cancer cells, p is the probability that a cancer cell is killed after such an interaction (instead of the T-cell), ω_{ij} is the probability of migration from organ i to organ j according to a Markov Chain model, and f and g determine the stimulation of immune cell proliferation by the presence of cancer cells. The Markov Chain used to determine the probabilities ω_{ij} depended on organ blood flows, volumes and by supposing that T-cells are three times more likely to extravasate to the organ that they were activated in than any other. They found that the distribution of T-cells across different lesions is very complicated. Addition of metastatic sites, for example, causes the main tumour site to grow larger due to T-cell redistribution. Removal of the main site can make metastatic sites grow larger, due to reduced T-cell proliferation. For their aim, they found that the success of generating an effective abscopal effect depends on which metastatic site is irradiated and where the main site is. They conclude that since the response to radiation is so unpredictable, metastatic sites for this purpose must be chosen carefully, and they advocate their model as a tool for choosing such targets. This model was extended in 2018 by some of the authors [85] to improve the precision of results, yielding similar conclusions to before.

Another model considered the PBPK of mesenchymal stem cells (MSC) with a similar methodology to the above models [86]. Their ODE system included the lungs, liver, spleen, kidney, heart and 'rest of body'. The ODEs were a simple mass balance argument, with arrest (effectively extravasation) and release (return) terms in the vasculature, and release and depletion (elimination) terms in the extravascular space. They matched model parameters to mouse data describing MSC abundance over 24 hours in various organs, and subsequently determined values of the cell arrest, release and depletion rates. They supported the model with a microscopy study, finding MSCs in the lung after 30 minutes as predicted, but not in the alveoli or associated with surrounding collagen. These cells then slowly deplete, as predicted. They evaluated the fit by using published data from mice, rats

and humans in normal or disease states, with timescales up to 24 hours or 10 days. The model performance was good, showing that the methodology can successfully describe the PBPK of MSCs in mice.

0.2.4 Age-structured models

In Chapter 4, I will describe a model in which the rate at which T-cells depart organ interstitial space depends on the length of time that they have been present, meaning that time and 'age' (time in the compartment) are both dependent variables. Such a PDE model is known as an age-structured model, or lifespan model. These have been studied extensively (examples of reviews: [87, 88]), but here I will consider just one example.

Chiorino *et al* (2001) [89] used a model to study how the cell cycle of cancer cells desynchronises. They performed experiments using ovarian carcinoma and leukaemia cells, initially in phase with each other, and measured the percentage of cells in S-phase as a function of time. This percentage falls from an initial 100% to 0% before rising again, but as the cells gradually desynchronise, the oscillations decay to an intermediate steady state distribution. The authors fitted a PDE system of the form,

$$\begin{aligned} &\frac{\partial p(a,t)}{\partial t}+\frac{\partial p(a,t)}{\partial a}=-\beta(a)p(a,t),\\ &p(0,t)=2\int_0^\infty \beta(a)p(a,t)\mathrm{d}a,\\ &p(a,0)=p_0(a), \end{aligned} \tag{0.2}$$

where $p(a,t)$ is the population of cells at time t and age a and $\beta(a)$ is the cell division rate as a function of age, $\beta(a)=\frac{f(a)}{1-\int_0^a \mathrm{d}a' f(a')}$, where $f(a)$ is a probability distribution for the division times. Defining a characteristic integral allows a solution to the equation to be found, which can be restated in a form that allows least-squares-fitting to experimental data. This yields values for the rate of convergence to steady state, the oscillation rate, and the mean and variance of the cell cycle duration for each cell type. Critically, these values can be found without assuming that the desynchronisation is due to a biological difference such as unequally sized daughter cells. They found that ovarian carcinoma has a shorter mean cell cycle time than does leukaemia, but it has a higher variability and hence loses synchrony faster.

0.3 Summary of DPhil

The studies reviewed in Section 0.2 exemplify how mechanistic models can elucidate hypotheses or carry out studies that would be difficult or infeasible experimentally, but whose predictions can be experimentally verified. They are ideal if the research question is 'why'

or 'how', whilst still maintaining some of the capacity of data-driven techniques to give quantitative results and address 'what', 'when', 'where' or 'how much'. Several also show how, if models are designed to address a biological hypothesis and their inputs and outputs are tied to (measurable) biological quantities, then they can be deeply integrated within an experimental or clinical framework. The examples above have focused mainly upon experimental science, but models have also been used to further clinical insight, for example to predict patient response rates in clinical trials by using historical data [62, 90, 91], to improve therapy schedules [61, 66–68] or to predict response in untested patient populations (e.g. paediatric patients) [92]. Though several of these models were mechanistic, many of them sought the answer to 'what': they predicted improved schedules or dosages, rather than investigating the reasons behind observed phenomena.

The objectives of this DPhil are to produce mechanistic models in order to propose potential reasons for observed variability or lack of clinical efficacy of immunotherapeutics ('why'), and, where feasible, to determine the limits of their efficacy ('how much'). We have modelled various steps of the cancer-immunity cycle, starting with a biological hypothesis at each step. We have focused in particular on models of T-cell activation in the lymph node and of T-cell trafficking, in order to investigate mechanisms of clinical failure for short peptide cancer vaccination and for CAR T-cell therapies. We aimed to investigate how the predictions of our models relate to clinical data, and so we ensured that the inputs and outputs of each model correspond to biological quantities. For example, parameter values are defined with values known from the literature rather than in a non-dimensionalised form, and outputs are at a 'systems' level that could be experimentally observed: We produced an *in silico* clinical trial to compare the response of virtual patients to the real ones, and we produced estimates of required dosages of CAR T-cells to compare to real clinical doses. This methodology also allows us to produce biological 'unit tests' (*e.g.* to check whether simple known biological phenomena are reproduced by the model) and to compare with real data, to build confidence in the models. Most importantly, it allows us to generate testable biological hypotheses and to fit our work within the context of biomedical research: it is not mathematical but biological novelty that we seek to produce.

In detail: the research chapters of the DPhil consist of two halves, each comprising two chapters. In the first half, we focus upon the first three steps of the cancer-immunity cycle and on short-peptide vaccinations for cancers. We hypothesise that, given the inconsistent success of such vaccinations, there may be quantifiable patient- or peptide-specific factors that limit the migration of dendritic cells or activation of T-cells (steps 2 and 3) and lead to variation in patient responses. To address this hypothesis, we developed an off-lattice, agent-based model of T-cell activation in the lymph node, detailed in Chapter 1. The

model quantifies how different patient-specific and peptide-specific factors influence T-cell activation after a short peptide vaccination. It was found that the off-rate of peptides from dendritic cell MHC-I and the time that dendritic cells take to migrate to the lymph node are the critical factors that determine T-cell activation. This model was then utilised to perform a virtual clinical trial on a virtual patient cohort, as detailed in Chapter 2. The aim of this chapter was to test whether the lymph node model is consistent with the observations of IMA901, a renal cell carcinoma vaccination clinical trial in which nine short peptides were introduced to the dermis of renal cell carcinoma patients. Given the predictions made from the lymph node model in Chapter 1, we hypothesised that the off-rates of the short peptides could be shown to have led to the lack of clinical efficacy observed, and that quantifying the model would allow us to suggest improvements to IMA901. The authors of IMA901 reported the number of patients responding to particular numbers of peptides; simulating the vaccination schedule using the lymph node model produced an expected number of peptide responses per patient consistent with these data. We present several potential reasons for the lack of efficacy seen in IMA901 and suggest how patient response could have been improved.

Chapters 3 and 4 focus upon steps 4 and 5 of the cycle, starting with a model of T-cell trafficking in Chapter 3. Using anatomical blood flows and volumes sourced from the literature, we produced a system of ODEs that can be used to find the maximum rates of delivery of T-cells across different organs and species (human, rat and mice), for both healthy and tumour tissue. We aimed to test the applicability of pre-clinical models to the clinic, to elucidate potential reasons for pre-clinical success failing to translate to the clinic, and why bioengineered/CAR T-cells have had greater clinical success for haematological than solid tumours. We hypothesised that the delivery rates of CAR T-cells to tissues and tumour would correlate with clinical observations of efficacy. We found that the delivery rates of an equal number of cells are up to 10,000-fold higher in mice than humans due to anatomical differences, but typical clinical dosages are only 100-fold less. We further found that clinical tumours are typically constrained by the perfusion of the tumour bearing organ, and that those organs with the highest perfusion (such as the kidneys) were those that some of the earliest successful immunotherapies targeted. Chapter 4 details how this model can be developed into a system of PDEs, introducing an organ-dependent residence time of T-cells and allowing temporal T-cell localisation to be quantified. We successfully fitted the model to rat lymphocyte data and found parameters consistent with biological expectations. The model was used to determine the extravasation rates of lymphocytes to different tissues. The model and associated data will be useful for identification of clinically relevant timescales for bioengineered T-cells, and determination of when either T-cell trafficking (step 4) or

extravasation (step 5) impede the clearance of solid tumours. The thesis will end with closing thoughts, discussion, and suggested future work.

Chapter 1

An *in silico* model of cytotoxic T-lymphocyte activation in the lymph node following short peptide vaccination

Note: much of the text in this chapter is reused from a publication of this work (written by myself; Liam V. Brown et al, J. Roy. Soc. Interface (2018) [93]).

1.0 Introduction

In Chapter 0, I discussed the cancer-immunity cycle and how it implies that any patient with an existing tumour most likely has a bottleneck somewhere in this cycle. In this Chapter and Chapter 2, I will discuss steps 1–3 of the cycle: the release of cancer antigen, presentation of those antigen to T-cells and T-cell activation. The most appropriate scenario to consider here are so-called cancer vaccinations, where cancer-associated antigen are introduced to the body in order to generate an immune response. Examples of these include mRNA vaccines, dendritic cell vaccines, viral vector-based vaccines, and 'short' peptide (meaning 9 amino acid long) vaccinations, in which the sequences associated with tumour antigen are introduced directly to the body with immune-activating agents. As previously discussed, such trials have seen limited success in the past. An example of a short peptide vaccine trial is IMA901 [52, 53], a clinical trial for renal cell carcinoma in which nine short tumour-associated peptides were administered to patients intradermally, aiming to raise multiple immune responses against the peptides and drive a tumour response. Surprisingly, most patients gave immune responses to only zero or one of the nine peptides, and none to more than three in any phase. It would be difficult to design an experimental study to address why these observations occurred, but an *in silico* model can generate explanatory

hypotheses by codifying all of the steps between vaccination and the activation of T-cells within the draining lymph node. The vaccine trialed in IMA901 contains both peptides and an adjuvant that drives activation and maturation of local dendritic cells (DCs). The peptides bind to MHC-I receptors on the DCs and the DCs migrate to lymph nodes, a process that takes between 20 and 30 hours [12, 13, 15, 94–98]. In the lymph nodes, DCs present peptides bound to MHC-I receptors on their cell surfaces to $CD8^+$ T-cells. T-cells that have an affinity for the peptide–MHC (pMHC) complex – T-cells for which this antigen is their *cognate antigen* – can become activated and begin to proliferate.

The objectives of this chapter are to develop a model that determines how the probability of successful T-cell response is influenced by vaccine properties, including the number of cognate antigen-specific T-cells and DCs, their dynamics, and biophysical peptide properties. We aim to generate a hypothesis linking one or more patient or peptide-specific parameters to vaccination success, that can explain the results seen in IMA901. We hence require a model both of the vaccination site and of cellular dynamics in the lymph node. In Chapter 0, I discussed several previous models of T-cell activation in the lymph node [19, 20, 22, 70–72]. One model in particular [19] predicts the minimum number of T- and dendritic cells required for successful T-cell activation. However, the focus of this model was limited to cell dynamics; to model a short-peptide-based vaccine, it is also necessary to consider the amount of administered peptide bound to DC MHC-I receptors and its dissociation rate. Hence, starting from this model, we have produced an integrated model of peptide competition in the dermis, DC migration and DC–T-cell interactions within the lymph node. In this chapter, I will introduce the model and the assumptions made in its creation. I will show results that justify many of these assumptions, and how their implications impact upon predicted probability of T-cell activation. I will use these results to suggest which properties of a short peptide vaccination most influence vaccine success. The model will then be utilised in Chapter 2 to investigate IMA901 in detail.

1.1 Methods

1.1.0 Model overview

Our model is summarised in Figure 1.0, showing all events from the introduction of peptide to the dermis, to T-cell activation in the draining lymph node. Peptide competition and clearance is modelled in the vaccination site and lymphatic vessels with ODEs, and T-cell–dendritic cell interactions in the lymph node are modelled with an agent-based model. In contrast to modelling frameworks which consider population level descriptions, for example by considering mean-field equations, agent based models track all individuals, for instance

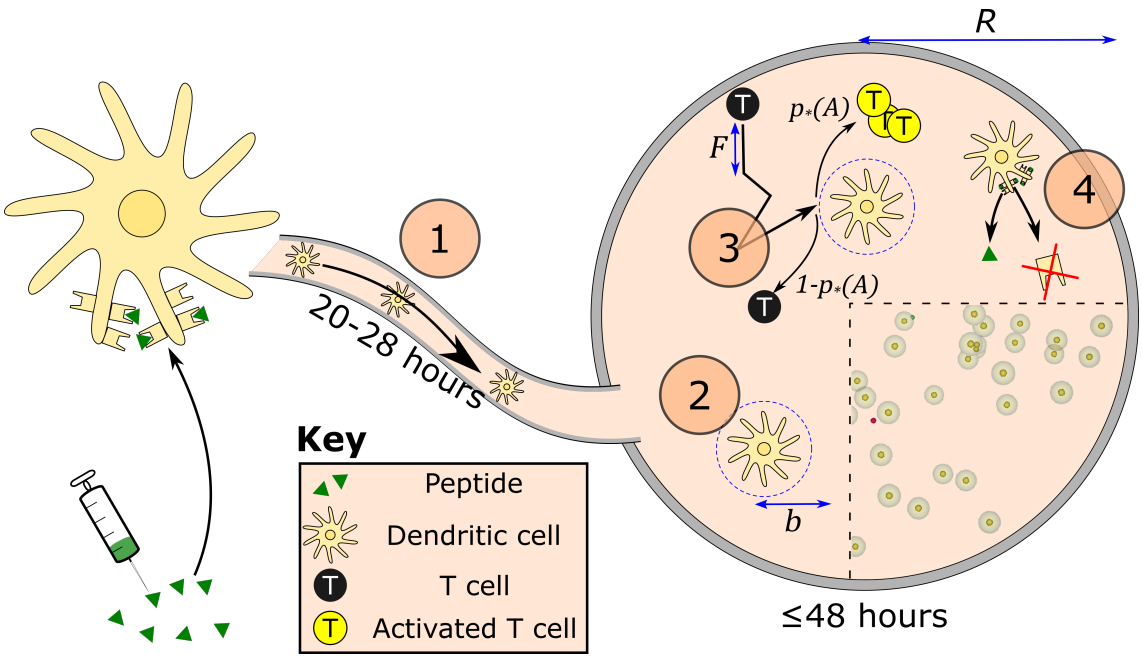


Figure 1.0. Model overview. **1)** Following short-peptide vaccination in the dermis, peptide binds to dendritic cell (DC) receptors (MHC-I). The DCs mature and then migrate to the lymph node via the lymphatics, which in total takes 20–30 hours following vaccination. They are then placed at random locations in the simulated lymph node. **2)** DCs in the paracortex may contact T-cells within a radius b. **3)** T-cells move with velocity v and mean free path F. If they contact a DC, they are activated with probability $p_*(A)$, where A is the proportion of MHC-I receptors on the DC with cognate antigen bound. **4)** A reduces over time due to unbinding and to MHC-I receptor turnover by the cell, depicted by the arrows indicating the removal of peptide and receptors.
Cutout: A modified screenshot of the graphical output of the model, which utilises GLSC3D [99]. Yellow cells are DCs, and the surrounding spheres are their contact radii. Red cells are T-cells, which turn green when activated.

cells, whose behaviour in space and time is governed by a set of rules. In the current context one agent might be a T-cell, which moves every time-step according to its stored velocity. If this cell approaches within a certain distance of a dendritic cell (DC) agent, then it will become activated or not based on parameters stored in the DC agent. Agent-based modelling is ideal when the individuals under consideration are not identical, undergo complex dynamics and are not present in extensive numbers, rendering standard reductions to population-level equations inappropriate. For instance one common technique, the Van Kampen expansion of the master equation [100], is impractical due to a combinatorial explosion in the dimension of the state space required to classify a system with numerous non-identical agents, and the absence of a thermodynamic limit. The use of an ABM allows comparison of results to work by other authors (see Section 0.2.1) and cellular dynamics to two-photon microscopy studies such as [26].

In particular, the framework we have developed is that of a human parameterised, agent based model of simulated T-cells interacting with dendritic cells (DCs) that carry antigen

within an off-lattice, spherical lymph node (LN). The LN is assumed to be non-inflamed and thus focus is on the probability of a first successful encounter between naïve T-cells and DCs. T-cells in the model are present in the LN at the beginning of the simulation, but DCs carrying varying proportions of peptide antigen are assumed to arrive gradually from the vaccination site at physiological rates. DC antigen decays exponentially over time according to its off-rate with MHC-I (receptors on the DC surface) and an estimated physiological MHC-I receptor turnover rate. When a T-cell interacts with a DC, it is assumed to be activated with a probability that depends on the amount of cognate antigen (*i.e.* antigen that the T-cell recognises) presented by the DC. The output of the model is the overall probability of CTL (cytotoxic T-lymphocyte) activation as a function of time up to 48 hours, the assumed lifetime of activated DCs within the LN [12–17]. It may be possible to calculate such probabilities with a simpler model such as ODEs or PDEs, by encoding the probability of T-cell–dendritic cell interaction with respect to lymph node geometry and considering antigen decay as a time limit in the model. Waves of arriving dendritic cells could be modelled by using an age-structured PDE. It may also be possible to derive an analytical expression for the probability of activation, by considering the first passage time until collision [101–103]. We postpone such investigations for further study.

We have systematically investigated the parameter regimes for which antigenic properties and/or cell interaction dynamics within the LN are important for the probability of CTL activation, and any "trade-off" between them. Initial positioning of cells, direction and magnitude of cell velocities, and the probability of successful T-cell activation all depend on random numbers, as detailed in the following sections. The model is hence non-deterministic, and all presented results are an ensemble average over a number of replicates specified in each case, which are sufficient to ensure the ensemble predictions used in this study are insensitive to including further replicates. The following sections give a more detailed outline of each stage of the model, from vaccination site to lymph node dynamics.

1.1.1 Vaccination site

Prior work [104] suggests that the rate-limiting processes for CTL activation following vaccination are diffusion, vascular clearance, binding of peptides to MHC-I receptors and the number of peptide-presenting DCs migrating to the draining LNs. Thus, we assumed that other vaccine components do not need to be explicitly modelled. We expressed peptide binding in the dermis as a set of ODEs, originally created by Mailen Kootsey, Jonathan

Wagg *et al* [104] and subsequently modified by us:

$$\begin{aligned}\frac{\mathrm{d}l}{\mathrm{dt}} &= k_{\mathrm{off}}c - k_{\mathrm{on}}rl - k_{\mathrm{clear}}l,\\ \frac{\mathrm{d}r}{\mathrm{dt}} &= k_{\mathrm{off}}c - k_{\mathrm{on}}rl,\\ \frac{\mathrm{d}c}{\mathrm{dt}} &= k_{\mathrm{on}}rl - k_{\mathrm{off}}c,\end{aligned} \tag{1.0}$$

where l, r and c are the concentrations of free peptide (**l**igand), free MHC-I **r**eceptors and bound peptide-MHC-I **c**omplexes respectively, and $r + c = r_{\mathrm{tot}}$ is constant. Clearance of free peptide is rapid (on a timescale of minutes [105–107]) and mediated primarily by the vasculature [108]. Additionally, DCs only begin to migrate after several hours, hence we assume that there is no free peptide present in the lymphatic vessels and that rebinding can be ignored, in which case these equations are effectively $\frac{\mathrm{d}c}{\mathrm{dt}} = -k_{\mathrm{off}}c$ in the lymphatic vessels. See Section 1.2.0 for consideration of the initial conditions of this ODE in the lymphatics. Note that to model nine peptides at once, l, c and k_{off} can be replaced by peptide-specific variables such as l_i and c_i, with $0 \leq i < 9$. In this case, the equations would be coupled because the total number of free and bound receptors is a constant, $r + \sum_{i=0}^{8} c_i = r_{\mathrm{tot}}$.

1.1.2 Agent-based lymph node model

We model the lymph node paracortex as a sphere with dendritic cells (DCs) and T-cells at random locations throughout the volume, such that no cell is initially within a contact distance b from another. DCs are added to their assigned locations over a finite period of time, arriving at either a constant rate or at times drawn from a Poisson distribution. The DCs present a variety of pMHC-I complexes, but only a proportion A of them are the antigen that are cognate to the modelled T-cells. Assuming that antigen are continually replaced with self-antigen, this proportion decays with time t according to the pMHC-I off-rate k_{off} such that,

$$\dot{A}(t) = -k_{\mathrm{off}}A(t). \tag{1.1}$$

The amount of antigen upon DC arrival within the lymph node is defined by $A_{\mathrm{dermis}}e^{-k_{\mathrm{off}}t_{\mathrm{arrival}}}$, where A_{dermis} is the cognate antigen ratio bound to the DC upon migration out of the dermis and t_{arrival} is the time taken for the DC to reach the draining LN. This initial ratio subsequently decays further according to Equation 1.1. DCs are generally assumed to be stationary in the LN, though this assumption is relaxed in some cases to confirm its validity. T-cells draw random velocities and free path lengths from gamma and Gaussian distributions, to match experimental data [26] as detailed in Section 1.1.4. The free path represents the distance that a T-cell moves before it stops for any reason and reorients, after which it draws a new free path value. If a T-cell moves within a contact distance b from a DC,

then its chance of activation is determined as detailed in Section 1.1.3 below. If activated, the T-cell is removed from the simulation, otherwise it stops and reorients in a direction away from the DC as outlined above. T-cells do not collide with any other simulated cells, as this effect is already included in the free path distribution. However, T-cells that reach the LN boundary stop and move in a new random direction. They continue to search the volume until one of three conditions becomes true:

- All T-cells have been activated, or
- No DCs have enough cognate antigen to activate a T-cell, or
- 48 hours have passed (approximate lifetime of activated dendritic cells in the LN [12–17]).

1.1.3 Probability of T-cell activation

For DC–T-cell contact to result in successful CTL activation, a threshold number of T-cell receptors (TCRs) must be ligated by antigen with sufficient affinity (*i.e.* their cognate antigen) to drive TCR clustering and subsequent CTL activation [109, 110]. DCs and T-cells have multiple interactions over several hours [12, 23, 26, 111] and "integration" of multiple such interactions leads to the final activation state of T-cells [12, 23, 26, 112], but we assume that if the first of such interactions fails, then all of them will fail. We assume that cognate peptide-MHC-I complexes, self peptide-MHC-I complexes and T-cell receptors (TCRs) are all expressed uniformly on the surface of DCs and T-cells, and that any new MHC-I receptors up-regulated during DC maturation are independent of the initial population of MHC-I receptors that could be bound to the short peptide. Suppose that the proportion of MHC-I receptors that are bound to cognate antigen is A and that N MHC-I receptors are sampled from this population, representing the sampling of the DC's receptors by a nearby T-cell. Then the probability that x from this sample are cognate is (approximately) given by the binomial distribution,

$$\frac{N!}{x!(N-x)!}A^x(1-A)^{N-x}.$$

Following this, the probability that the number of cognate receptors x within the sampled population N surpasses some threshold $n \leq N$ is given by,

$$p_*(A) = 1 - \sum_{x=0}^{n-1} \frac{N!}{x!(N-x)!}A^x(1-A)^{N-x}. \tag{1.2}$$

Supposing that the threshold number of TCRs that must be ligated for activation is n, $p_*(A)$ represents the conditional probability of activation given that a cognate T-cell has made

contact with a DC, with N the total number of MHC-I receptors in the T-cell-DC contact area that were expressed on the DC surface at the time of vaccination. This conditional probability of activation is plotted against the cognate antigen ratio A in Figure 1.1 with other parameters taking the values in Table 1.0. Since the number of MHC-I receptors in the T-cell–DC contact region N is large, there is a sharp transition in $p_*(A)$, which approaches unity rapidly for $AN > n$. This is because for large N, comparatively small changes in A lead to large changes in AN. For example, if $n = 20$ and $N = 500$, $p_*(AN{=}32) = 0.9923$.

Adopting this equation for T-cell activation and ignoring the multiple interactions that T-

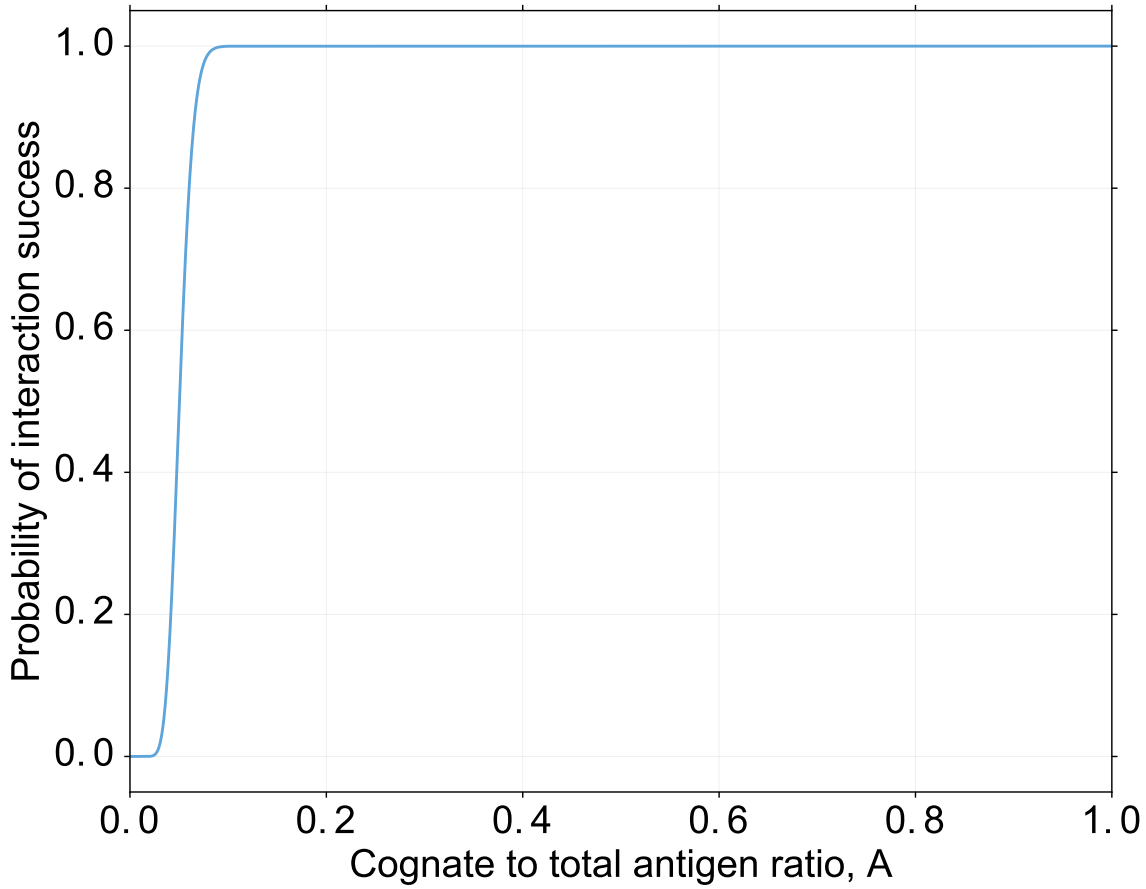


Figure 1.1. Probability of success for a single T-cell–dendritic cell (DC) interaction in terms of the proportion of receptors A on the DC's surface that are bound to cognate antigen, as given by Eq 1.2, with other parameters as listed in Table 1.0, including a threshold of $n = 20$. There exists a transition value, A_*, significantly below which there is almost zero probability of interaction success, and significantly beyond which success is essentially certain. The reason for this transition follows from properties of the binomial distribution, as explained in Section 1.1.3.

cells have with DCs avoids one additional complexity: when T-cells begin to proliferate, the lymph node becomes inflamed, affecting both cell densities and dynamics. We also ignore varying TCR affinities for different antigen: we assume instead that there is a precursor frequency ϕ of T-cells capable of recognising the antigen with approximately equal affinity.

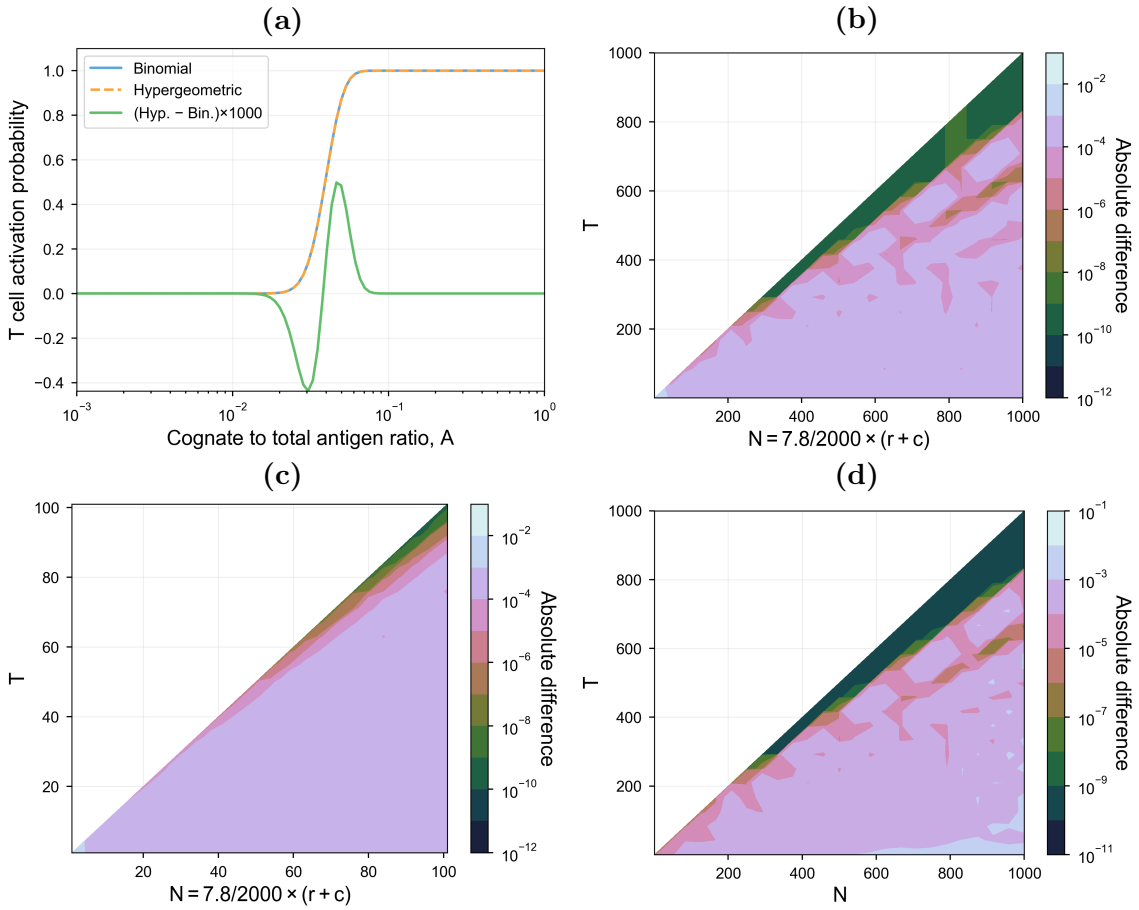


Figure 1.2. A comparison of the estimated probability of T-cell activation when it depends on the binomial distribution (Equation 1.2) or the hypergeometric distribution. **a)** The values of both distributions and the difference between them, as a function of the cognate-to-total antigen ratio A for typical parameter values ($n = 20$ and $N = 500$). **b)** The largest absolute difference between the two distributions for any A, as a function of the number of antigen in the T-cell–dendritic cell (DC) contact region N and the T-cell activation threshold n. The size of the contact region relative to the DC's surface area is held fixed, such that the total number of antigen on the DC $r + c = 2000/7.8N$. **c)** The same plot with a shortened x-axis. **d)** The same plot as in (b), but with a constant $r + c = 10^5$.

Finally, the role of $CD4^+$ T-cells in licensing DCs for activation [23, 113] is not considered and our model quantifies only the probability that CTL activation fails due to insufficient quantities of migrating DCs or cognate antigen. All of these assumptions are in accordance with our objective to find the regimes in which CTL activation is impossible, rather than the precise prediction of LN output.

One final assumption of mathematical importance is our use of the binomial distribution, which represents sampling of pMHC-I by the T-cell with replacement. A more exact distribution is the hypergeometric distribution, which samples *without* replacement. Suppose that there are c MHC-I on a DC bound to cognate antigen and r that are not, so that the proportion of bound antigen for the cell is given by $A = \frac{c}{r+c}$. Further, suppose that N MHC-I are sampled from the population of the DC without replacement, representing the

sampling of the DC's receptors by a nearby T-cell. Then the probability that x from this sample are cognate is given by the hypergeometric distribution,

$$\frac{\binom{c}{x}\binom{r}{N-x}}{\binom{r+c}{N}}.$$

In the limit $r, c \to \infty$, this reduces to the binomial distribution given above. We assume that $r + c = 10^5$ [10, 114–117] and so for all values $\frac{n}{N} < A < 1$, the fractional difference between the hypergeometric and binomial distributions is less than one part in a thousand for typical parameter values ($n = 20$ and $N = 500$; Figure 1.2). With $N = 7.8/2000 \times \eta$, corresponding to the assumed ratio of the T-cell–DC contact region area to the DC's surface area, this holds for all values of n and N between 1 and 1000. Without this restriction on contact areas, the difference rises as N approaches η. However, we never vary our estimate of the contact region or DC surface area, and always work with values of $n < 50$ and $300 < N < 700$. We hence proceed with the simpler and more intuitive binomial distribution to calculate probabilities of T-cell activation.

1.1.4 T-cell dynamics

We used previously published (mouse) data [26] to fit velocity distributions to T-cells and DCs, as shown in Figure 1.3. We expect the velocities of T-cells and DCs to be similar between mice and humans and so use this data to parameterise our human model. For each distribution and each data set, the distribution parameters were selected from the best least-squares-fit predicted by SciPy [118]. The gamma distribution provides the best fit to both T-cells and DCs, so we chose this distribution for both cell types in our model.
No theoretical distribution that we considered provides a good fit to velocities of T-cells in the absence of DCs. The authors [26] suggest that the T-cell velocity distribution changes in the presence of DCs due either to factors released by DCs or to cell-cell contacts, though it occurs regardless of whether the DCs carry cognate antigen. Since they note that the change in T-cell motility occurs only after first contact and because the first contact is the limit of our model, the correct distribution to use is in the absence of DCs. For the reasons above, we chose to fit the gamma distribution to this distribution. This ensures that the shape of the velocity distribution is similar to other data and that the observed mean velocity is correctly reproduced.

1.1.5 Dendritic cell dynamics

Following the estimates and assumptions of Teichgräber *et al* [104], we assume that the number of DCs per square millimetre of dermis is 600 [119] and that the initial injected vaccine solution dispersion area is 0.6cm^2. The number of these DCs that successfully reach

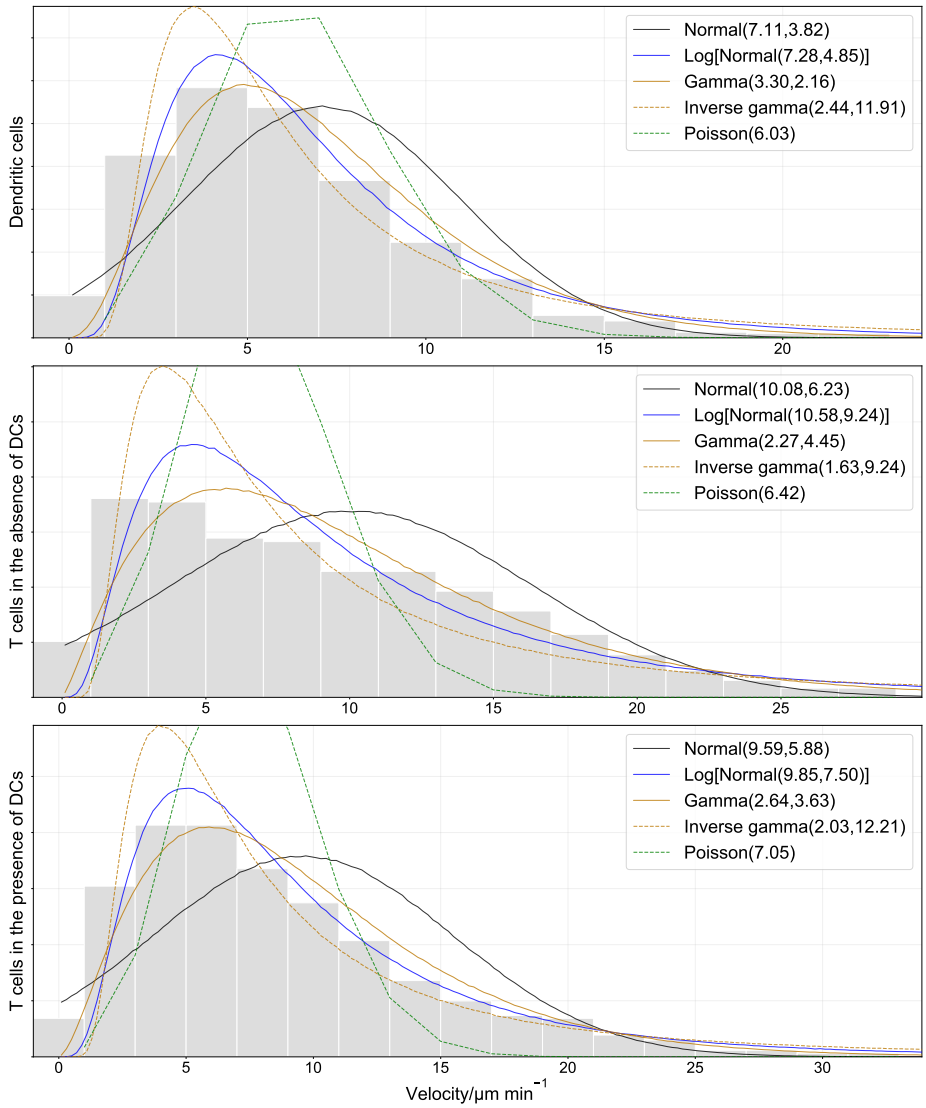


Figure 1.3. Histogram of velocity measurements of T-cells and dendritic cells in mice as in figures 1–2 of Mempel *et al*'s letter [26], with various velocity distributions fit to them. The parameters used for each distribution are indicated in the figure legends.

the draining LN is small [120]; we assumed it to be of order 1% [120] and that they reach the LN at different times. We considered cases where DCs arrive at a constant rate between the times P and $P + p$ or are Poisson distributed (see Section 1.2.5).

Another potentially important consideration is the movement of DCs within the lymph node. When DCs arrive from an afferent lymph vessel, they are deposited into the interstitial space surrounding the LN. They move randomly before entering the node at essentially any point, with points closest to the afferent lymph vessel more likely. After entry, DCs exhibit directed random motion towards the high endothelial venule at the LN centre, and then to the paracortex. Once they have settled into the paracortex, they 'run and tumble' akin to T-cells with lower motility. We have considered the entry of DCs into LNs and their

random motion separately, to quantify the effect of including or ignoring either.

DC arrival motion: The entry of each DC is modelled by choosing random positions on the outer surface of one hemisphere of the virtual LN. The DC then moves at constant velocity towards a random location within the LN and halts. Alternatively, DCs are placed directly at a random location such that it is not in contact with another DC.

DC random motion: Movement of DCs once within the paracortex is modelled by assigning velocity and mean free path distributions to DCs, as described for T-cells in Section 1.1.4. Most sources report DC velocities of around 2–6µm/min (*e.g.* [12, 13, 26]), so we consider this range for both active migration and random motion.

1.1.6 Parameter selection

Model parameter values were selected and/or estimated from relevant literature, as summarised in Table 1.0. We matched values to measurements in primary citations where possible, otherwise we have implemented a value within the reported range in the experimental literature. We have used mouse parameters where human data was absent and the parameter is not expected to change substantially between mouse and human (for example, cell velocities). T-cell numbers were chosen by multiplying observed T-cell densities with a precursor frequency ϕ, the proportion of T-cells with an affinity for the modelled cognate antigen, and DC numbers were chosen as described in Section 1.1.5.

1.1.7 Sensitivity analyses

To identify important parameters and ensure that the model is well-behaved, we performed a sensitivity analysis. We utilised three techniques for this: the random forest model [134], calculation of the variance of model outputs under variation of parameters and a simple local analysis. Random forest was implemented using the python package Scikit-learn [135] and the variation of parameters made use of extended FAST [136–138]. Sensitivity analyses for random forest and eFAST were performed over a Latin Hypercube with sides up to $\pm 25\%$ of the default values given in Table 1.0, with 13 parameters (listed in Table 1.2), 84000 samples (a number calculated to suffice for eFAST) and 10 repeats per parameter set. The local sensitivity analysis was performed by calculating the probability of T-cell activation for a single parameter set with 1000 repeats, then altering each parameter by 5% and calculating the change in the probability of activation. R^3 and b^3 were altered in lieu of R and b, as cell counts and interaction volume are proportional to the cube of these respective parameters.

The random forest technique can yield relative importance factors for each parameter, whose sum adds up to one. eFAST yields two indices for each parameter. The first of these

Sym	Parameter	Typical value	Reference
A	Ratio of cognate to total antigen on DC surface	Initially 10%–20%	**HS:** [121]
b	Contact radius	20µm	**MM:** [122–126] **HS:** [22] **CM:** [72]
D	Number of simulated DCs	500–1000	See Section 1.1.5
F	T-cell free path	25µm	**MM:** Calculated from [18, 26]
F_{DC}	DC mean free path	3–5µm	**MM:** Calculated from [26]
k_{off}	Antigen off-rate from MHC-I receptors	N/A	
N	Antigen in T-cell–DC contact area	500	**MM:** [13, 17, 18, 95]
n	T-cell activation threshold	20	**MM:** [109, 110, 127–130]
P	Time of first DC arrival	18 hours	**MM:** [15, 94–96]
p	Time between first and last DC arrival	6 hours	**MM:** [12, 13, 15, 97] **HS:** [98]
ϕ	T-cell precursor frequency	10^{-4}–10^{-6}	[19, 131–133]
-	Density of T-cells in paracortex	$\mathcal{O}(10^5$–$10^6)/\mathrm{mm}^3$	**MM:** [12, 15, 17, 20] **CM:** [72]
T	Number of T-cells	Derived from above	
R	LN (paracortex) radius	500µm	[19, 122–126] **HS:** [22] **CM:** [72]
r_{tot}	Number of MHC-I receptors per DC	100k	Internal+**MM:** [10, 114–117]
v	T-cell speed	10µm/min	**MM:** [18, 25]; [26] in particular
w	DC speed	2–6 µm/min	**MM:** [12, 13, 26]

Table 1.0. Model parameters. References for each parameter are organised by species: **MM**, mouse; **HS**, human; **CM**, monkey.

is the "sensitivity index", which represents the proportion of an output's variance that is due *solely* to a given parameter, *i.e.* first order contributions. The sum of these indices is usually less than one, due to missing higher order terms, which involve combinations of parameters. The "total sensitivity index", is calculated from the proportion of an output's variance due to all parameters except one. This index contains all of the contributions that involve a given parameter at any order. The higher order terms are counted in multiple indices, so the sum of the indices is usually greater than one.

The difference between the total sensitivity and sensitivity indices is the contribution of a parameter's variance due only to higher order terms, indicating the degree of "interaction" that a parameter has with others. Total sensitivity indices can be large even for unimportant parameters, if its higher order terms have contributions from other, more important parameters. The local sensitivity analysis can be compared to these values by dividing the changes in activation probabilities by their sum.

1.2 Results

1.2.0 Peptide dynamics

1.2.0.0 Perturbation analysis

The ODE system in equation 1.0 can be separated into three asymptotic regimes (where a 'regime' refers to a timescale over which a particular equation well approximates the solution to the ODE system), because 1) there are three rate parameters (k_{off}, k_{on} and k_{clear}) with different orders of magnitude and 2) the number of ligands is much greater than the number of receptors in the system. The three regimes represent the binding of peptide to receptors, the clearance of peptides with receptors at maximum occupancy, and the gradual decay of bound receptors when peptide has been fully cleared. They are associated with the timescales $\frac{r_{\text{tot}}}{k_{\text{clear}} l(0)}$, ${k_{\text{clear}}}^{-1}$ and ${k_{\text{koff}}}^{-1}$, respectively. Approximate solutions in each regime can be found by taking values for relevant rate parameters as follows: $k_{\text{off}} = 0.1\text{h}^{-1}$ (an example off-rate of order of one of the stable peptides in IMA901 from MHC-I), $k_{\text{on}} = 100\text{h}^{-1}(\mu\text{M})^{-1}$ (the assumed on-rate for all IMA901 peptides to MHC-I) and $k_{\text{clear}} = 10\text{h}^{-1}$ (the clearance rate of peptides from the skin by the vasculature). The concentration of peptides at the vaccination site is assumed to take the constant value of 200µM. The concentration of dendritic cells is assumed to be 2×10^{-7}µM, each with $\eta = r + c = 10^5$ receptors, for a receptor density of 2×10^{-2}µM.

The middle (middle-time) and inner (short-time) solutions can be found by rescaling the number of ligands to $l' = l/l(0)$ and the number of bound complexes to $c' = c/r_{\text{tot}}$, and choosing the timescale $T = \frac{k_{\text{clear}} l(0)}{r_{\text{tot}}} t$. The middle solution can be found by choosing the timescale $\tau = k_{\text{clear}} t$. These yield the equations,

$$\begin{aligned} \textbf{Inner:}\ & l(T) = l(0), && c(T) = r_{\text{tot}} \left(1 - e^{-(N+F)T}\right), \\ \textbf{Middle:}\ & l(\tau) = F\, l(0) + (\text{const}) e^{-\tau}, && c(\tau) = r_{\text{tot}}, \end{aligned}$$

where $F = \frac{k_{\text{off}} r_{\text{tot}}}{k_{\text{clear}} l(0)}$ and $N = \frac{k_{\text{on}} r_{\text{tot}}}{k_{\text{clear}}}$. At long times in the middle solution, $l(\infty)$ becomes equal to $Fl(0)$. Rescaling the number of ligands to $l'' = l/\left(Fl(0)\right)$ and the number of complexes to $c' = c/r_{\text{tot}}$, and choosing timescale $\theta = k_{\text{off}} t$ gives the outer (long-time) solution,

$$\textbf{Outer:}\ l(\theta) = F\, l(0) e^{-\theta}, \qquad c(\theta) = r_{\text{tot}} e^{-\theta}.$$

Matching each of the three solutions at their boundaries allows the constants to be fixed,

$$\begin{aligned} \textbf{Inner:}\ & l(t) = l(0), && c(t) = r_{\text{tot}} \left(1 - e^{-(k_{\text{on}} l(0) + k_{\text{off}})t}\right), \\ \textbf{Middle:}\ & l(t) = Fl(0) + l(0)\,(1 - F)\, e^{-k_{\text{clear}} t}, && c(t) = r_{\text{tot}}, \\ \textbf{Outer:}\ & l(t) = Fl(0) e^{-k_{\text{off}} t} && c(t) = r_{\text{tot}} e^{-k_{\text{off}} t}, \end{aligned}$$

which can then be combined by adding them together and removing the values of the solution where they overlap, which yields the composite,

$$\begin{aligned} l(t) &= l(0)e^{-k_{\text{clear}}t} + \frac{k_{\text{off}}r_{\text{tot}}}{k_{\text{clear}}}\left(e^{-k_{\text{off}}t} - e^{-k_{\text{clear}}t}\right) \\ c(t) &= r_{\text{tot}}e^{-k_{\text{off}}t}\left(1 - e^{-k_{\text{on}}l(0)t}\right). \end{aligned} \tag{1.3}$$

These two equations are plotted in Figure 1.4, together with the numerical solution of the ODE system with one peptide. The solution is indistinguishable by eye. The initial binding of peptides to MHC-I happens over a timescale less than a second, the clearance of peptides from the vasculature over minutes and the unbinding of peptides from pMHC-I (for this peptide off-rate) over hours. We assume that peptide is cleared overwhelmingly by the vasculature rather than the lymphatics (see discussion) and so is not encountered by the DCs as they move through the lymphatic vessels. In other words, k_{clear} is large and removes peptide from the system, rather than sending peptides to the same compartment as migrating DCs. Hence, the ODE system in equation 1.0 can be replaced by the single equation $\frac{\mathrm{d}c}{\mathrm{d}t} = -k_{\text{off}}c$ in the lymphatics. Alternatively, dendritic cells reach the lymphatics at times $t \gg \frac{1}{k_{\text{on}}l(0)}$, which reduces the composite equation for c to $c(t_{\text{lymph entry}}) \approx r_{\text{tot}}e^{-k_{\text{off}}t}$. The initial condition for this reduced ODE can be determined from r_{tot} and the length of time that DCs spend 'idle' in the dermis before leaving, which is typically around 16 hours: $c(t_{\text{lymph entry}}) \approx r_{\text{tot}}e^{-k_{\text{off}}16\text{h}}$. This analysis could be altered to account for multiple peptides with different off-rates. The equations would be coupled as before (through the condition on the total number of receptors, $r + \sum_{i=0}^{8} c_i = r_{\text{tot}}$) but the timescales and initial conditions involved would yield a similar result, except for peptides where k_{off} is not much less than k_{clear}, for which $l(t)$ and $c(t)$ would reach zero values more quickly. In any case, dendritic cells in the lymphatic vessels either have zero or an exponentially decaying initial amount of each antigen, where we assume that rebinding of peptides can be ignored (see discussion). Under these assumptions, the exponential decay equations for each peptide become independent.

1.2.0.1 Time course

Modelling an expanded version of the ODE system in Equation 1.0 with nine peptides and biologically relevant parameters produces the results in Figure 1.5. As before, free peptide rapidly drops to zero (not explicitly shown), resulting in a distribution of bound peptides that depends only on time and peptide off-rate. The results in this chapter do not depend on peptide competition; they consider only a single peptide at a time. For these results, we may drop this ODE system and work instead with a defined peptide off-rate and the proportion of pMHC-I bound to the peptide of interest, $A(t) = \frac{c}{c+r}(t)$.

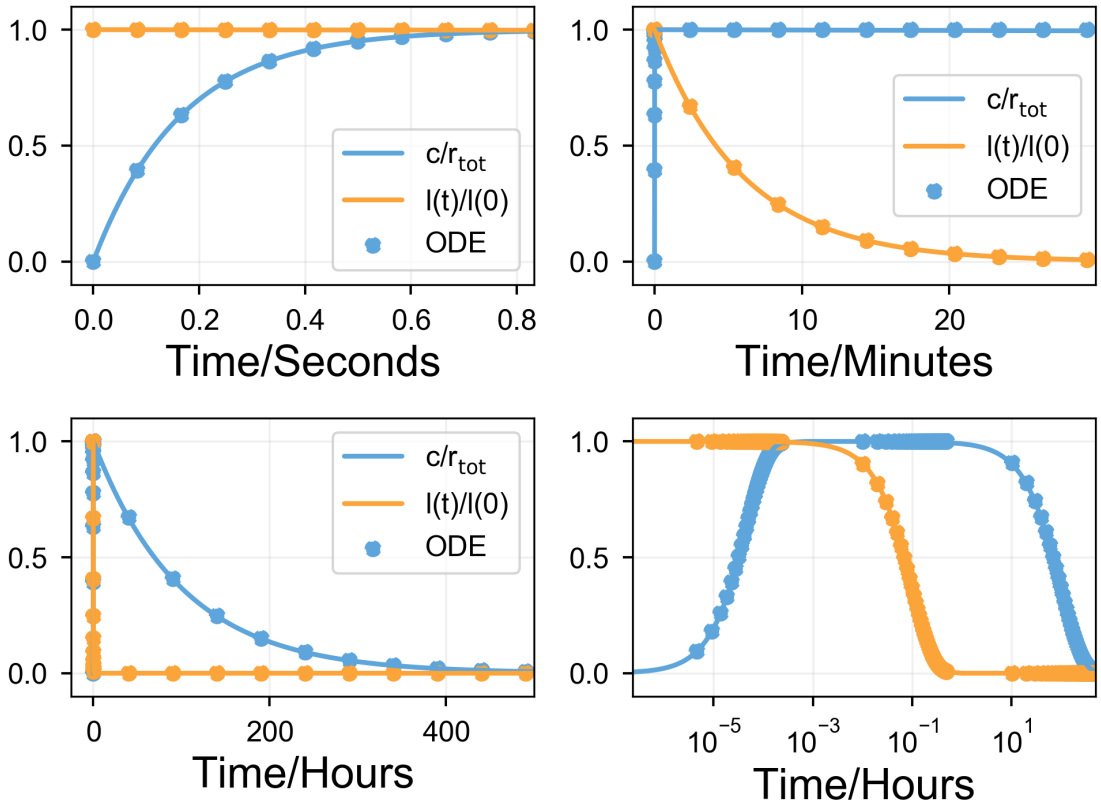


Figure 1.4. The composite equation, Equation 1.3, derived from matched asymptotic analysis of equation 1.0 (in the text). Lines show the composite equation, and scatter points show the numerical solution of the ODE system that it is approximating. Both are plotted over three different timescales: the first panel shows them up to time $5/(k_{\text{on}}l(0))$, the second up to $5/k_{\text{clear}}$ and the third up to $5/k_{\text{off}}$. The fourth and final panel is a repeat of the third with a log scale. The three timescales are very rapid initial binding of peptide to MHC-I (in less than a second), then rapid clearance of the peptide by the vasculature (in minutes) and finally unbinding of peptide from pMHC-I over many hours.

1.2.1 CTL activation probability against numbers of T-cells or DCs

The simulated effect on cytotoxic T-lymphocyte (CTL) activation probability (the fraction of simulated T-cells that become activated) due to the numbers of T-cells or dendritic cells (DCs) in the lymph node (LN) is shown in Figures 1.6a–b, which confirm previously reported results [19]. In these results, the peptide off-rate is set to zero, so that all DC–T-cell interactions lead to successful CTL activation in the simulated time period. This represents a "best-case" scenario in which properties of the administered short peptide are not limiting for T-cell response and increasing the number of either cell type causes activation probability to increase until saturation.

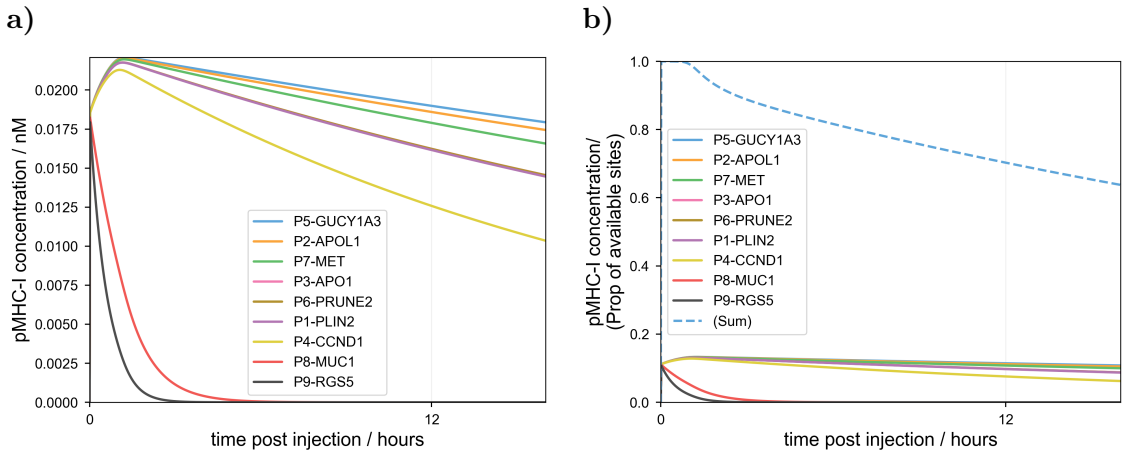


Figure 1.5. Dynamics of the ODE system, Equation 1.0, that describes peptide competition for MHC-I receptors on dendritic cells in the dermis and lymphatics. Vaccine peptides are introduced at time 0. **a)** The concentration of different pMHC-I complexes. They all rise sharply at first, but after free peptide has cleared, peptides with very high off-rates are quickly eliminated. The remaining peptides compete and form an evolving distribution on the dendritic cell's surface. **b)** The proportion of all MHC-I receptors bound to each vaccine peptide in turn. The total of these fractions rapidly falls from 1.0 after free peptide has been cleared through the vasculature, indicating that there will be a time at which these peptides cannot be presented to T-cells.

1.2.2 CTL activation probability against antigen off-rate

The simulated effect of cognate antigen on the probability of CTL activation is shown by Figures 1.6c–d as a function of the antigen off-rate. The probability is characterised by a sigmoidal increase followed by a saturation. The location of the sigmoidal region on the off-rate axis shifts to higher off-rates ('left' in the figure) if A_{dermis} is increased, as shown between panels **c** and **d**. A decrease in P (the time taken for the first DC to migrate from the dermis to the LN) also shifts the region left, according to $A_{\text{LNentry}} = A_{\text{dermis}} e^{-k_{\text{off}} P}$. Changes to the number of antigen in the T-cell–DC contact region, N, or the threshold for response, n, also shift the region, due to the corresponding shift in the transition (Figure 1.1) for the conditional probability of a T-cell–DC interaction leading to CTL activation. Changing the number of DCs changes only the saturation height of the slope, as the shaded regions in panels **c** and **d** indicate.

1.2.3 Comparison to previous work

A component of our model is related to the *in-silico* mouse model reported by Celli *et al*: namely, that dendritic cells (DCs) and T-cells interact in an off-lattice sphere. It is possible to reproduce their results for a consistency check by changing suitable elements of our model, as summarised in Table 1.1. In particular, the T-cell velocity is made to imitate diffusion instead of drawing velocities and free paths from defined distributions, the antigen off-rate

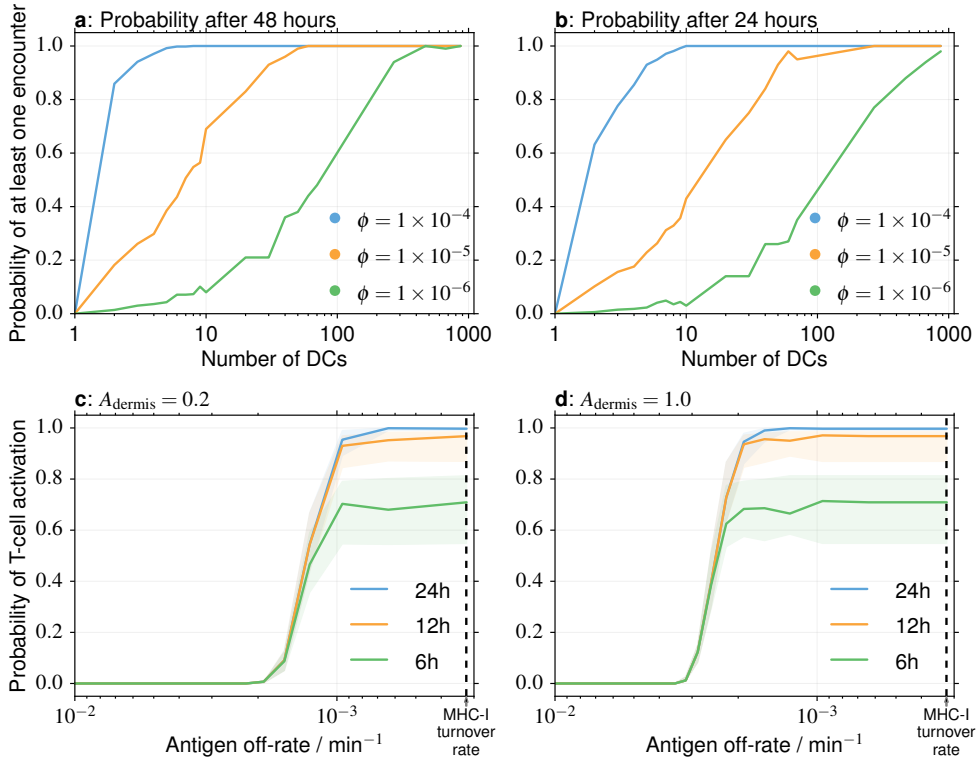


Figure 1.6. The probability of cytotoxic T-lymphocyte (CTL) activation against numbers of dendritic cells or antigen off-rate, k_{off}. In all plots, time $t = 0$ is defined as the time at which the first dendritic cell (DC) arrives in the lymph node. Where parameters are not specified, their values are as presented in Table 1.0. **a, b:** Probability of activation of at least one simulated T-cell in (a) 48 hours or (b) 24 hours, against number of DCs. A very low antigen off-rate is used such that any interaction between cells results in successful CTL activation. Curves for T-cell precursor frequencies between 10^{-6} and 10^{-4} are shown. The greater the numbers of either kind of cell, the greater the chance of successful CTL activation. Both plots are an ensemble average over 1000 replicates. **c, d:** Probability of T-cell activation after the specified number of hours versus peptide off-rate after (c) 20% or (d) 100% of MHC-I receptors bind to cognate antigen at the site of vaccination. The probability is characterised by a sigmoidal increase followed by a saturation. Solid lines show the probability for 720 DCs and shaded regions show the probabilities between 500 and 1000 DCs. The dotted line in each plot indicates the turnover rate of MHC-I receptors [139]. Both plots are an ensemble average over 100 replicates.

is set to zero and the initial cognate to total antigen ratio to unity, so that interactions are always successful. The first comparison between Celli *et al*'s results and our own is shown in Figure 1.7. The left panels show the case in which simulated dynamics (velocity and free path distributions) are made to match Celli *et al*'s; the predicted activation probabilities subsequently also match. The right panels show the case in which simulated dynamics are as used in our other results, which incurs a difference between our results and Celli *et al*'s. Figure 1.8 shows another comparison. As before, left and right panels show Celli *et al*'s and our own chosen dynamics, respectively. It should be expected that the curves on both of the left panels match. Panel (i) in each case shows the probability of T-cell activation as a function of time up to one hour. There is a large difference in the probabilities predicted

Parameter	**Our model**	**Celli *et al*'s model**
T-cell motion	Run-and-tumble	Diffusion
Velocity distribution	Gamma	Gaussian
Mean 3D speed	10µm/min	22.6µm/min
Variance of 3D speed	44.3µm/min^2	90.4µm/min^2
Mean free path	25µm	0 or $v\Delta t$
Antigen off-rate	k_{off}	N/A (zero)
DC contact radius	20µm	12µm
LN radius	500µm	400µm
Modelled animal	Human	Mouse
Modelled cells	CD8$^+$	CD4$^+$

Table 1.1. A summary of the changes that must be made to the model in order to reproduce the results of Celli *et al*.

by Celli *et al* and our models regardless of chosen dynamics, but the reasons for this are unclear. Panel (ii) in each case shows the numbers of T-cells and DCs required for a 50% probability of at least 1 T-cell activation in 24 hours, for both our model and an analytical expression by Celli *et al*. The data and the expression generally agree when diffusion-like cell dynamics are chosen.

1.2.4 Motion of dendritic cells

To approximate the error due to assuming DCs are stationary, we also considered a version of the model in which DCs are allowed to move. Figures 1.9a and 1.9c show the change in probability of CTL activation when DCs are allowed to move at random throughout the sphere. The greater the DC velocity w, the greater the increase in CTL activation probability. However, for velocities $w < 6$µm/min, a relatively high DC velocity according to experimental estimates, the impact on modelling predictions is weak and comparable to the standard error of the stationary model's output (for 100 statistical replicates per datum). Higher DC velocities up to 10µm/min still do not increase predicted probabilities by more than 0.1. The impact on predicted probabilities is most pronounced in a regime where antigen is not limiting, as in Figure 1.9a.

Figures 1.9b and 1.9d show the difference in CTL activation probability when initial DC migration is explicitly modelled instead of adding simulated DCs directly to random positions. In general, ignoring migration causes an overestimate of activation probability at low velocities and an underestimate at high velocities. At low velocities, DCs are confined to the outer edges of the LN for long times and do not sample as much of the volume. At high velocities, DCs move rapidly through the LN and sample a larger volume than a stationary DC would for the migration period. The effect of numbers of DCs can be explained similarly. Ignoring migration causes an under-/over- estimate for small/large

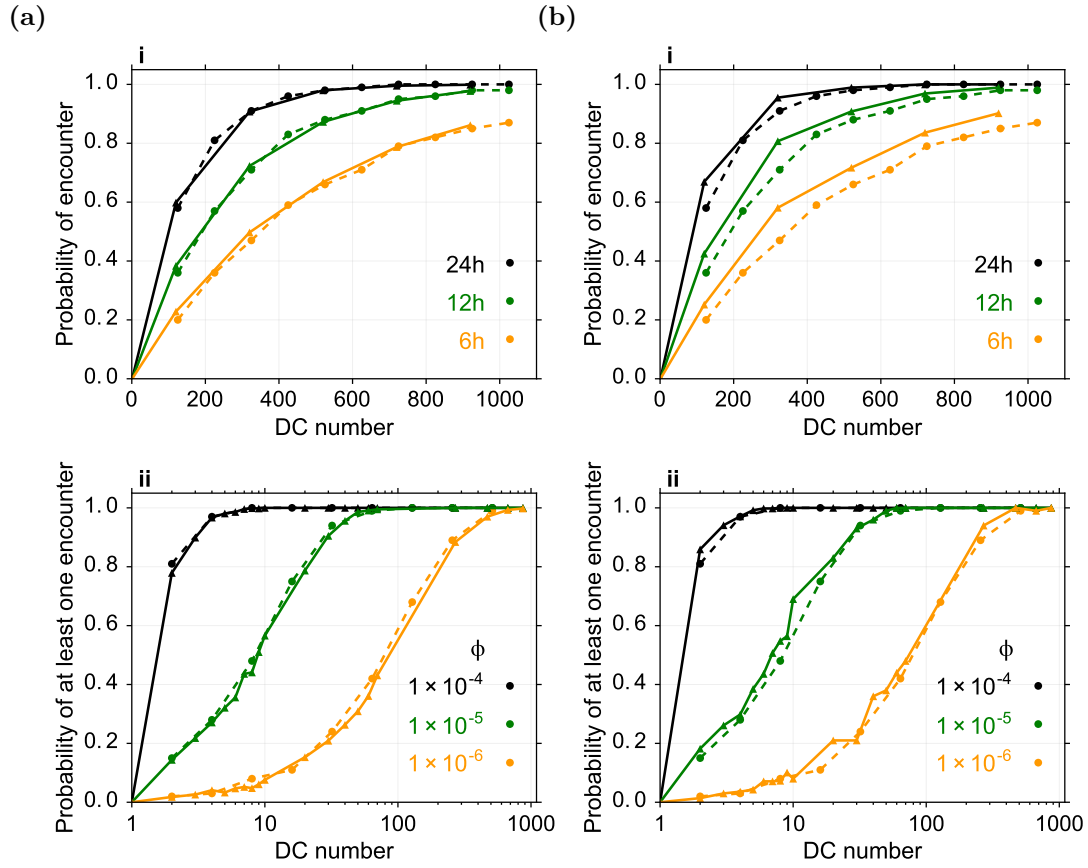


Figure 1.7. Comparison with figure 2 of Celli *et al*'s report. Dotted lines with circles show their results, solid lines with triangles are the results of the model presented here, with parameter values as discussed in the text. Time 0 is defined as the time that the first dendritic cell enters the lymph node. Simulated cell dynamics in **a)** match those of Celli *et al*'s model, whilst the plots in **b)** use our own choice of dynamics and thus results differ slightly. In particular, the velocity is sampled from a gamma distribution with mean 10.1µm/min and cells move unimpeded in one direction for a duration drawn from a Gaussian distribution, to match observations of cell motility [26].

numbers of DCs compared to simulations with intermediate velocities (Figure 1.9b). Moving DCs sample a greater volume of the LN than stationary ones, so activation probability is increased for small numbers of DCs. However, if the number of DCs is large, then the volume surveyed by each arriving DC overlaps. Randomly scattering the DCs in this case increases the surveyed volume at early times and so yields a higher activation probability. Considering changes in antigen off-rate with 720 simulated DCs, as presented in Figure 1.9d, there is essentially no difference in activation probability with changes in motility for very low or very high off-rates. In the transition region, the volume of the LN surveyed by DCs before losing their cognate antigen becomes important, so activation probability differs with velocity as described above. In both panels b and d, the predictions by the stationary model lie between the predictions for $w = 2$µm/min and $w = 6$µm/min for the entire parameter region, indicating that the error due to ignoring movement during DC migration is no

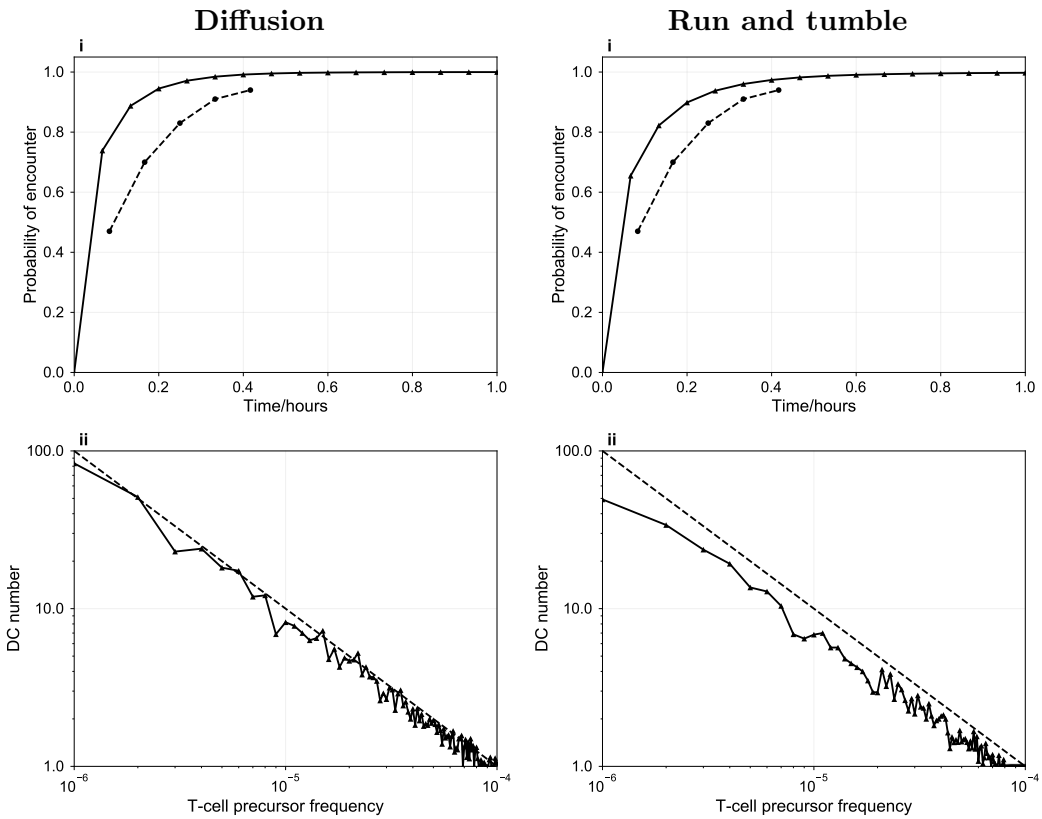


Figure 1.8. Another comparison with figure 2 of Celli *et al*'s report. As before, dotted lines with circles show their results and solid lines with triangles are the results of the model presented here, the left plots use simulated dynamics as in Celli *et al*'s model and the right plots use our own choice of dynamics. Panel (i) in each case shows the probability of T-cell activation as a function of time up to one hour, in the presence of 20,000 DCs. Panel (ii) in each case shows the numbers of T-cells and DCs required for a 50% probability of at least 1 T-cell activation within 24 hours. In this panel, the points are predictions from our model, whereas the dotted line is an analytical expression produced by Celli *et al*.

greater than the error in our uncertainty of the DC's true velocity.

1.2.5 Stochastic entry of dendritic cells

In all results shown in this chapter, DCs are assumed to reach the lymph node at a constant rate $\lambda = (D-1)/p$, where D is the total number of DCs and the last DC arrives at time $t = p$. $(D-1)$ is used because the first DC arrives at time $t = 0$ by definition. It would be more realistic to assume that they arrive according to a Poisson distribution with rate λ. Figure 1.10 shows that the probability of T-cell activation after 48 hours is unaffected by the choice between these two assumptions. All other results thus use the simpler assumption that DCs arrive linearly spaced in time.

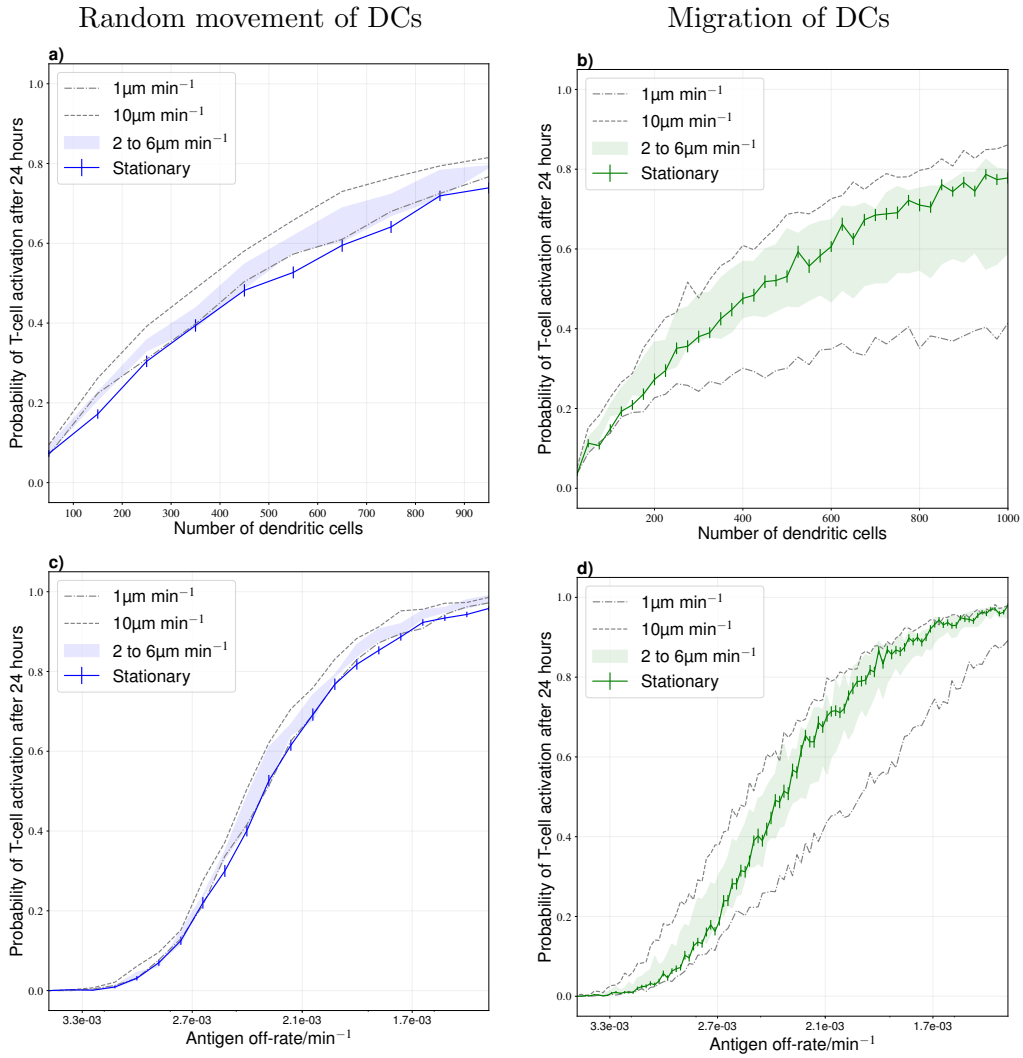


Figure 1.9. The difference made to cytotoxic T-lymphocyte (CTL) activation probability by modelling dendritic cell (DC) motion (see Section 1.1.5). All plots represent an ensemble average over 100 replicates. Left plots (**a** & **c**) show the difference due to random DC motion, and right plots (**b** & **d**) show the difference made by explicitly modelling entry of DCs into the lymph node (LN) from its outer surface. The probability is plotted 24 hours after the first DC arrival in the LN for a range of values of **(c,d)** the peptide antigen off-rates and **(a,b)** numbers of dendritic cells. The solid line shows the probability of activation for the basic model, in which stationary DCs are placed at random positions throughout the LN. The shaded region shows the probabilities for the altered model with velocities between 2 and 6μm/min, and the dotted lines show it for velocities of 1μm/min and 10μm/min respectively. Standard error bars are included for the static model, to compare deviations due to DC movement with statistical noise.

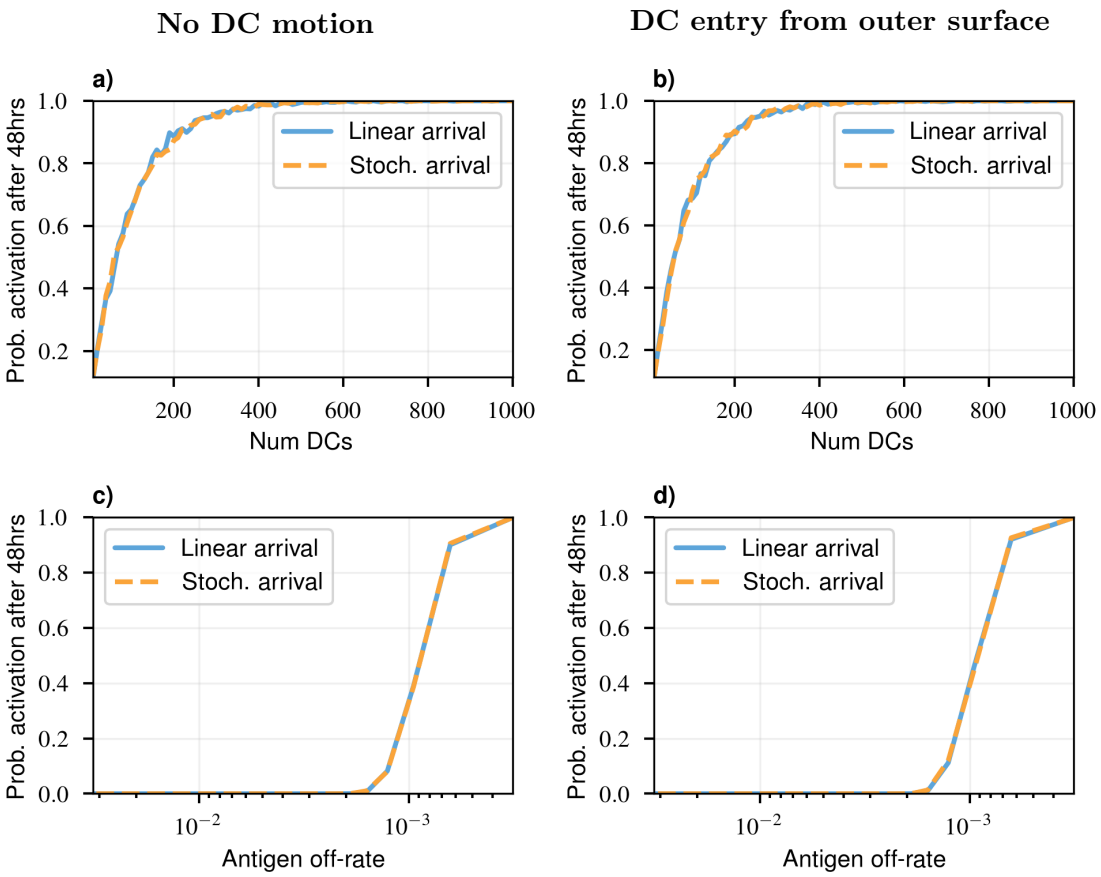


Figure 1.10. The difference in results caused by making dendritic cell (DC) arrival times stochastic instead of linearly spaced. The probability of T-cell activation as a function of **(a, b)** the number of DCs and **(c,d)** antigen off-rate is plotted. Left plots show this for the model in which DCs arrive instantaneously at random positions. Right plots show it for the model in which DCs travel from the LN surface to those positions (as in Section 1.2.4). The number of simulated T-cells is 50, the number of simulated repeats is 100, the time between the first and last arrival is 48 hours and the number of DCs in the lower plots is 720. In all results, no visible difference can be seen between the two cases.

1.2.6 Analysis of the sensitivity of CTL activation probability

We performed sensitivity analyses on the model as described in Section 1.1.7 and shown in Figure 1.11. Exact values of indices are displayed in Table 1.2. We used three techniques: random forest, eFAST and a simple local analysis. Random forest yields estimates of the importance of each parameter, eFAST yields a 'sensitivity' and a 'total sensitivity' index for each, and the local analysis produces the relative impact of perturbing each parameter on the probability of T-cell activation for a specific parameter set. We performed analyses with two ranges of k_{off}: one within and one beyond the transition region described in Section 1.2.2.

As expected, only factors relating to cell interactions are important beyond the transition region, where antigen off-rate does not influence CTL activation at long times. Within the

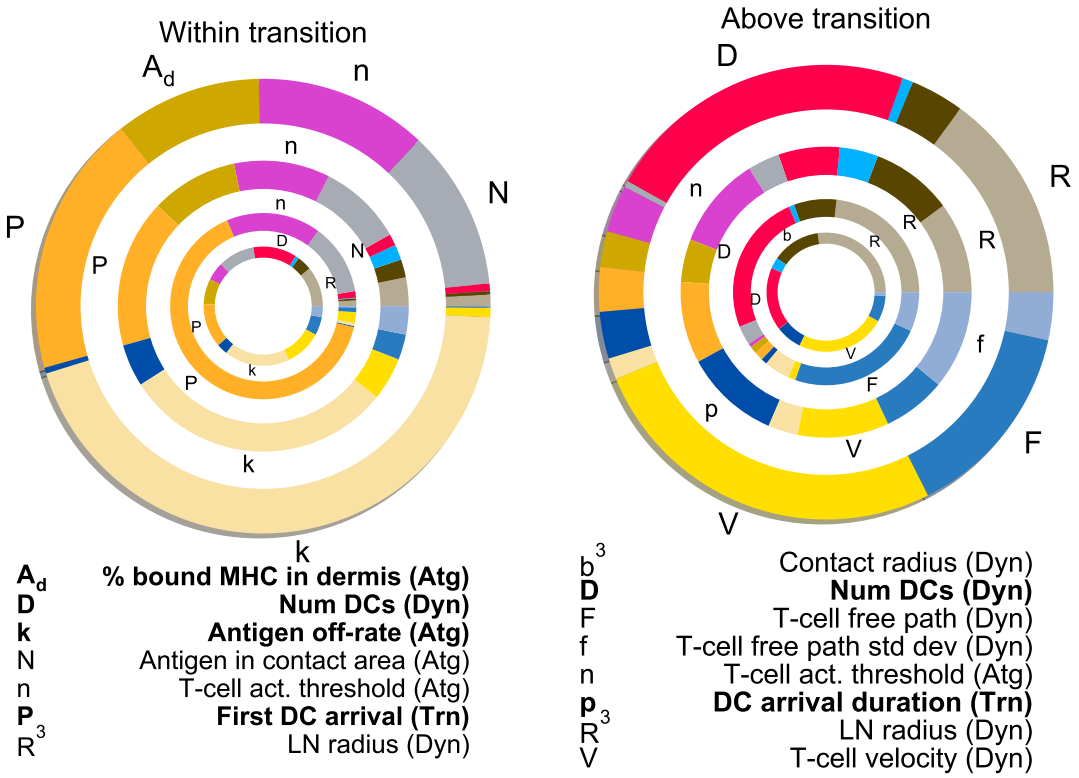


Figure 1.11. Sensitivity analysis of the model either 'beyond' the transition region shown in Figures 1.6c-d or within it, with parameters varied as described in Section 1.1.7. All simulations are an ensemble average over 10 replicates. For each plot, rings from the outside to the inside respectively show proportional sensitivity according to sensitivity indices from eFAST, total sensitivity indices from eFAST, importance factors from a random forest and a local sensitivity analysis. Parameters that occupy more than 10% of the arc length of a ring are labelled and listed below the plots. Bold parameters indicate those controllable by vaccination design (administration or composition). **(Atg)**, **(Trn)** and **(Dyn)** indicates parameters related to properties of the **a**ntigen, **t**ransit of dendritic cells and T-cell–dendritic cell interaction **d**ynamics respectively. The colours of the segments refer to the same parameters in each ring.

transition, CTL activation depends upon sufficient antigen being present, so factors relating to antigen and rapid transit of DCs are most important. Interestingly, total sensitivity indices for many parameters are of a similar size, indicating that there is a high degree of correlation between model parameters.

1.3 Discussion and conclusions

1.3.0 Precise dynamics of single peptide binding in the dermis are not an important factor for T-cell activation probability

The ODE system of peptide dynamics in the dermis, Equation 1.0, has several flaws that have been ignored. Firstly, the original model by Mailen Kootsey, Jonathan Wagg *et al*

Parameter	Sensitivity index	Total sensitivity index	Random forest importance	Local sensitivity
Within transition				
LN radius cubed	0.01	0.05	0.01	0.11
Contact radius cubed	0.00	0.03	0.00	0.03
Num T-cells	0.00	0.03	0.00	0.01
Num DCs	0.00	0.02	0.01	0.12
Antigen in contact area	0.09	0.15	0.12	0.10
T-cell act. threshold	0.10	0.18	0.16	0.05
Dermis cog. ag ratio	0.09	0.16	0.00	0.07
First DC arrival	0.16	0.27	0.65	0.12
DC arrival duration	0.00	0.08	0.00	0.03
Antigen off-rate	0.37	0.50	0.00	0.18
T-cell velocity	0.01	0.08	0.02	0.10
T-cell free path mean	0.00	0.05	0.01	0.05
T-cell free path std. dev.	0.00	0.05	0.00	0.03
Sum	0.83	1.64	1.00	1.00
Above transition				
LN radius cubed	0.03	0.78	0.23	0.29
Contact radius cubed	0.01	0.70	0.07	0.07
Num T-cells	0.00	0.33	0.01	0.02
Num DCs	0.05	0.53	0.25	0.15
Antigen in contact area	0.00	0.27	0.03	0.00
T-cell act. threshold	0.01	0.79	0.01	0.00
Dermis cog. ag ratio	0.01	0.37	0.01	0.00
First DC arrival	0.01	0.70	0.02	0.00
DC arrival duration	0.01	0.82	0.01	0.08
Antigen off-rate	0.00	0.26	0.04	0.00
T-cell velocity	0.06	0.78	0.01	0.24
T-cell free path mean	0.03	0.54	0.24	0.13
T-cell free path std. dev.	0.01	0.85	0.07	0.02
Sum	0.22	7.73	1.00	1.00

Table 1.2. Results of the sensitivity analyses described in Section 1.1.7 and presented in Figure 1.11. Data are organised into columns for sensitivity indices from eFAST, important factors from Random Forest or relative sensitivity calculated from a simple, local sensitivity analysis. The table is split into two; each part represents an analysis performed in a different range of k_{off}.

[104] treats receptors r as "dendritic cells" in order to address cell diffusion, but the key biological entity here is actually MHC-I receptors. Since each cell should have approximately the same number of receptors and we know that this quantity is of order 100,000 [10, 114–117], we ignore the terminology of the model, removed terms related to diffusive entry of new dendritic cells and assume that we are dealing with receptors. The second issue is that we ignore the process by which pMHC-I complexes are internalised and recycled, taking the bound peptide with them. However, because the clearance rate of free peptides is so fast (10/hour [107]), the recycling rate may be considered a simple addition to the peptide off-rate. The third issue arises when multiple peptides are considered. Peptides in the dermis compete not just with each other for MHC-I receptors, but also with "self"-peptides: the peptides usually expressed on MHC-I are fragments of peptides that have

been produced by the cell itself. Cytoplasmic peptidases cleave peptides into fragments, and these fragments can be chaperoned by a transporter named TAP into the endoplasmic reticulum and bound to MHC-I. The resulting peptide–MHC (pMHC) complex is then transported to the cell membrane. For a vaccine peptide to bind to MHC-I on the cell surface, this self peptide must first unbind. The presence of these self-peptides and their unbinding is, however, not modelled. The ensemble of self-peptides is vast and each has its own off-rate, making accurate modelling difficult. Instead, we use an experimental result that a desired peptide will typically occupy 10–15% of MHC-I receptors on dendritic cells after peptide introduction and clearance [140, 141] to define a factor A_{dermis}, the proportion of MHC-I on the dendritic cell bound to cognate antigen when they leave the dermis. A related issue is that we do not consider whether short peptides may be taken up by the dendritic cell and 'cross-presented' onto MHC-I. Short peptides have a half life of seconds within dendritic cells [105–107], and once they are cleaved a single time, they become too short to be presented on MHC-I (which only binds peptides that are 8-11 amino acids long). We hence ignore this route of presentation for the administered self-peptides. Considering both this and that free peptides in the dermis are rapidly cleared by the vasculature, the distribution of vaccine peptides quickly becomes characterised by their relative off-rates from pMHC-I complexes.

We also do not consider the fate of peptide that is cleared away from the dermis, nor peptide that unbinds from dendritic cells. There are multiple reasons for this. The peptides were injected subcutaneously, where there are few lymphatic vessels. Most clearance of small molecules in this region is hence through the vasculature, not the lymphatics [108]. It was found in sheep that "the rate of filtration and reabsorption of fluid across the vascular capillaries is high (20–40 L/day) in comparison to the small quantities of interstitial fluid drained by the lymph (2–4 L/day). Consequently, small molecules ($< 2,000$ molecular weight) which can access lymph and blood capillaries equally are predominantly cleared by the blood vessels, whereas particulates and molecules of increasing molecular size appear to favo[u]r drainage into the more open capillaries of the lymph." [108]. The authors of another study in sheep found that more than 96% of clearance of small molecules is through the vasculature rather than the lymphatics [108]. By comparison to the aforementioned cutoff molecular weight of 2000, short peptides of size 9 amino acids have an average molecular weight of around 990, and therefore we would expect the vast majority of such peptides to be cleared by the vasculature.

Another factor that would reduce the number of short peptides reaching the lymph node is due to the cells lining the lymphatic vessels, that each themselves express MHC-I, including antigen presenting cells such as macrophages. These cells are likely to capture many peptides

that reach the lymphatics. Similarly, those peptides that are swept into the interstitial space surrounding the lymph node are likely to be taken up by the macrophages in this space before penetrating the lymph node itself. Finally, vascular clearance of molecules the size of short peptides within the lymph node is also very fast, and so the small number that do reach the lymph node are likely to be cleared before binding to dendritic cells present. However, the number of remaining peptides that bind to DCs, though small, may not be no smaller than the number of peptides that are presented on MHC-I after uptake by dendritic cells, and so may be worth explicitly modelling in future work, for example with ODEs describing their transport.

Another critical assumption we have made is that the on-rates of all vaccine peptides are the same, in the absence of any data about them. This can be justified by a previous finding that peptide binding affinity for MHC-I is controlled mainly by off-rate [142]. Little variation between on-rates has been found previously (unlike off-rates) [142, 143], and the timescale for binding is very short compared to the timescale for unbinding, particularly due to the high peptide concentrations used. Possibly, on-rates are diffusion-limited or the initial contact event is a similar rate among all peptides, and the rate of the subsequent conformational change would be irrelevant as long as the on-rate is fast compared to the off-rate. We hence assume that the peptide on-rates are equal. This chapter of the thesis considers vaccination with only a single peptide at a time, but where peptide competition is important (as in Chapter 2), this assumption could affect results.

Following these assumptions produces an ODE system in which peptide rapidly binds to all available sites (typically 10–15% [140, 141]) on the MHC-I (in roughly equal proportions if multiple peptides are considered). All free peptide is rapidly cleared by the vasculature, and no peptide that enters the DCs (if any) is subsequently presented on MHC-I. This leaves a distribution of pMHC-I complexes on dendritic cells that depends only on time and pMHC-I off-rate. The ODE system is thus not important to model in the absence of peptide competition, and may instead simply be summarised by a value A_{dermis}, the proportion of MHC-I that are bound to cognate antigen upon DC migration from the dermis.

1.3.1 Model observations are not sensitive to dendritic cell motion

Section 1.2.4 contains a description of the implementation of DC movement into the LN and of random movement of DCs within the LN. Measurements of a typical DC velocity vary. For example, one measurement using two-photon microscopy [26] gave a median 3D velocity of 6.6µm/min, whereas another [111] gave a value corresponding to 3.3µm/min. This discrepancy could be due to the time of measurement; DCs in the LN are known to lose motility over time before stopping after 48 hours [12, 12–17]. This loss of motility

could be due to changes in either velocity or mean free path, but either case improves the approximation of DCs as stationary.

We found that the difference made to model predictions by ignoring DC migration gives a prediction equal to one in the range of migration velocities $2 \leq w \leq 6$µm/min. However, we modelled DC migration as straight-line movement from the perimeter, which is not true to the biology. It neglects collisions with cells in the node, and the fact that DCs move first to the high endothelial venules. On the other hand, as we do not model T-cell ingress, HEV modelling is unnecessary. Straight-line movement represents a first approximation, whose effect on predicted probabilities of T-cell activation is expected to be similar to that of modelling more detailed DC movement. We hence conclude that it is valid to ignore movement of DCs during their migration through the LN.

Random motion of DCs has a greater impact on model predictions than does linear DC migration; the higher the velocity, the greater the predicted probability of T-cell activation. However, DC velocities $w \leq 6$µm/min have a relatively weak impact on model predictions (increasing predicted probabilities by no more than 0.1 over the parameter regimes tested). Considering both this and the uncertainty in DC velocities reported in the literature, we choose to ignore the random motion of DCs and use a model in which only T-cells can move, thereby greatly increasing computational efficiency. All other results presented in this work use the 'basic' model with no DC movement.

We also considered whether results differ if DC arrival times are assumed to be stochastic (Poisson-distributed) rather than linearly spaced. It was found that the two assumptions produce almost identical results; however, this is likely due to the high number of model repeats (100) used to generate results. If fewer results were to be used, both of the expressions for the probability of T-cell activation would become noisy, but some of the noise in one of the data sets would be due to the stochastic entry times. In all analyses, we use a high number of repeats and thus it would be valid to ignore stochastic entry times of dendritic cells. When a smaller number of repeats is used or stochastic variability is important, however, it might be prudent to not ignore this source of noise.

1.3.2 The role of peptide off-rate and DC/T-cell numbers in CTL activation

The primary aim of the work in this chapter was to determine the factors that influence the probability of cytotoxic T-lymphocyte (CTL) activation following short peptide vaccination. Figure 1.6 shows that the predicted probability increases with time, the number of dendritic cells (DCs), T-cell precursor frequency and decreasing peptide-MHC (pMHC)

off-rate. The fundamental factor is the off-rate, against which activation probability exhibits three characteristic behaviours (Figures 1.6c-d): a region of zero probability for fast off-rates, a region of high probability for slow off-rates, and a transition region between the other two that is highly sensitive to changes in off-rate. This behaviour is inherited from the sharp transition in the binomial equation (Figure 1.1). Such transition-like behaviour has been observed in the amount of cytokine produced by CTLs as a function of pMHC density when the latter are immobilised on a surface [109, 144], which is consistent with conclusions drawn from our model.

For sufficiently slow off-rates beyond the transition, the probability of CTL activation is sensitive to patient-specific factors such as the number of recruited DCs or the precursor frequency, as shown in Figures 1.6a-b. In this region, activation probability is not sensitive to or limited by properties of the antigen.

Within the transition, small changes in off-rate are predicted to cause dramatic changes in CTL activation probability. Thus, vaccination with short peptides with off-rates in this range may show higher levels of inter-patient variability in T-cell activation response. Activation probability within the transition is insensitive to the number of recruited DCs, seen in the narrow thickness of the shaded regions in Figures 1.6c-d. CTL activation probability in this parameter regime could be increased by shifting the transition region, by reducing the time taken for DCs to reach the lymph node (LN) or increasing the proportion of MHC-I receptors bound to cognate antigen in the dermis (by more efficiently loading peptides).

Below the transition, off-rates are too high and there is no prospect for improving CTL activation probability without first reducing the off-rate. We conclude that the initial events in a short-peptide vaccination that determine CTL activation depend foremost on pMHC presentation and off-rate. To improve the probability of CTL activation, peptides should be optimised for slow off-rates, high presentation at the vaccination site and fast migration time to draining LNs. For example, performing multiple single vaccinations instead of a single multiple-peptide dose would reduce competition for MHC-I receptors and thus increase MHC-I receptor occupancy for each peptide. DC migration time could be improved with alternate delivery mechanisms, such as intranodal DC vaccination.

1.3.3 Differences between this model and prior work

To check for consistency with previous work, we altered our model parameters and cellular dynamics to match those utilised by Celli *et al*, and attempted to reproduce their predictions. Celli *et al* modelled $CD4^+$ T lymphocytes, whereas we model $CD8^+$ cells. We could not find a reason to suggest that the naïve dynamics of these cells differ, so we assume that comparison of our results is valid. Similarly, Celli *et al* parameterised their model for

mice rather than humans, but after scaling parameters and assuming that cell velocities are unchanged between the species, the models are equally applicable to both species. Figures 1.7 and 1.8 show the differences in predictions made by their model and ours, after our parameters and cell dynamics are changed to match their results. Most data match except short-time data (Figure 1.8i). The reasons for short-time predictions not matching, even with diffusion-like motion, are not clear. The differences incurred by the use of gamma-distributed velocities and non-zero free paths can be explained with the expected range of T-cells. The reduction in mean velocity and choice of gamma distribution lowers overall cell displacement. However, the non-zero free path F increases displacement. This can be seen by relating the mean free path to motility through the time-step: assuming a constant mean free path F and velocity v, we can define a timestep of motion $\Delta t = \frac{F}{v}$, as in a random walk. The motility $M \propto \frac{F^2}{\Delta t} = Fv$, and the expected 3D displacement after a time $t \gg \Delta t$ is $\sqrt{6Mt} \propto \sqrt{Fvt}$. The changes to the velocity and mean free path together slightly reduce the volume that T-cells can search in finite time and thus the probability that they become activated by a DC.

1.3.4 T-cell movement and range

Simulated T-cells in the lymph node model draw velocities from a gamma distribution and free paths from a Gaussian distribution. Initial simulations assumed Brownian motion, but this is not appropriate for motion of T-cells. It has been reported that T-cells follow a Lévy walk [145], characterised by many local steps interspersed with occasional longer steps. The weakness of the Lévy distribution for this application is that it has an infinite mean, which complicates control of T-cell velocity as a parameter. The distribution is a special case of the inverse gamma distribution, which does not necessarily have an infinite mean, so this was one of the distributions we attempted to fit to experimental data of instantaneous T-cell velocities [26] in Figure 1.3. It is, however, a much poorer fit than the gamma distribution. The gamma distribution is qualitatively similar to the inverse gamma distribution in that it is heavy-tailed, which reproduces the desired behaviour of many small jumps interspersed with rarer large jumps. With this in mind and because it gave the best fit, we draw T-cell velocities from a gamma distribution in our model.

An additional property of interest is the free path of T-cells, or the distance of velocity jumps. This can be determined by fitting to measurements of the displacement of T-cells in a finite time t. As argued in the previous section, this can be expected to be $\sqrt{\chi Fvt}$, where v is the velocity, F is the free path and χ is a constant. Fitting this function to the displacement of simulated T-cells gives a constant of proportionality as approximately 4.3, and hence a free path of T-cells of around 15-40µm, depending on experimental conditions.

We choose a Gaussian distributed free path of 25±3µm for our simulations. A more detailed calculation would require an analytical expression for the motility and range of T-cells in a velocity jump process, but this is non-trivial. We will revisit this problem in Chapter 2.

1.3.5 The sensitive parameters in our model support expectations

To ensure model robustness, sensitivity analyses were carried out both within and beyond the transition. The parameters found to be important were as expected from intuition. Within the transition, where the pMHC off-rate is barely sufficient for CTL activation, the model is most sensitive to parameters relating to antigen-MHC-I receptor binding such as A_{dermis} or the time taken for the first DC arrival, P. Beyond the transition, where activation probability is not limited by off-rate, only parameters relating to T-cell–DC interactions such as the number of DCs are important. This supports the conclusions given in Section 1.3.2.

In most cases, the sensitivity indices and random forest importances described in Section 1.2.6 agree on which parameters are important. The differences likely stem from the details of each algorithm: the growth process for the random forest chooses parameters that are best able to split data into two groups, whilst eFAST compares how the variance of different parameters affects the variance of the output. For example, $e^{-k_{\text{off}}P}$ determines the location of the transition and thus P can effectively split data that crosses the transition, but does not contribute greatly to output variance away from the transition. Additionally, the sensitivity index for P quantifies output variance due to changes in P alone, but P appears only as a product with the off-rate, k_{off}, whilst k_{off} is used alone to calculate decay of antigen within the LN (Eq. 1.1).

Total sensitivity indices also disagree with the other two sets of indices, but we choose to discount these as they are heavily influenced by interactions between parameters. Sensitivity to other parameters is low, indicating that the model and its conclusions are robust.

The simple local sensitivity analysis, which involved finding the probability of activation for a particular dataset and then increasing parameters by 5% one at a time, generally agrees with the analysis by eFAST (the parameters identified as important are the same). This is likely to be because both techniques consider how the change in output compares to changes in each input, which is distinct from random forest (as described above).

1.3.6 Conclusions and future work

The objectives of Chapters 1 and 2 are to identify a mechanism that could hypothetically explain the lack of efficacy seen in IMA901, representing a cancer-immune response with a potential bottleneck in the first three steps of the cycle. We produced a model of the events leading to initiation of CTL activation within the lymph node and quantified various

assumptions made by us and others. We analysed the sensitivity of parameters in our model to ensure that our model's conclusions are robust. By considering the properties of antigen bound to DCs and how this affects CTL activation probability, we found that peptide-MHC off-rates that are significantly faster than a certain transition value lead to zero probability of CTL activation. For slightly slower off-rates, there is a rapid transition to a probability that is independent of further decreases in off-rate. For these slower off-rates, the only factors naturally varying among individuals that CTL activation probability depends on are the numbers of interacting T-cells and DCs. The transition off-rate depends upon factors such as the time taken by DCs to migrate to the draining LN.

The off-rate hence confirms our hypothesis that a vaccine property can significantly affect the probability of successful T-cell activation, and generates the new hypothesis that the off-rate of peptides from pMHC-I complexes is critical for short peptide vaccinations. This work could have significant impact on short-peptide cancer vaccine design. We predict that certain tumour-associated peptides used in vaccines may have zero probability of ever activating CTLs in patients, based only on properties of the peptide. This may explain the lack of success in many cancer vaccine clinical trials to date. Peptides should be selected first for a pMHC off-rate low enough to have a chance of CTL activation, and then patient-specific factors should be optimised to improve this chance. The vaccination site may be cooled to reduce peptide clearance, multiple vaccination sites that drain to the same lymph node may be used to increase numbers of recruited DCs or reduce competition for MHC-I receptors in multi-peptide vaccinations, or intranodal DC vaccination may be considered.

There is considerable scope for further work from this study. We could consider additional complexities such as T-cell ingress, egress and receptor affinity to quantify CTL proliferation and compare to the work of others [70, 72, 146]. The model could be applicable to other types of vaccination such as mRNA or DC delivery, or to tumour-draining LNs, though these would all require consideration of cross-presentation. The important of geometry could be considered: notably, human lymph nodes are made of several independent lobules, but each lobule is approximately the size of a mouse lymph node. This might indicate that the size of the functional unit of the lymph node is conserved or close to an optimum, which may be interesting to study. This also lends to the study of what happens in the lymph node under inflammation after T-cell activation and how this affects further lymphocyte–dendritic cell dynamics. Other aspects of lymph node geometry such as its shape (which is closer to a bean than a sphere) or the fibroreticular network present within could be investigated. We have shown that various simplifications and assumptions in the model are valid, thus a simpler first-passage model as previously reported [19] could be utilised in future, by comparing expected interaction time with antigen off-rate. However,

such an approach should take into account complexities such as staggered arrival of DCs and a 'run and tumble' description of T-cell dynamics.

Although we have quantified how short-peptide vaccinations (can fail to) drive CTL activation, we aimed more generally to identify why patients in the short-peptide cancer vaccination, IMA901, responded to so few of the constituent peptides. The model in this chapter has indicated that peptide off-rates from MHC-I could be an important factor, but no link to IMA901 has yet been established, so this hypothesis cannot be confirmed or refuted. To proceed in Chapter 2, we will utilise the model to simulate a virtual clinical trial and attempt to reproduce the real-world results of IMA901 and explain the observed results.

1.A Appendix: Sensitivity analysis techniques

In this section, we give a brief explanation of the sensitivity analysis techniques used.

1.A.0 Random Forest

A random forest is a forest of decision trees. Given a set of N data with P parameters and O outputs, a forest is grown for each output o. Each tree is grown from N samples with replacement. At each node of the tree, $p < P$ random parameters are chosen and used to split the node into branches that best explain the change in o. This process is repeated until the tree reaches its largest extent [134]. The decision tree is then able to classify or predict o given a new parameter set. The classification or prediction of o by a forest of such trees is defined as the mode or mean, respectively, of the constituent trees.
The growth process of the tree results in an estimate of the importance of each parameter, based on their relative influence on node splits. These estimates can be used for a sensitivity analysis.

1.A.1 Variance-based sensitivity analysis

An alternative approach is to consider the contribution of each parameter x_i to the variance of the expected model output y_o, $V_i = V(E(y_o|x_i))$ [136–138]. The total variance can be written,

$$V_{\text{total}} = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \sum_i \sum_{j>i} \sum_{k>j} V_{ijk} + \ldots + V_{0\cdots P}, \tag{1.A.0}$$

where $V_{ij} = V(E(y_o|x_i, x_j))$, with other terms defined by analogy. We define the sensitivity index S_i by,

$$S_i = \frac{V_i}{V_{\text{total}}}, \tag{1.A.1}$$

where $0 \leq S_i \leq 1$ gives an estimate of the proportion of the variance that is due to parameter i. However, by Eq. 1.A.0 we see that $\sum_i S_i < 1$ if the interaction terms are non-zero. The total sensitivity index takes into account these higher order terms,

$$S_{Ti} = 1 - \frac{V_{-i}}{V_{\text{total}}}, \tag{1.A.2}$$

where V_{-i} is the sum of all terms not including i,

$$V_{-i} = V_{\text{total}} - V_i - \sum_j V_{ij} - \ldots - V_{0\cdots P}. \tag{1.A.3}$$

We present two methods for determining these indices [136–138]. The first is to generate N samples of data from a model f and to estimate the variance as,

$$\hat{V}_{\text{total}} = \frac{1}{N}\sum_{m=1}^{N}[f(X_m)]^2 - \hat{f_0}^2, \tag{1.A.4}$$

where $X_m = (x_{1m}, x_{2m}, x_{3m}, \ldots, x_{Pm})$ is the mth sample of parameter values and f_0 is the mean of the data $\frac{1}{N}\sum_{m=1}^{N} f(X_m)$. V_i is estimated from the covariance of two data sets where parameter x_i is held fixed,

$$\hat{V}_i = \frac{1}{N}\sum_{m=1}^{N}[f(X^{(1)}_{(-i)m}, x_{im}) f(X^{(2)}_{(-i)m}, x_{im})] - \hat{f_0}^2, \tag{1.A.5}$$

where $X_{(-i)m}$ indicates the values of all parameters except i, and ${X^{(1)}}_{(-i)m}$,${X^{(2)}}_{(-i)m}$ are different samples of each of the parameters $x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_P$. Similarly, V_{-i} is calculated as the covariance of two data sets where everything *except* parameter x_i is held fixed,

$$\hat{V}_{-i} = \frac{1}{N}\sum_{m=1}^{N}[f(X_{(-i)m}, x^{(1)}_{im}) f(X_{(-i)m}, x^{(2)}_{im})] - \hat{f_0}^2. \tag{1.A.6}$$

From these, S_i and S_{-i} can be estimated.

1.A.2 Extended FAST

A second approach is to use extended FAST [137, 138]. Every parameter is associated with an integer frequency ω_j. To calculate the sensitivity index S_i, ω_i is given a high frequency whilst the other parameters take low frequencies. Parameter values are calculated according to,

$$x_j(s) = x_j^{\min} + (x_j^{\max} - x_j^{\min})\big[\frac{1}{2} + \frac{1}{\pi}\sin^{-1}\big(\sin(\omega_j s + \phi_j)\big)\big], \tag{1.A.7}$$

with $0 \leq s \leq 2\pi$ and ϕ_j is a random phase shift for each parameter j. N_ϕ different sets of values of the ϕ_j are produced and the corresponding sets of parameter values are used to

find the outputs $\boldsymbol{y}(s) = f(\boldsymbol{x}(s))$ from a model f. Each set of observables $\boldsymbol{y}$ is then Fourier transformed and the power spectrum $\Lambda_o^2(\omega)$ calculated, which in turn is averaged over the sets of ϕ_j.

It can be shown that the total variance is the sum of all terms in the power spectrum except the zero-frequency term [137],

$$V \approx \hat{V} = 2 \sum_{\omega=1}^{M\omega_i} \Lambda^2(\omega), \tag{1.A.8}$$

where the factor of two comes from the symmetric negative frequency terms and M is the maximum number of harmonics to be considered of the highest frequency, ω_i. If an observable y_o is strongly dependent on a parameter x_i, then it will exhibit high amplitude oscillations at frequency ω_i. This fact allows us to estimate the sensitivity index S_i by considering the harmonics of ω_i,

$$V_i \approx \hat{V}_i = 2 \sum_{p=1}^{M} \Lambda^2(p\omega_i). \tag{1.A.9}$$

The terms of Eq. 1.A.0 with no dependence on x_i are associated with small integer frequencies. If ω_i is chosen such that $\omega_i > 2M$, then the contribution of all such terms can be found by considering all frequencies $\omega < \frac{\omega_i}{2}$,

$$V_{-i} \approx \hat{V}_{-i} = 2 \sum_{\omega=1}^{\frac{\omega_i}{2}} \Lambda^2(\omega). \tag{1.A.10}$$

This procedure is repeated for each sensitivity index S_i in turn.

Chapter 2

What determines efficacy of short-peptide vaccinations? Simulating mechanisms of clinical failure

Note: This chapter is based on a manuscript in preparation for publication. Note also that the methods subsection 2.1.2 and the experimental work detailed within were written and carried out by collaborators (Rachel Darley, Andy van Hatern and Tim J Elliott).

2.0 Introduction

The lymph model model in Chapter 1 yields the prediction that the pMHC-I off-rate is a critical determinant of the probability of T-cell activation following a short-peptide vaccination. In this chapter, I will describe an *in silico* experiment to test this hypothesis with a virtual clinical trial of IMA901. We will subsequently use the virtual clinical trials to attempt to identify why Phase III did not replicate the successes in Phases I and II, and to determine if changes to the vaccine could have improved upon results. Simulated virtual clinical trials are not a novel concept. For example, the overall survival in a phase III clinical trial has been accurately predicted by parameterising a survival model with the phase III results from a different drug and a tumour-kinetic model from the phase II data of both drugs [62], and the effectiveness of a treatment for bone disease in a paediatric population was estimated using a virtual clinical trial [147]. As described in Chapter 1, renal cell carcinoma (RCC) patients in IMA901 were given nine 'short' (defined as 9 amino acid long) peptides intradermally in a combined injection, but most patients gave immune responses to zero or one of the nine peptides. No patients responded to more than three in any phase. To enable comparison to this data, we ran virtual clinical trials using the lymph

node model on a virtual cohort, and quantified how the number of patients giving 0, 1, 2 or 3 immune responses depends on parameters such as the pMHC-I half-life on the surface of dendritic cells. We aim to determine if the hypothesis that the dissociation rate of short peptides from MHC-I forms a mechanism of failed T-cell activation that is consistent with the results of IMA901, to determine if there is a quantifiable difference between the phases of the trial that may explain why phase III did not result in a benefit to overall survival of patients over the standard of care, and to identify changes to the vaccine design that could have improved results.

2.1 Methods

2.1.0 IMA901 trial procedure

Phase	Num. patients	Num. peptides	Visit days
I	27	9+1	1, 2, 3, 8, 15, 22, 36, 64
II	61	9	1, 2, 3, 8, 15, 22, 36, 57, 78, 99, 120, 141, 162, 183, 204, 225, 246
III	96	9	1, 2, 3, 8, 15, 22, 43, 64, 85, 106

Table 2.0. Quantitative details of IMA901 that differ between each phase: the number of patients enrolled, the number of peptides administered to each patient and the days on which (all) peptides were administered. Patients were given approximately 4mg of each peptide on each visit in each phase.

For results to be consistent, virtual clinical trials must match the logistics of the trials of IMA901. Each phase recruited a different number of renal cell carcinoma (RCC) patients and administered short peptides to them in multiple visits. Table 2.0 shows the number of patients recruited and on which days the patients were administered the vaccine, beginning with day 1. There are also operational differences between the phases. In phase I, patients were given a marker peptide from hepatitis B virus (HBV), and in phase III, one arm of patients were given Sunitinib, the standard of care for RCC. Both of these points will be discussed later in the text. We could not identify any differences between the patients recruited for each phase of IMA901. Each had a similar median age, had an advanced stage of the disease, had similar criteria on the previous treatments that they had received, had similar prognostic scores and similar incidences of metastases. The one exception is that Phase II had a lower mean and median time since diagnosis, however, this was not reflected in a change to the prognostic score for this phase.

2.1.1 Virtual clinical trial procedure

The model described in Chapter 1 was used to predict the probability of T-cell activation as a function of input (patient) parameters. It was extended to allow the simulation of multiple waves of dendritic cells (overlapping or not), in order to simulate the vaccination

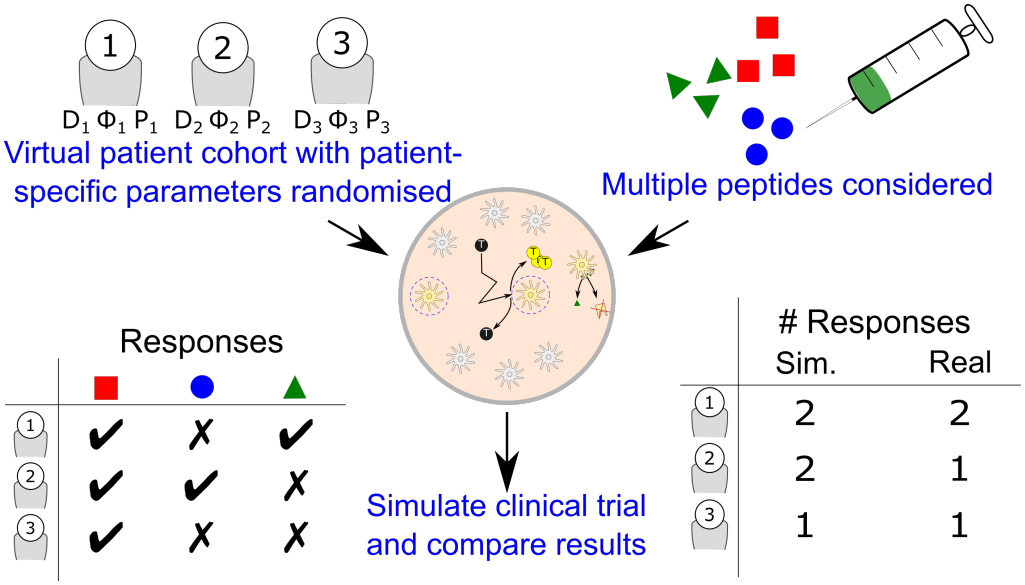


Figure 2.0. Virtual clinical trial schematic: For each virtual trial, a virtual patient cohort is produced with random patient-specific parameters, such as the dermal dendritic cell density. The lymph node model is simulated with each peptide for each virtual patient, and the expected number of peptide responses is calculated for each patient. Three distinct virtual trials are performed to gain an estimate of stochastic variability of results, then the numbers of patients responding to zero, one, two or three peptides are compared to IMA901's results.

schedule of IMA901. The extended model is used to run virtual clinical trials as shown in Figure 2.0 and as follows: a virtual patient cohort equal in size to one of the phases of IMA901 is created, and each patient is given normally distributed random values for each parameter expected to differ between individuals (dendritic cell migration efficiency, T-cell precursor frequency, lymph transit time and spread, and the proportion of MHC-I available to vaccine peptide in the dermis). The ODE system of peptide competition in the dermis described in Section 1.1.1 (page 38) is used to determine the initial abundance of each peptide, using their measured pMHC-I off-rates and assuming equal concentrations and on-rates.

The lymph node model is then run for each virtual patient and peptide in turn (using predicted initial abundances of each peptide), yielding a predicted probability that patients respond to each peptide after the entire vaccination schedule. The sum of these gives the expected number of peptide responses for each patient and in turn the number of patients expected to respond to 0, 1, 2 or 3 peptides, which can be compared to the data reported by the authors of IMA901. The means and standard deviations of the Gaussian distributions for patient parameters (*e.g.* there is a distribution of possible lymph transit times, from which each patient draws a value) can be controlled to find the set of parameter distributions for which virtual trial results match the result of IMA901. For presented results, the lymph node model is repeated 10 times to reduce uncertainty of the probability

of T-cell response. The entire virtual trial is additionally repeated 3–5 times, each with different random patients, in order to gain an estimate of the variability of results due to stochasticity in patient parameters.

Although the phase I trial of IMA901 contains the marker peptide from HBV, the number of responses made to it are not reported or of interest. We wish to match the data on the number of responses made to the nine RCC peptides to predicted responses, and hence HBV is not modelled – its presence would represent only a reduction in the maximum number of MHC-I sites bound in the dermis, which in any case is determined by fitting of parameters to the reported data (see Section 2.2.7).

2.1.2 Measurement of peptide-MHC-I off-rates

The off-rates of each of the peptides used in IMA901 were measured by collaborators (Rachel Darley, Andy van Hatern and Tim J Elliott). They carried out *in vitro* experiments with immobilised human MHC-I, and with mature human B-cells that present MHC-I. In the first experiment, the peptide of interest is added, the cells are incubated until binding has been achieved, then free peptide is washed away and a fluorescent label added. As the peptide unbinds, fluorescence increases, which can be used to measure the peptide off-rate. In the case of the B-cells, Brefeldin A (BFA) is added to prevent the creation of new MHC-I complexes and the abundance of MHC-I is measured over time with anti-HLA-A2 antibodies. The experimental methodology, *written by our collaborators*, follows in the rest of Section 2.1.2.

2.1.2.0 Fluorescence of immobilised MHC-I

Production of MHC-I proteins: A pHN1+ plasmid encoding the mature human beta 2-microglobulin protein (β2m hereafter) was obtained from Prof P Moss. A pET22b plasmid encoding HLA A*02:01fos was obtained as described in [148]. Peptide-loaded MHC-I complexes were obtained as in [149] by refolding solubilised inclusion bodies of MHC-I heavy chains with solubilised inclusion bodies of human β2m and UV-labile MHC class I specific peptide.

Peptides: The following HLA-A*02:01 binding peptides were used: the UV-labile peptide KILGFVFjV (j represents 3-amino-3-(2-nitro) phenyl-propionic acid), the fluorescent peptide FLPSDC*FPSV (C* denotes TAMRA labelled cysteine), NLV (NLVPMVATV), NAV (NAVPMVATV), ADF-1 (SVASTITGV), ADF-2 (VMAGDIYSV), APO-1 (ALADGVQKV), CCN-1 (LLGATCMFV), GUC-1 (SVFAGVVGV), K67-1 (ALFDGDPHL), MET-1 (YVDPVITSI), MUC-1 (STAPPVHNV), HBV-1 (FLPSDFFPSV) and RGS-1 (LAALPHSCL) were synthesised by GL Biochem.

Fluorescence polarisation experiments: Fluorescent polarisation measurements were taken using an I3x (Molecular Devices) with rhodamine detection cartridge. Binding of TAMRA-labelled peptides is reported in millipolarisation units (mP) and is obtained from the equation $1000\frac{S-GP}{S+GP}$, where S and P are background-subtracted fluorescence count rates (S is the polarisation emission filter is parallel to the excitation filter; P is the polarisation emission filter is perpendicular to the excitation filter; and G (grating) is an instrument and assay-dependent factor). All experiments were conducted at 25°C in duplicate and used PBS supplemented with 0.5 mg/ml bovine gamma-globulin (Sigma), in a volume of 60µl. Peptide-receptive HLA-A*02:01fos complexes were obtained by mixing 75 nM HLA-A*0201fos loaded with UV labile peptides with 1.5 µM human β2m (Fitzgerald) and exposing to ≈360nm light for 20 min at 4°C ("UV exposed"). UV exposed MHC class I molecules were incubated with 75nM of unlabelled peptide overnight at 25°C. The next day 12nM FLPSDC*FPSV TAMRA labelled peptide was added and binding of FLPSDC*FPSV was measured for approx. 200 hours. Non-linear regression was performed, using the one phase association model in Prism.

2.1.2.1 B-cell MHC-I BFA decay

Cell lines: The T2 cell line [150] was maintained in RPMI (Sigma-Aldrich) with 10% fetal bovine serum (Globepharm), 2 mM L-Glutamine (Sigma-Aldrich) and 10mM HEPES (Lonza) at 37°C with 5% CO_2.

Peptides: ADF-1 (SVASTITGV), ADF-2 (VMAGDIYSV), APO-1 (ALADGVQKV), CCN-1 (LLGATCMFV), GUC-1 (SVFAGVVGV), K67-1 (ALFDGDPHL), MET-1 (YVDPVITSI), MUC-1 (STAPPVHNV), HBV-1 (FLPSDFFPSV) and RGS-1 (LAALPHSCL) were synthesized by GL Biochem and re-constituted in 100% DMSO to give 10mM stocks.

BFA Decay: T2 cells were incubated in 24 well plates at 3×10^5 cells/ml in AIM V Serum Free Medium (ThermoFisher) for 18hrs at 26°C with 5% CO_2 + peptide at 20µM or DMSO control. Peptide loaded cells were washed with PBS to remove excess peptide and re-suspended in 0.2 ml AIM V Serum Free Medium containing Brefeldin A (Sigma-Aldrich) at 5µg/ml. Incubated in 96 well plates at 37°C with 5% CO_2 for the specified time points. At the end of the time course each well was washed with PBS + 0.5% fetal bovine serum (Globepharm), stained with anti-HLA-A2 antibody BB7.2 [151] followed by goat anti-mouse conjugated with FITC (Sigma-Aldrich) and analysed by FACS (BD Caliber).

Analysis: Mean fluorescence intensity (MFI) measurements were determined using FCS Express software. HLA-A2 was expressed as the percentage of mean channel fluorescence at time point 0. Half-lives were determined from non-linear regression curve fits using GraphPad Prism.

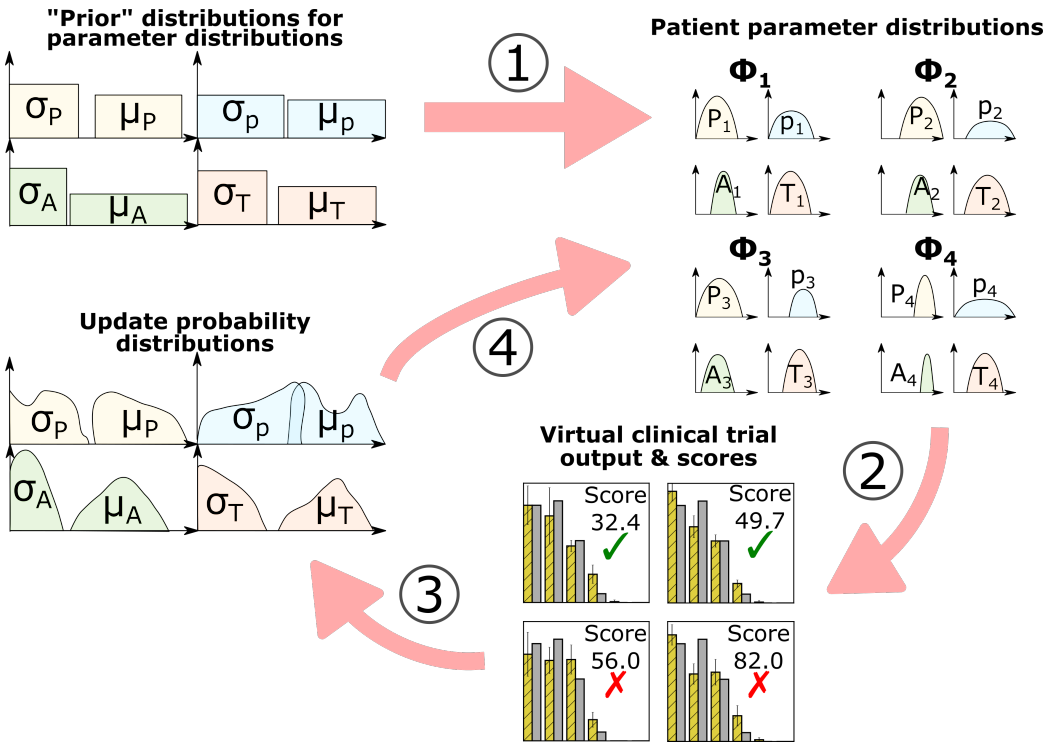


Figure 2.1. A schematic to explain how Approximate Bayesian Computation is used to find all patient parameters sets that give virtual clinical trial results consistent with IMA901. A detailed explanation, making use of the figure labels, is given in the text. **1)** Random patient parameter distribution means (μ_i) and standard deviations (σ_i) are initially uniformly drawn over a range. These yield many samples Φ_i, each containing a set of Gaussians for patient parameters. **2)** These parameter distributions are used to generate patient parameters, from which a virtual clinical trial is run. Each trial is assigned a score S_i based on its difference from IMA901; low scores are accepted and high scores rejected. **3)** The accepted parameter sets are used to change the probability distributions for the means and standard deviations of patient parameter distributions, and **4)** the process is repeated.

2.1.3 Finding all parameter distributions that match results via Approximate Bayesian Computation

Approximate Bayesian Computation may be used to find the range of patient parameters that give virtual clinical trial results consistent with IMA901. The procedure to find these values is presented graphically in Figure 2.1 and proceeds as follows, where list numbers correspond to the figure:

1. Patient parameters (P, the lymph transit time, p the spread in transit times, A, the proportion of dendritic cell MHC-I available to vaccine peptides in the dermis and T, the number of cognate antigen-specific T-cells in the lymph node, are shown) are drawn from Gaussian distributions. These distributions are in turn described by their means μ_i and standard deviations σ_i, which are initially unknown. The distributions for the means and standard deviations are hence initially flat, *i.e.* any value with a range is equally likely to be drawn. For each sample Φ_i, a random value of the

mean and standard deviation of each parameter distribution is drawn, such that each sample contains a Gaussian distribution for each patient parameter.

2. A virtual cohort of patients (*e.g.* 27 in phase I) is formed from each sample, each taking random parameter values (*e.g.* lymph transit time) from the parameter distributions. A virtual trial is run with each cohort, leading to a prediction of the number of patients that respond to 0, 1, 2 or 3 peptides. The difference between these virtual results and IMA901's results yields a score S_i:

$$S_i = \sum_{i=0}^{5+} (y_i - \tilde{y}_i)^2 \,,$$

where y_i is the number of patients predicted to respond to i peptides (or precisely, between $i - 0.5$ and $i + 0.5$ peptides) in the virtual trial and $\tilde{y}_i$ is the number in IMA901. An initial target score of $\frac{\sum_i \tilde{y}_i}{2}^2$ (=364 for phase I) is selected. Scores below the threshold are accepted. Those above are rejected.

3. The accepted samples Φ_i form the new probability distributions used to select parameter means and standard deviations. In other words, the next sample of means and standard deviations will be closer to the 'good' values in the previous sample. Then, the 75th percentile of the sample scores S_i is used to define a new target score.

4. This process is repeated 25 times with 10,000 (accepted) samples per iteration, iteratively updating the posterior distribution. More precisely, we are hence utilising ABC-*Population Monte Carlo* to find a final posterior distribution of patient parameter values that give virtual clinical trial results consistent with IMA901.

2.1.4 Hastening model evaluation

The model underlying the virtual clinical trials is an agent-based model and is thus relatively slow. Virtual clinical trials require, taking phase III as an example, 96 patients times 9 peptides times 3 trial repeats times 10 lymph node model repeats, which are each a simulation over 107 days with 3,081,600 timesteps of 0.05 minutes (not accounting for time 'skips' between vaccinations). Testing a large number of these virtual trials for Approximate Bayesian Computation (Section 2.1.3) or a sensitivity analysis (Section 2.2.11) would take a considerable amount of time. To remedy this problem, we attempted various techniques to speed up model evaluation: 'emulation' of the lymph node model (meaning replacement by a trained machine learning algorithm), an analytical solution and pre-calculation of the probability of T-cell–DC interaction as a function of time. The former two techniques

were not successful and are discussed in Appendix 2.A. The latter technique, however, is described here.

To calculate the cumulative probability of activation at time t, $p_A(t)$, we require both the cumulative probability of an interaction between T-cells and DCs $p_I(t)$ and the probability of T-cell activation given this interaction, $p_{A|I}(t)$, such that $p_A(t) = \int_0^t p_{A|I}(t)\frac{\mathrm{d}}{\mathrm{d}t}p_I(t)\mathrm{d}t$. The latter expression is simply the binomial formula already given, Equation 1.2:

$$p_{A|I}(t) = 1 - \sum_{x=0}^{n-1} \frac{N!}{x!(N-x)!} A(t)^x \left(1 - A\left(t\right)\right)^{N-x},$$

where n is the T-cell activation threshold, N is the number of MHC-I in the T-cell–dendritic cell contact area, and A is the ratio of cognate to total pMHC-I receptors on the dendritic cell's surface. Only a minimal amount of storage space is required to pre-calculate the cumulative probability of interaction p_I between a single T-cell and a number of DCs between 1 and 3000: 300 equally spaced double floating point values of the mean and standard deviation (over 300 repeats) up to 48 hours requires $3000 \times 600 \times 64/8\mathrm{B} \approx 14\mathrm{MB}$. This means that both $p_I(t)$ and $p_{A|I}(t)$ are known, and they can be integrated to calculate $p_A(t)$. This is achieved by finding a time t_{before} just above the transition seen in Figure 1.1 (page 41) such that $p_{A|I} = 1 - \epsilon$, and a time t_{after} just below the transition such that $p_{A|I} = \epsilon$. We assume that the cumulative probability of activation at time $0 \leq t < t_{\text{before}}$ is just the cumulative probability of interaction in that time, $p_1 = p_I(t_{\text{before}})$, because above the transition, all interactions lead to successful activation. The cumulative probability of activation for times after that is $p(t) = p_1(t_{\text{before}}) + p_2(t)$, where $p_2(t) = \int_{t_{\text{before}}}^t \mathrm{d}t' p_{A|I}(t')\frac{\mathrm{d}p_I(t')}{\mathrm{d}t}$.

This approach to calculating $p_A(t)$ requires the gradient of p_I at t. For this, we use a proxy model. There are two potential models for this:

1. **Proxy model 1:** Set $A_{\text{dermis}} = 1.0$ and $k = 0.0$, such that every interaction results in successful activation.

2. **Proxy model 2:** Set $A_{\text{dermis}} = 0.0$ and remove the condition that the simulation ends when DCs cannot activate T-cells, such that no interaction results in successful activation.

Proxy model 1 provides the cumulative probability of T-cell interaction over time. This leads to an accurate prediction of $p_2(t)$ when the initial value of $p_{A|I}$ is high, but otherwise it results in an underestimate of the probability of activation. The reason for this may be seen by considering a simulation of the proxy and full models with an identical random seed and identical parameters other than A and k. Suppose that the very first interaction is unsuccessful in the full model. Then, the current simulated probability of interaction

(activation) in the proxy model is $p_I(t) > 0$, but $p_A(t) = 0$ in the full model. Thus, the proxy model now has fewer simulated T-cells to interact compared to the full model, leading to an underestimate of $\frac{\mathrm{d}p_I}{\mathrm{d}t}$ for later times. This may be remedied by selecting values of $\frac{\mathrm{d}p_I}{\mathrm{d}t}$ as a function of p_I instead of time, but the method still yields a slight underestimate if the initial value of $p_{A|I}$ is low.

Proxy model 2 provides the expected total number of interactions a T-cell has in time t, and hence yields the cumulative probability of an interaction occurring $p_I(t) = 1 - e^{-(\text{interaction rate})t}$. This method is inaccurate if the initial value of $p_{A|I}$ is high, because once an activation has succeeded in the full model and the simulated T-cell is removed, the rate of interactions should fall. However, this regime of accuracy complements proxy model 1. We implemented an ad-hoc hybrid solution that uses proxy model 2 if $p_{A|I}(t = 0) \leq 0.1$ and proxy model 1 otherwise. A comparison of the two methods and the hybrid method with the full model is shown in Figure 2.9a.

This calculation must be further refined to account for various features of the model. The most trivial is that we are interested not in the probability of activation per T-cell, but the probability of at least one activation, and this is $p_{\mathrm{A}\geq 1} = 1-(1-p_A)^T$ for a number of T-cells T. The second modification is that DCs arrive over a finite arrival period p. It transpires that assuming that effectively $D/2$ DCs are present during the arrival period, where D is the final number of DCs, yields a good estimate for p_A in practice. This can be motivated by considering the analytical solution for the probability of at least one interaction described in Section 2.A.0.0 under the assumption of diffusive transport, $P_{I\geq 1}(t) = 1-e^{-\alpha DTt}$, where α is a geometrical factor. This analytical solution does not fit the simulated model, but were it correct, then it could be modified for p as follows:

$$
\begin{aligned}
t_* =& \frac{1}{k}\ln\left(\frac{NA_{\mathrm{dermis}}}{n}e^{-kP}\right) \\
p_{I\geq 1} =& \begin{cases} 1 - e^{-\alpha T d(t)(t-\frac{d(t)-1}{2}\frac{p}{D-1})} & t < p < t_*, \\ 1 - e^{-\alpha TD(t-\frac{p}{2})} & p < t < t_*, \\ 1 - e^{-\alpha TD(t_*-\frac{p}{2})} & t_* < t, \end{cases}
\end{aligned}
\tag{2.0}
$$

where k is the pMHC-I off-rate, P is the time spent by dendritic cells in the lymphatics, $d(t) = 1 + \frac{t}{p}(D-1)$ is the (continuous) number of DCs present at time t and t_* is the time at which $A(t_*)N = n$, which is used as an estimate for t_{after}. The result for the probability of activation during or after the arrival period can be obtained through a product of terms. The arrival of the j^{th} DC occurs at time $p(j-1)/(D-1)$. Hence, the probability of *no*

interaction after the j^{th} arrival is,

$$1 - p(I \geq 1) = \prod_{i=1}^{j-1} \underbrace{\left(e^{-\alpha T i \frac{p}{D-1}}\right)}_{\substack{\text{Contribution} \\ \text{by } i \text{ DCs in} \\ \text{time interval} \\ \frac{p}{D-1}}} \underbrace{e^{-\alpha T j (t - \frac{(j-1)p}{D-1})}}_{\substack{\text{Contribution} \\ \text{by } j \text{ DCs in} \\ \text{time since} \\ (j-1)^{\text{th}} \text{ arrival}}} = e^{-\alpha T j (t - \frac{j-1}{2} \frac{p}{D-1})},$$

in which we exchange j for $d(t)$ as a continuous approximation. The expression for the probability of interaction after time t_* (bottom line of equation 2.0) can be seen by inspection to be equivalent to a reduced number of DCs $D/2$ being present during the arrival period p. Though the analytical solution is not a good approximation to the full model, the assumption that $D/2$ DCs are present during the arrival period does work for the simplified model that we use.

The final model simplification addresses the fact that multiple vaccinations are given in IMA901. Some of the resulting 'waves' of dendritic cells entering the lymph node are close enough in time to overlap (*i.e.* within 48 hours of each other – the lifetime of DCs in the lymph node [12–17]). This is handled correctly in the simulated model, but we assume that each wave is independent in the simplified model to ease calculation. The effective probability of activation at time t is hence $1 - \prod_{w=1}^{\text{num waves}} (1 - p_A^w(t))$, where the probability of activation for wave w, p_A^w, may be at its final value, an intermediate value or 0 (wave not yet occurred) at time t.

2.1.5 Fitting experimental dendritic cell migration data

The mean and standard deviation of the distributions of patient specific parameters, such as the time taken for dendritic cells to migrate to the lymph nodes, are allowed to vary and are used to fit virtual clinical trials to the results of IMA901. The one exception is the migration efficiency of dendritic cells, which has been measured and previously reported to be of order 1% [120]. We fitted an exponential distribution to these data (shown in Figure 2.2) and use this distribution for all virtual clinical trials, unless otherwise stated.

2.1.6 Sensitivity analysis

The sensitivity of the lymph node model was analysed in Section 1.2.6, but the sensitivity of the virtual clinical trial may be subtly different, as might be the simplified model. We performed two types of sensitivity analysis. The first was a global sensitivity analysis of the predicted probability of T-cell activation using the random forest technique, as described in Chapter 1, on two parameter sets. The first parameter set was smaller and evaluated with both the full and simplified model, to allow for comparison between them (this parameter

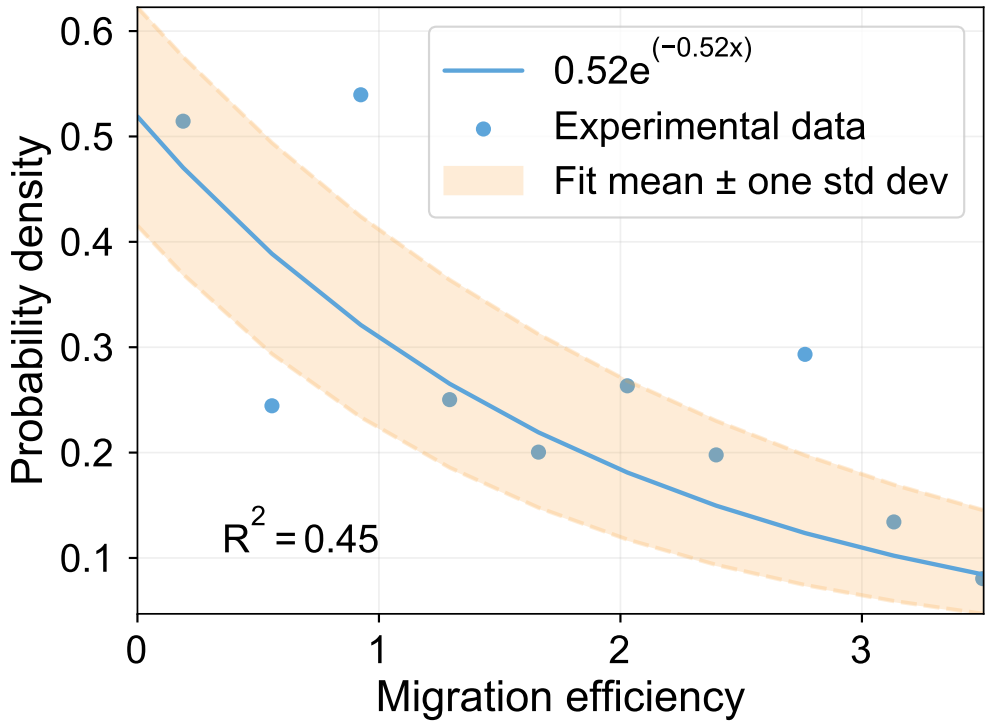


Figure 2.2. Probability density of dendritic cell migration efficiency (equivalent to a histogram). Points show reported experimental data [120] and the curve shows the best fit using a normalised exponential distribution. The shaded region shows the fits obtained with a decay constant that is a standard deviation above or below the best fit value.

set is also analysed with other techniques in Section 2.2.6). The second parameter set is much larger, but is evaluated with only the simplified model.

We also performed 'local' sensitivity analyses about each of the virtual clinical trial results (bar charts) presented in this chapter, wherein each parameter value was increased by 1% in turn, and the corresponding effect on the score (describing the fit of the data to IMA901, defined in Section 2.1.3) was calculated. This enables us to see which parameter most perturbs fits of the model to IMA901's results.

2.2 Results

2.2.0 Prediction of peptide response for one virtual patient

The lymph node model described in Chapter 1 yields the probability of a T-cell response against each peptide, given patient and peptide parameters. The model predictions for a single peptide and two random patients are shown in Figure 2.3, which may be compared to analogous figures shown in Chapter 1 such as Figure 1.6 (page 51). The left plot shows the probability of an interaction occurring between migrating dendritic cells (DCs) and T-cells

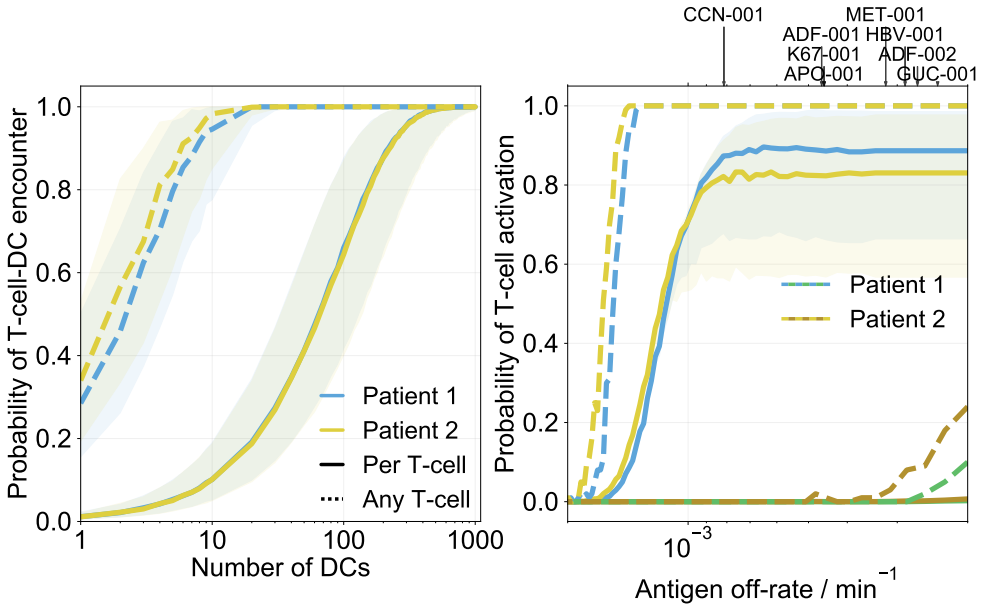


Figure 2.3. Left: The probability of simulated T-cells interacting with a variable number DCs after migration to the LN for two virtual patients, whether it results in activation or not. Solid lines indicate the probability for each (cognate) T-cell and dotted lines the probability of at least one (cognate) T-cell interacting. The shaded regions indicates probabilities between 6 and 24 hours, whilst the central lines indicate the probability after 12 hours.
Right: The probability of simulated T-cell activation in response to tumour-peptide carrying dendritic cells against pMHC-I off-rate for two virtual patients. The numbers of migrating dendritic cells for each patient have been randomly drawn from an experimentally determined distribution [120]. The definitions of the solid and dotted lines and shaded regions are as before. Blue and yellow lines represent the probability when only one peptide is injected into the dermis, and green and brown lines represent the probability for a single peptide after competition for pMHC with eight other injected peptides in the dermis. Arrows above the top axis indicate the off-rates of IMA901's constituent peptides.

in a finite period of time, as a function of the number of migrating DCs. Dotted lines show the probability that at least one T-cell interacts with a DC. Patient 2 has a higher precursor frequency (the number of simulated T-cells assumed to recognise the peptide), and so this probability is (slightly) higher for them. Solid lines show the probability per T-cell. No modelled patient-variable parameter except for the numbers of T-cells and DCs significantly affects the probability of T-cell interaction in a fixed period of time, so these probabilities are identical for each patient.

The right plot shows the probability that a T-cell is successfully activated in each patient in a fixed period of time versus the pMHC-I off-rate. As before, the dotted lines show the probability that at least one T-cell becomes activated, and solid lines the probability per T-cell. The upper four (yellow and blue) lines on the plot show a scenario in which only one peptide is injected into the dermis, and so the peptide can occupy as many available MHC-I as possible. As in Chapter 1, the transition-like behaviour in the probability of T-cell activation can be seen.

Parameter	Value
DCs migrated to lymph node	Distributed according to literature[120]
Number of antigen-specific T-cells	40 ± 11
Time until first DC leaves dermis	(6.00 ± 0.25) hours
Lymph transit time	(24.63 ± 1.76) hours
Spread of dermal departure times	(1.64 ± 0.33) hours
Fraction of MHC-I bound in dermis	0.2307 ± 0.0083
MHC-I in T-cell–DC contact region	500 ± 60

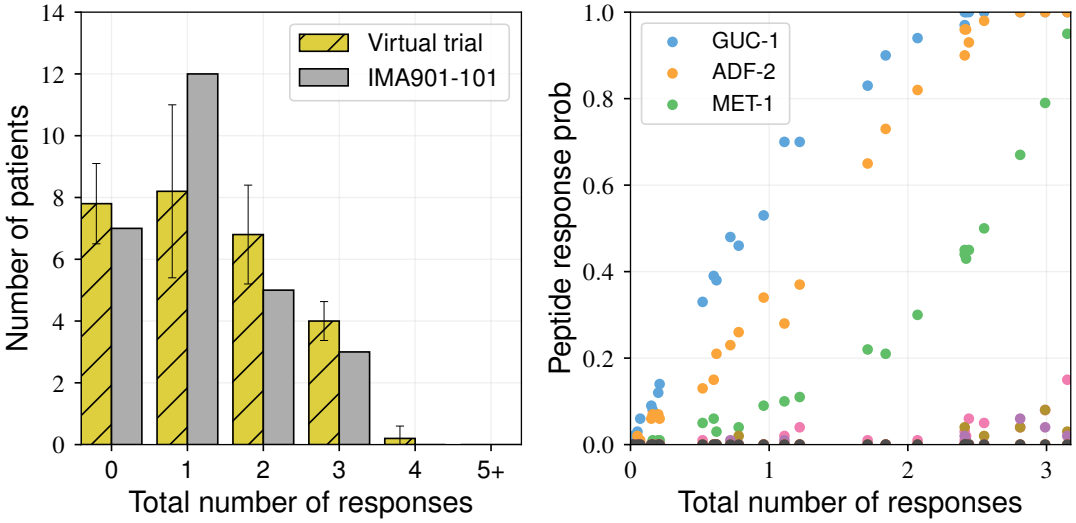


Figure 2.4. Results of matching a virtual clinical trial to the phase I trial of IMA901. **Table:** Parameters values for each patient are drawn from the distributions indicated, where the dendritic cell distribution is matched to literature data [120] (see Section 2.1.5). **Bar chart:** The number of patients who responded to 0, 1, 2 or 3 peptides, in the real trial and in the virtual trial. Error bars for the virtual trial are the standard deviation over three repeats. **Scatter plot:** The probability of responding to each peptide as a function of the total expected number of responses for each patient. For example, a patient expected to respond to 2.0 peptides has around a 90% chance of responding to GUC-001 and around an 80% chance of responding to ADF-002.

The lower (brown and green) curves on the right plot show the same simulations, but with the initial amount of MHC-I bound by the peptide divided by 9, to represent peptide competition. This corresponds to slightly over 3 half-lives, a large loss of initial peptide which results in a much smaller chance of peptide response. In a virtual clinical trial, the initial abundances of the nine peptides of IMA901 are calculated from an ODE system of peptide competition, resulting in unequal amounts of each peptide bound to DC MHC-I. These peptide- and patient-dependent probabilities of T-cell activation are used to calculate the results of virtual clinical trials, discussed in Section 2.2.1.

2.2.1 Simulated clinical trial results

The lymph node model was used to simulate a virtual clinical trial on a cohort of virtual patients, as described in Section 2.1.1. Representative results from the simulated clinical trials,

Parameter	Value
T-cells	90 ± 9
First DC exit	(6.00 ± 0.25) hours
Lymph transit	(29.41 ± 0.75) hours
Transit spread	(0.58 ± 0.50) hours
Dermal MHC-I	0.2079 ± 0.0012

Parameter	Value
T-cells	26 ± 9
First DC exit	(6.00 ± 0.25) hours
Lymph transit	(25.86 ± 1.57) hours
Transit spread	(2.31 ± 0.38) hours
Dermal MHC-I	0.2225 ± 0.0063

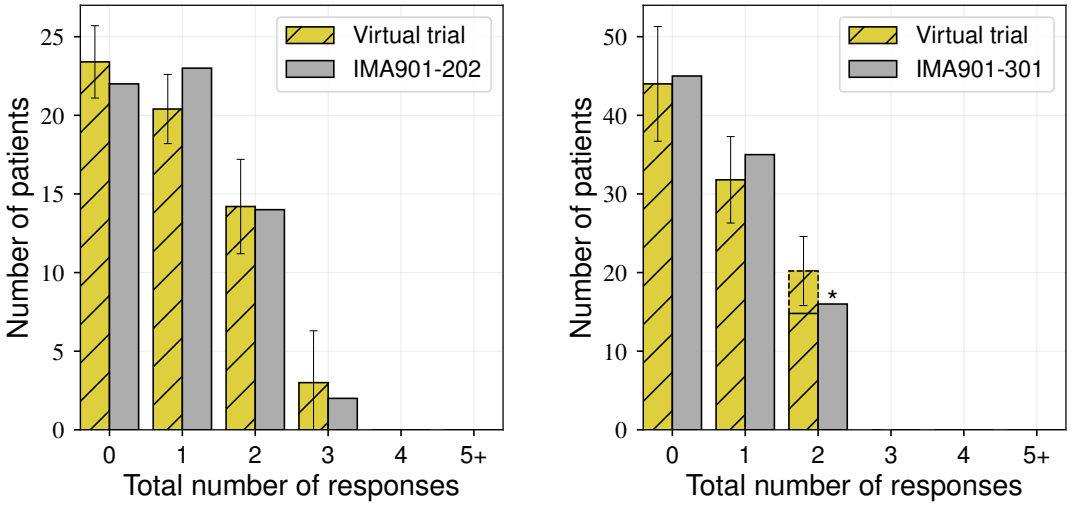


Figure 2.5. Results of the virtual clinical trials that match phases II (left) and III (right). In the tables indicating parameters, rows for the number of dendritic cells (DCs) and the amount of MHC-I in the T-cell–DC contact region have been omitted and other rows abbreviated for clarity. Note that the parameter distributions are different to each other and to phase I, and (*) that the number of patients responding to 2 or 3 peptides are combined in the data for IMA901-301 (phase III). The dashed bar is the prediction of the number of patients who respond to 3 peptides in the virtual trial.

compared to the real ones, are shown in Figures 2.4 (phase I) and 2.5 (phases II and III). In each case, virtual results are visually consistent with their real-world counterparts. The error bars shown are obtained by repeating the virtual trial three times and re-randomising patient parameters each time, in order to estimate the variability of results for a given parameter set. For the phase I data, an additional plot is presented, showing which peptides patients are expected to respond to (not reported by the authors of IMA901). For all three phases, it is expected that patients respond to a mix of GUC-001, ADF-002 and MET-001, with response probability falling in that order. The parameter distributions used for the virtual clinical trials in phases I, II and III are different. This is because parameter distributions are chosen such that the virtual and real-world trial results match, and moreover multiple parameter distributions can fit the data. We performed a parameter sweep to find all parameter distributions that can fit the observed data and the difference between phases I-III, if any. These results are presented in Section 2.2.7. In all cases, patients are expected to respond best to those peptides with the lowest peptide–MHC-I off-rates, particularly at late times, as those peptides occupy the most MHC-I sites. Thus, accurate values of these

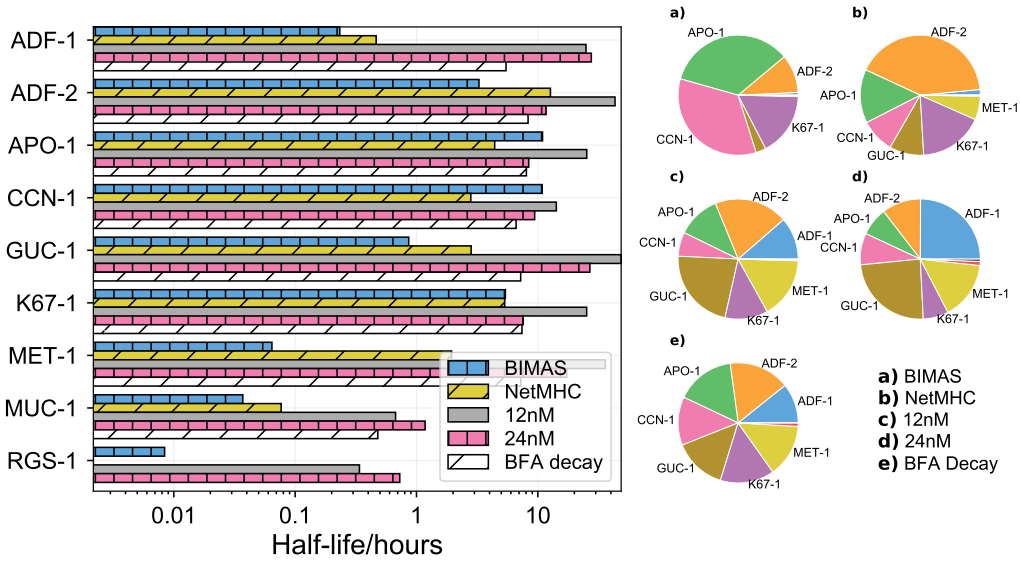


Figure 2.6. A comparison of the half-lives for each peptide, according to two algorithms ([152, 153]) and measurements by collaborators. Note that the rank order of each peptide differs for each technique, and that the average half-life measured by experiment differs greatly from the predicted values.

off-rates are critical to check for consistency with the results of IMA901. These values were obtained both from online utilities and from experimental measurements.

2.2.2 Measurement of the pMHC-I off-rates of the IMA901 peptides

The half-lives of the peptides of IMA901 were measured by our collaborators; the results are shown in Figure 2.6. They were measured with two techniques, and the first was repeated with two different concentrations of fluorescent marker. Measured values are very different from those predicted by BIMAS and NetMHC, both in order of magnitude (bar charts in Figure 2.6) and in the distribution of peptide half-lives (pie charts). However, experimental values also differ. The 24nM data, for example, predicts a more even distribution of half-lives than does the same experiment performed with 12nM of tracer. However, this may be because it has a much lower dynamic range (difference between the positive and background data) than the 12nM data, and so the 12nM data was selected rather than the 24nM data. The BFA decay suffers from a similar problem as the 24nM data; five peptides have very similar half-lives, and the remaining two – judged to be very poor binders by every other measure – have a significantly lower or undetectable half-life. As a result, the 12nM estimate for the half-lives of peptides are used in all presented results, unless otherwise stated.

	BIMAS		NetMHC	
Parameter	**Value**	**Modified value**	**Value**	**Modified value**
T-cells	30 ± 20		91 ± 22	
First DC exit	(6.00 ± 0.25) hours		(6.00 ± 0.25) hours	
Lymph transit	(19.52 ± 1.08) hours		(21.47 ± 3.62) hours	
Transit spread	(0.74 ± 0.12) hours		(0.37 ± 0.23) hours	
Dermal MHC-I	0.53 ± 0.06	0.80 ± 0.06	0.53 ± 0.004	0.90 ± 0.004

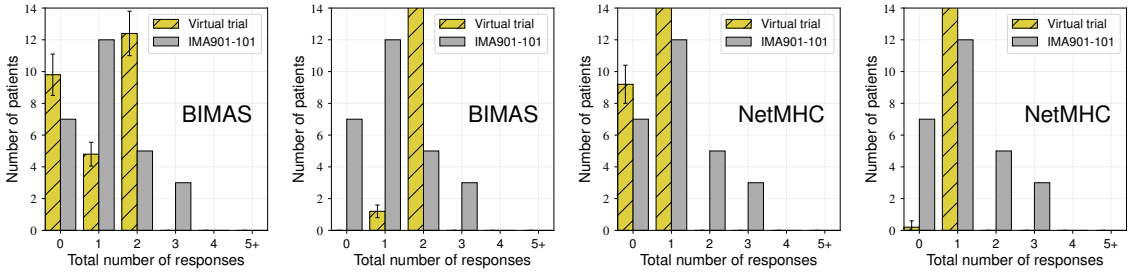


Figure 2.7. Results obtained by matching a virtual clinical trial to the phase I trial of IMA901, as in Figure 2.4, but values used for the pMHC-I off-rates for the peptides in IMA901 are obtained from online utilities for estimating pMHC-I off-rates, BIMAS [152] or NetMHC [153]. The left plot of each pair shows one of the best fits (by the sum of squared differences between bar heights) for that algorithm. The right plot shows the result of greatly increasing the proportion of MHC-I bound by peptide in the dermis, in order to try and gain probability content in the next column along without losing probability content in the bar for 0 responses.

2.2.3 Simulated clinical trial results using pMHC-I off-rates estimated from BIMAS and NetMHC

Initial attempts to investigate the results of IMA901 made use of online utilities to estimate the off-rates of the vaccine's constituent peptides from DC MHC-I. We made use of BIMAS and NetMHC. Virtual clinical trial results obtained using BIMAS and NetMHC's pMHC-I off-rates are shown in Figure 2.7. NetMHC produces affinities of peptides for MHC-I (K_D), so for comparison to BIMAS, we assumed an on-rate of 2.3×10^5/nM/min, such that the calculated off-rates, $k_{\text{off}} = k_{\text{on}} \times K_D$, had an average equal to the predictions from BIMAS. The two algorithms yield a different rank order of off-rates and neither are able to reproduce the results of IMA901. Regardless of parameter values it is not possible to simultaneously fit the number of patients responding to 0 and to 3+ (BIMAS) or 2+ (NetMHC) peptides, because the difference between predicted peptide off-rates is too large. To remedy this, we attempted to combine the predictions of many algorithms. We took a large ensemble of peptides and used several algorithms in turn to predict their binding affinities for MHC-I. The lowest predicted affinity for each algorithm was subtracted from all its predictions, such that the affinity of the worst peptide had a score of zero. Each set of predictions were then rescaled by the largest predicted affinity in each set, resulting in a list of scores from 0 to 100 from each algorithm for the peptides. The score of the peptides of IMA-901 could thus be compared between multiple algorithms, and their relative affinity for MHC-I calculated.

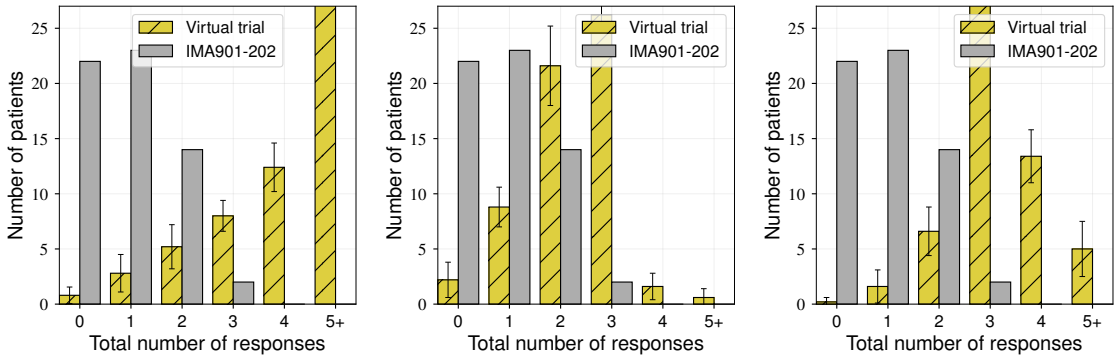


Figure 2.8. Virtual clinical trials of phase II using the parameters given in Figure 2.5, but with alterations as follows: **Left:** Lymph transit time is reduced from 30 to 10 hours, **Middle:** the 3 peptides with the fastest measured off-rates are removed and **Right:** the 4 peptides with the fastest measured off-rates are removed. The model predicts that faster lymph transit time (*e.g* by intramuscular injection) or reduced peptide competition would improve responses to a short peptide vaccine. Note that the grey bars indicate the results of the real trial, but approximately match those of the unmodified virtual trial.

However, even considering relative affinity this way, different algorithms rarely agree on peptide-MHC-I affinity and their predictions do not strongly correlate, raising questions about the reliability of the off-rates predicted by any algorithm. We thus measured the pMHC off-rates required for our model.

2.2.4 Alterations to clinical trial design

Virtual clinical trials allow for rapid testing and comparison of potential changes to vaccine design. Using the parameter distributions for phase II shown in Figure 2.5, various modifications were made to test different vaccine formulations. The left panel of Figure 2.8 shows the effect of reducing the lymph transit time from 30 to 10 hours, representing intramuscular injection. The middle and right panels shows the effect of using fewer peptides, by eliminating those with the fastest pMHC-I off-rates. Each case results in dramatically improved patient response distributions.

2.2.5 Difference between observed results in phases I, II and III of IMA901

IMA901 was successful in phases I and II, but failed to show clinical benefit over the standard of care, Sunitinib, in phase III. Furthermore, the distribution of patients responding to 0, 1, 2 or 3 peptides is different in each case. An interesting question is whether these differences are due to chance or to actual differences between the patient populations. However, although average values of details such as patient age were given by the original authors [52, 53], the complete dataset was not published and requests to access the data were not successful. Furthermore, no estimate of the error in the determination of patient response was given;

log(p value) with phase I probabilities

	0	1	2	3	Sum
Phase I	-0.3	-0.3	-0.3	-0.3	-1.2
Phase II	-1.5	-0.8	-0.7	-1.6	-4.6
Phase III	-5.9	-1.2	-1.2	-1.8	-10.1

log(p value) with phase II probabilities

	0	1	2	3	Sum
Phase I	-0.9	-0.6	-0.5	-2.0	-4.0
Phase II	-0.3	-0.3	-0.3	-0.3	-1.2
Phase III	-1.9	-0.4	-2.1	-0.5	-4.9

log(p value) with phase III probabilities

	0	1	2	3	Sum
Phase I	-1.8	-0.7	-0.8	-1.5	-4.8
Phase II	-1.3	-0.4	-2.2	-0.4	-4.3
Phase III	-0.3	-0.3	-0.3	-0.3	-1.2

Table 2.1. Base 10 logarithms of the p-values obtained by assuming that the number of patients responding to (for example) 0 peptides in phase I forms a Bernoulli random variable, occurring with probability $\frac{7}{27} = 0.26$. The difference between the expected and observed number of responses in phases I, II and III and hence p-values can be calculated. Each table shows the values obtained by using the probabilities from phase I, II and III in turn. Each row gives the p-values for the observed data, e.g. row 1 of the 2nd table is the p-value of the phase I results occurring using probabilities from the phase II results. All p-values are very low.

patient response was determined by manual appraisal and thresholding of flow cytometry for each patient/peptide combination, and only a part of these data was published.

An alternative method to determine if there are any differences between the results of each phase seen is to assume that each of the bars in one of the histograms are Bernoulli random variables. For example, if 7/27 patients responded to 0 peptides in phase I, then we assume that the probability of a patient responding to 0 peptides is $\frac{7}{27} = 0.26$, and the standard deviation of the number of patients expected to respond to 0 peptides in phase I is $\sqrt{0.26(1-0.26) \times 27} = 2.28$. We then calculate the expected number of patients that respond to 0 peptides in phases II and III (for example, 15.8 in phase II), calculate the difference between this and the observed result and thus calculate a z-score (for example, 1.81). This z-score allows us to calculate the corresponding p-value (for example, 3.5×10^{-2}). We perform this analysis for the data from phase I, phase II, phase III, phases (I+II), phases (I+II+III) and so on. The p-values obtained for the generation of the phase I, II and III results given probabilities derived from phase I, II or III alone are shown in Table 2.1, and all are much lower than should be expected. This may indicate that the approach is unsuitable, as the underlying process that generated the results is clearly not Bernoulli. We instead used Approximate Bayesian Computation to identify all parameter sets that match the results of each phase of IMA901, and compared the parameter sets to quantify the differences between the phases.

2.2.6 Simplified model accuracy

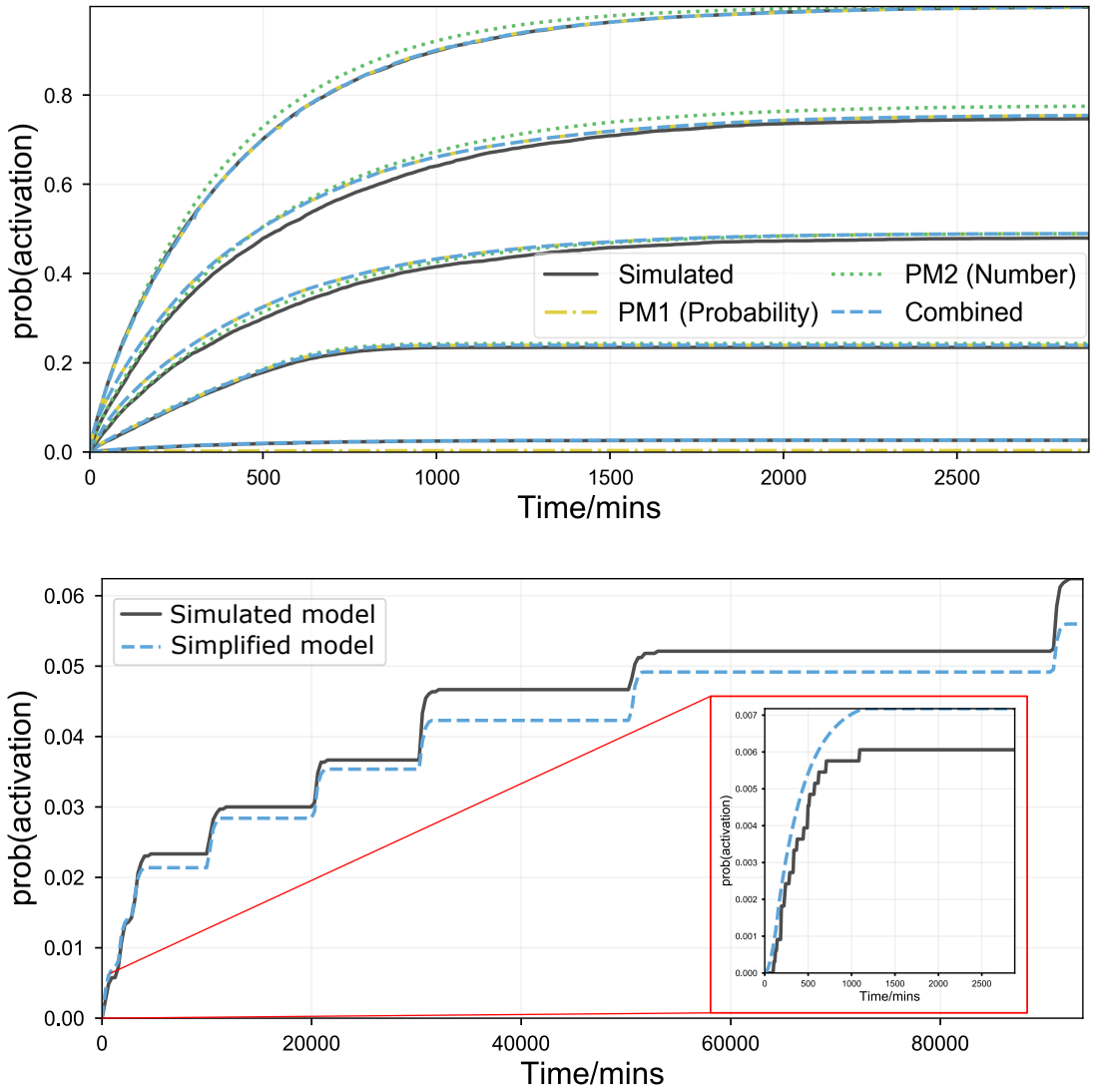


Figure 2.9. Plots comparing the simulated model to the simplified model described in Section 2.1.4. **Above:** Comparisons of different techniques for calculating the probability of interaction p_I at time t. Each set of curves is a different set of parameters. The bottom 4 curves have a finite off-rate and gradually increasing initial amounts of antigen. The top curve has zero off-rate. Solid black lines indicate the probability of activation predicted by the lymph node model. Dash-dotted yellow lines are the predictions obtained when proxy model 1 (PM1) is used to calculate p_I. Dotted green lines are the predictions obtained when PM2 is used for p_I. Dashed blue lines show the predictions from a hybrid algorithm that selects one of the other two predictions based on parameters. **Below:** The predictions by the simplified model (using the combination of PM1 and PM2) and the full model of the probability of activation for a single patient/peptide pair in a virtual trial of IMA901. The inset shows the probability after a single vaccination (wave of dendritic cells), whereas the main plot shows the probability after every vaccination of IMA901. Note that after one wave, the simplified model overestimates the probability of activation, but after many such waves, it underestimates it.

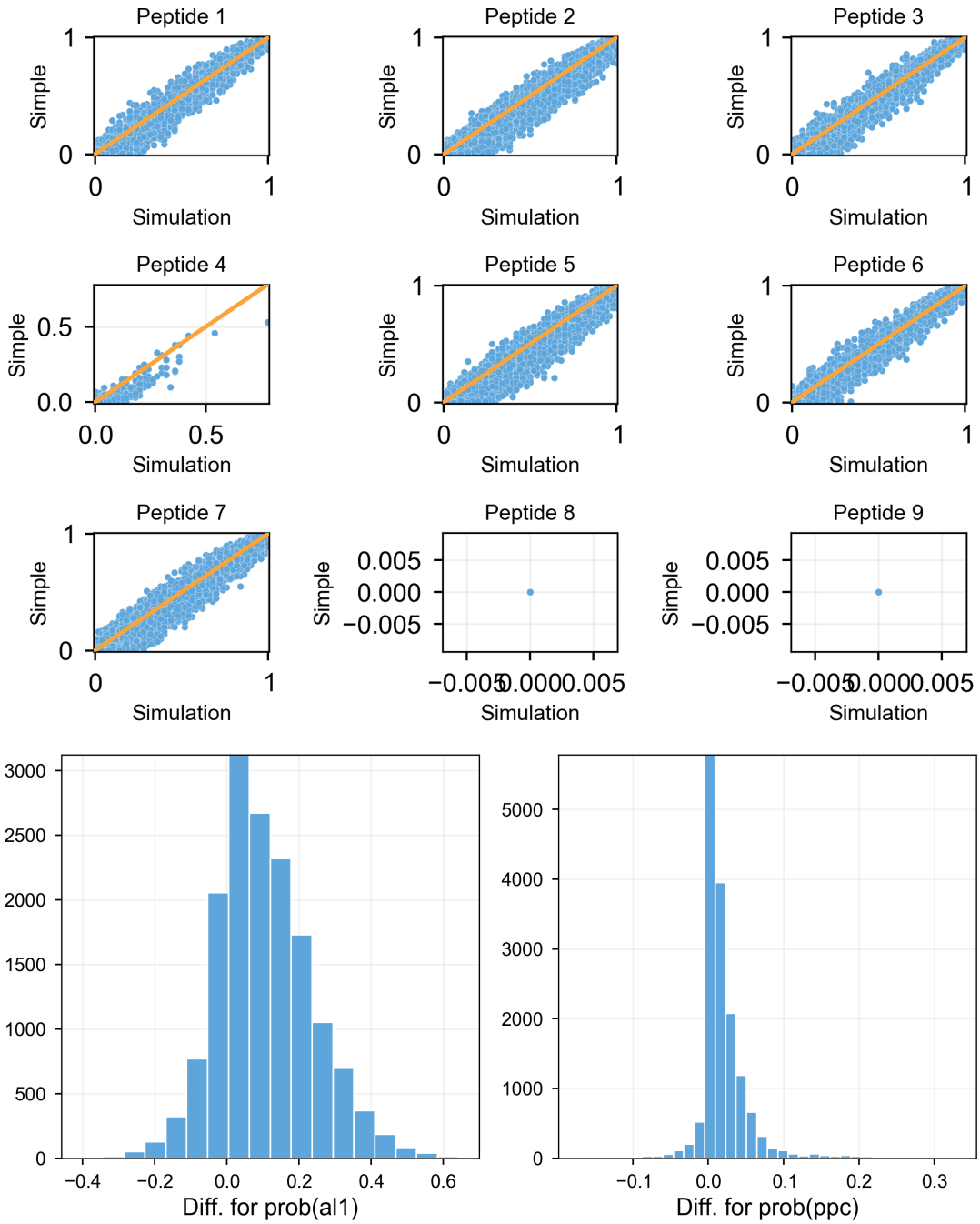


Figure 2.10. The difference between the lymph node model and an alternative, simplified expression used for analysis such as Approximate Bayesian Computation. The top plots show the probability of at least one T-cell activation for 15625 parameter sets predicted by the simplified expression, against the probabilities predicted by the simulated model. Each subplot shows the predictions for each peptide in turn. The bottom row of plots show histograms of the differences between these simplified and simulated predictions, for both the probability of at least one T-cell activation (al1) and the probability of activation per T-cell (ppc).

requires many evaluations of the model, which would take an infeasibly long amount of time. We thus created a simplified version of the model that could be evaluated much more quickly. Before proceeding with analysis using this simplified model, we had to confirm that its predictions matched those of the simulated model. Shown in Figure 2.9a is a set of comparisons of the predictions of the two models against time. The problems that arise when calculating the cumulative probability of interaction were described in Section 2.1.4. It can either be calculated from the cumulative probability of interaction in a zero-off-rate proxy simulation, or from the rate of interactions in a zero-antigen proxy simulation. The plot shows the predictions by both of these methods, in addition to a line indicating which of the two methods would be selected by our algorithm in each case. There are minor discrepancies, but the predictions seem to match well in these examples. Figure 2.9b shows how the two models' predictions of the probability of T-cell activation against time for a single patient/peptide pair from IMA901 compare. There are also small discrepancies between these predictions; the simplified model overestimates the probability after a single vaccination (wave of dendritic cells) in IMA901, but underestimates it after successive vaccinations.

To check whether this is a problem for just this single iteration or whether it occurs across all of parameter space, we generated a Latin Hypercube of patient parameters over relevant ranges, with 5 points per dimension (15625 total). We simulated the full model with 50 repeats and calculated the simplified expression at each point, then calculated the difference between the two models at each point. The results are shown in Figure 2.10. The two methods are strongly correlated, with coefficients of 0.88, 1, 0.89, 0.58, 1, 0.89 and 0.98 for the probability per T-cell for the first 7 peptides, and the last 2 peptides have 0 probability for all parameters. The differences in predicted probabilities of T-cell activation per cell reach 0.1, though 89% of the histogram's content is within ± 0.05. However, these differences accumulate for the probability of at least one T-cell activation. With numbers of T-cells ranging between 5 and 100, the largest differences in the probability of at least one activation are 0.6, with 73% of the histogram's content between ± 0.2. Since the two models' results are highly correlated and most differences are small, we proceeded with the simple expression to find patient parameter values consistent with the results of IMA901 using ABC, and then we performed a translation step in which we insert predicted parameter values into the full model and perturb parameters to a new local minimum.

2.2.7 Range of parameter values matching results of IMA901

Multiple sets of patient parameter distributions yield virtual clinical trial results that fit the results of IMA901, and hence it is not appropriate to present a single example of a fit

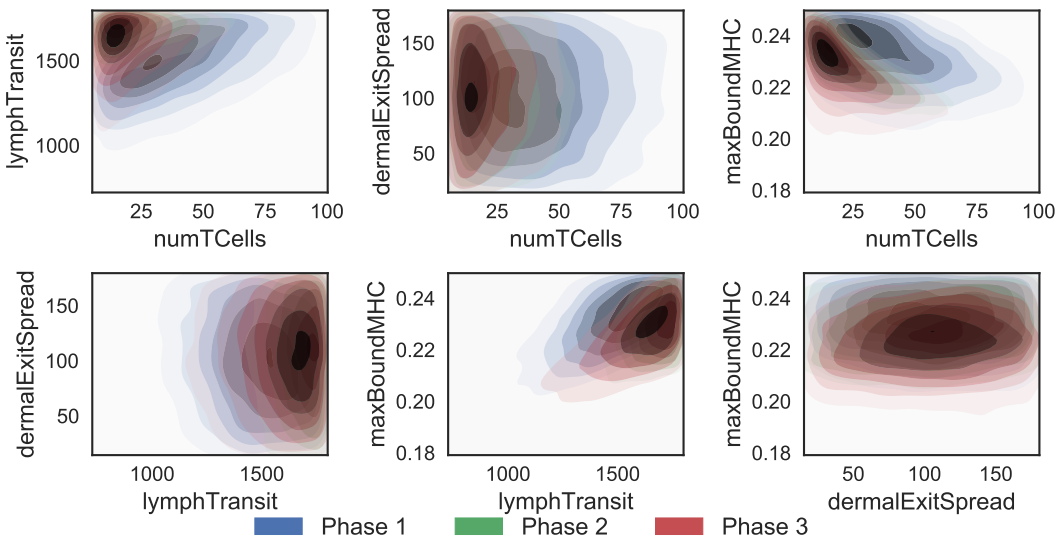


Figure 2.11. A density plot of the posterior distribution obtained through Approximate Bayesian Computation (see text). 10000 sets of random values for all parameters are taken from a 'basket' (prior distribution) of possible sets. Those that yield virtual clinical trial results closest to the real-world results are used to form a new basket. This process is repeated 25 times, yielding the posterior distribution shown. Areas of higher density indicate regions of parameter space that match the real-world data best. The density plots of phases I, II and III are shown together, allowing the difference between the phases to be visualised.

as in Section 2.2.1. In order to find the ranges of parameter distributions that can fit the data, to see how robust the model is and to quantify any difference in patient parameter distributions between the trials, we utilised Approximate Bayesian Computation (ABC) as described in Section 2.1.3. This procedure produces a density map where darker values correspond to parameter distributions that best fit the results of IMA901. Figure 2.11 shows projections of these densities onto 2D planes, to compare the values of pairs of parameters. It can be seen that a wide range of parameter distributions yield good fits to IMA901, and that phase I requires slightly different patient parameter distributions for a good fit.

2.2.8 Perturbation of local solutions to fit each phase

Whether there is a difference between optimal patient parameters for each phase of IMA901 can be seen more clearly by utilising ABC in a controlled fashion. Rather than allowing all parameters to change as in Section 2.2.7, a particular starting solution is chosen and then exactly two parameters are allowed to change at a time. The 2D density thus obtained across the two parameters is free of interference from the values of other, unplotted parameters. This analysis was performed for every possible pair of patient parameters. An example of the final 2D densities obtained for each phase is shown in Figure 2.12. As before, it is not phase III but phase I that splits off from the other two phases. This can be quantified further by integrating the densities. Integrating a single density plot should produce a value

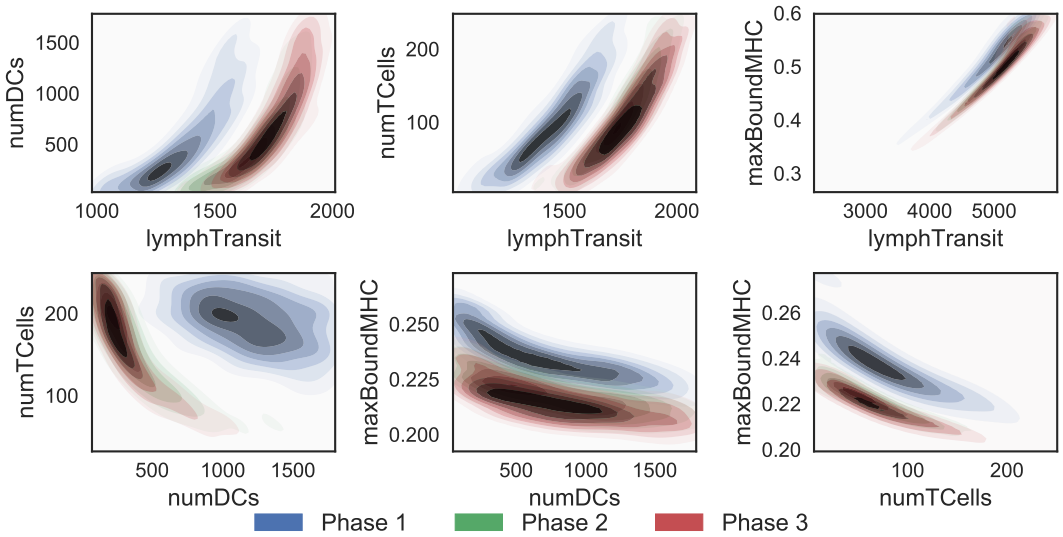


Figure 2.12. Density plot of the posterior distributions obtained through Approximate Bayesian Computation, where only two parameters are allowed to change at a time (indicated in each subpanel). This is unlike Figure 2.11, where every parameter is allowed to change at once and 2D projections of the resulting posterior distributions are plotted. As before, areas of higher density indicate regions of parameter space that match the real-world data best. Densities for phase I, II and III are plotted together in each panel. Note that the density of parameters that match phase I is distinct from those for phases II and III.

close to 1.0. Integrating the absolute differences between two density plots would produce 0 if they were identical, and a value close to 2 if they were completely different. The difference in integrals over density plots for the means of parameter distributions fitting phases I and II average 1.68. They average 1.73 between Phases I and III. They average 0.38 between phases II and III. In Figure 2.11, these values are 0.80, 0.69 and 0.39.

2.2.9 Approximate Bayesian Computation result accuracy

In order to determine that a sufficient number of (successful) samples per iteration were generated to find the correct density of parameters that match IMA901, we calculated the integral of the difference in density between two density plots generated by ABC with 10,000 samples. We then calculated the integral of the difference between two grids of random noise with the same size. If the densities were identical, this integral would be zero. If they were opposite, then the difference would be 1 (or -1). The average of the integral of the difference between the two density plots for each pair of parameters was 0.08, whereas the integral of the difference between the random data was 0.004. This indicates that more samples are required, but iterations with 50,000 samples take prohibitively long to compute (as the system must find 50,000 parameter sets better than the threshold every iteration, and the acceptance ratio had reached 0.008 by the 18th iteration): more than a week with the simplified system for a test run, versus half a day for 10,000 samples.

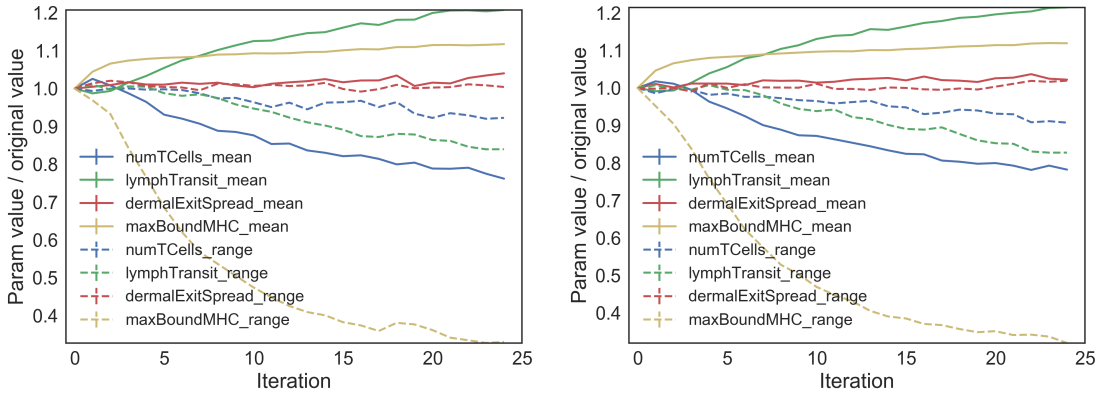


Figure 2.13. Convergence of the mean value of each parameter in the posterior distribution determined by Approximate Bayesian Computation against iteration number. Each panel shows a fit of the model to the phase I results of IMA-901, starting from the same prior distribution. 10,000 accepted samples were required per iteration. Solid and dotted lines are respectively (the mean value of) the mean and standard deviation of parameter distribution means. In both plots, the values seem to be converging to similar values; they differ by no more than 2%. With a very high number of samples, these final values should be expected to be the same.

Parameter	Value
DCs migrated to lymph node	Distributed according to literature[120]
Number of antigen-specific T-cells	30 ± 5
Time until first DC leaves dermis	6.00 ± 0.25 hours
Lymph transit time	30.00 ± 2.00 hours
Spread of dermal departure times	2.00 ± 0.38 hours
Fraction of MHC-I bound in dermis	0.2500 ± 0.0050
MHC-I in T-cell–DC contact region	500 ± 60

Table 2.2. Patient parameter distributions used to generate the synthetic data for Section 2.2.10, with both the simplified and full models.

However, when ABC was run with 10,000 samples with a uniform prior twice, the estimated mean for the eight parameters (that is, means of the mean and standard deviation of four patient parameter distributions) are no more than 2% different, with an average difference of 0.4%. In addition, the iterative change in these mean values can be confirmed on visual inspection to have 'levelled off' in both cases (see Figure 2.13), indicating that predictions are likely a good approximation to the required distribution. We thus proceeded with 10,000 samples per iteration.

2.2.10 Approximate Bayesian Computation result reproducibility

To increase our confidence that the patient parameter distributions produced by ABC are meaningful, we checked whether synthetic data produced by the simplified and simulated models (with the same parameter distribution parameters) can be recovered using the procedure, an idea utilised in the past in cardiac modelling (*e.g.* [154]). We ran a virtual

clinical trial of phase III of IMA-901, with patient parameter distributions as presented in Table 2.2. The data were produced with 250 repeats of the virtual trial and 50 repeats of the underlying lymph node model (in the case of the synthetic data produced by the full model).

We used the same priors as for the three phases of IMA-901 and used ABC to determine the parameter distributions that provide the best match to the synthetic data, the same number of virtual trials repeats per sample (three), the same number of successful samples per iteration (10000) and the same number of iterations (25). As before, we extracted the points of highest density and selected representative results that provide a good fit to the synthetic data by eye. The density plots produced by this procedure are plotted in Figure 2.14 together with a scatter plot of the true parameter values (shown in Table 2.2). Extracted representative results are shown in Figure 2.15. The parameter distributions for these representative results do not match those used to generate the data, but the true parameter values are within the region of high density that is predicted to match trial results. We can quantify the differences between the predicted and true values by comparing the predicted parameter distribution means and standard deviations to their true values. These data are presented in Table 2.3. Predictions from ABC with three trial repeats for each sample differ from the true values by 24% on average for parameter distribution means, and 50% for parameter distribution ranges. In case the difference was due to the small number of repeats of the virtual clinical trial per sample, we repeated the analysis with the simplified model with 100 trials per sample. The resulting density plots are compared to those obtained with the simplified model with 3 repeats in Figure 2.16. The predictions in this case differ from the true values by 10% on average for parameter distribution means and 31% for parameter distribution ranges.

2.2.11 Model sensitivity analysis

We performed both local and global analyses of the sensitivity of the model. The results of the local sensitivity analysis, the percentage difference in the score (fit of a virtual trial to IMA901) after a 1% change in each parameter, are shown in Figure 2.18. Some changes are no greater than would be expected of random variation of the unperturbed parameter sets (*i.e.* are within the Gaussians shown in the figure), but some parameters cause large changes in the score with relatively small changes in parameter value. The model is particularly sensitive to the proportion of MHC-I bound to peptide as dendritic cells leave the dermis and the number of pMHC-I complexes in the T-cell–DC contact area, for which a 1% change in parameter value causes the score to change by tens of percentage points. This may indicate that the model is not robust, and is possibly due to the assumption that short-peptides

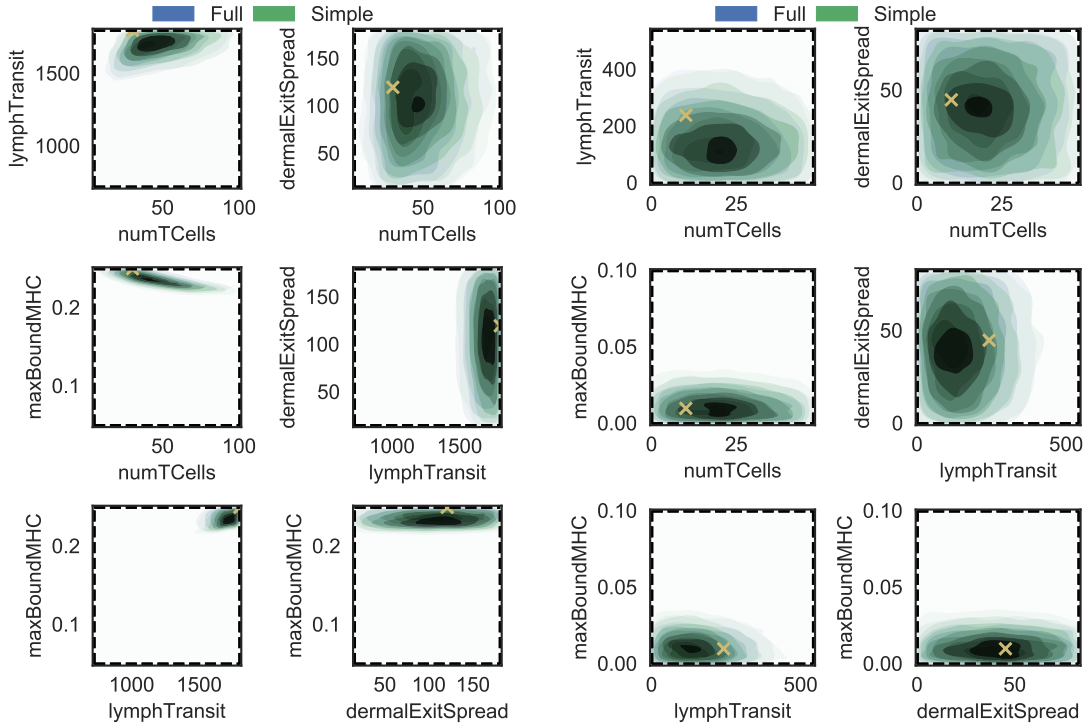


Figure 2.14. Density plots of the posterior distribution obtained through Approximate Bayesian Computation when matching to the synthetic datasets produced by the full and simplified models (*i.e.* the agent-based lymph node model and the proxy model described in Section 2.1.4). The left six plots show matches to mean values of the parameter distributions, and the right six to standard deviations of the distributions. As before, areas of higher density indicate regions of parameter space that match the synthetic data best. Both density plots have been plotted together, allowing the difference between the phases to be visualised. Yellow scatter points indicate the true parameter values used to produce the synthetic data. The dotted line bordering the plot indicates the boundaries of the uniform prior used to initialise ABC.

do not get loaded onto pMHC-I complexes in the endoplasmic reticulum (see discussion section).

The data produced to compare the simplified and simulated models in Section 2.1.4 were used to perform a global sensitivity analysis with the random forest technique. The results of this analysis are shown in the left part of Table 2.4. It can be seen that both the simulation and the simplified model have very similar sensitivities to different parameters, but that the difference between the two expressions is not more sensitive to any one parameter than another. The latter observation is contrary to what is seen when a simple correlation coefficient is calculated between each parameter and the difference between predictions (Figure 2.17), which indicate that the number of T-cells and the amount of antigen in the T-cell–dendritic cell contact area are particularly correlated with model discrepancy, and that the spread in departure times from the dermis has no effect on model discrepancy. To improve confidence in the global sensitivity analysis, and since the simulated and simplified

Parameter	Value
T-cells	46 ± 10
First DC exit	(6.00 ± 0.25) hours
Lymph transit	(27.65 ± 1.41) hours
Transit spread	(1.72 ± 0.34) hours
Dermal MHC-I	0.2359 ± 0.0063

Parameter	Value
T-cells	50 ± 11
First DC exit	(6.00 ± 0.25) hours
Lymph transit	(27.79 ± 1.43) hours
Transit spread	(1.73 ± 0.34) hours
Dermal MHC-I	0.2331 ± 0.0072

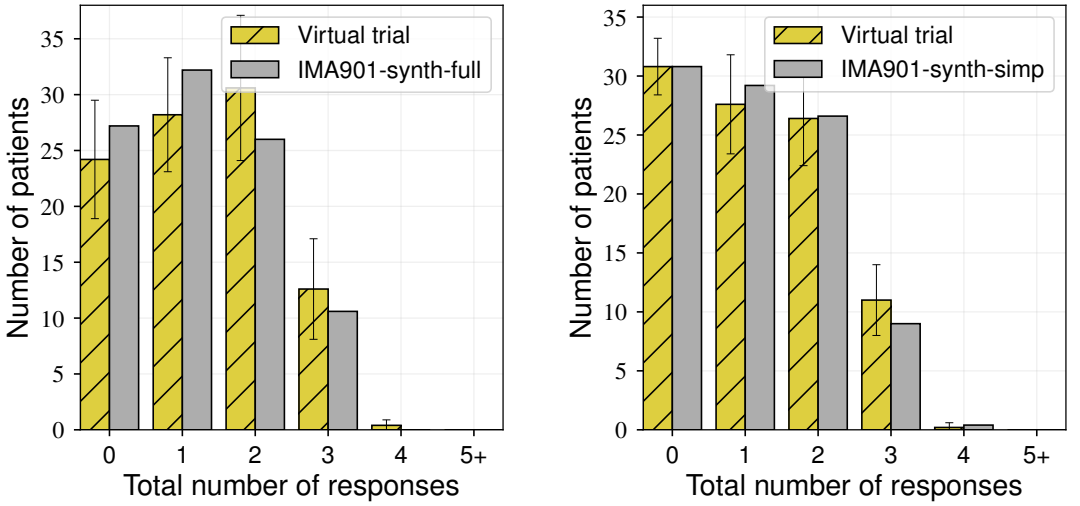


Figure 2.15. Representative results of the virtual clinical trials that match the synthetic phase III data produced by the full, simulated model (left) and the simplified model (right). In the tables indicating parameters, rows for the number of dendritic cells (DCs) and the amount of MHC-I in the T-cell–DC contact region have been omitted and other rows abbreviated for clarity. Note that the predicted parameter distributions are different to those used to generate the data, Table 2.2.

models are highly correlated, we repeated the analysis using only the simplified model but taking 1,000,000 parameter sets. These results are shown on the right of the table; sensitivity indices remain very similar to those obtained from the smaller dataset.

2.3 Discussion

2.3.0 Peptide–MHC-I off-rates

As a result of the 'transition' in activation probability predicted by the lymph node model, the results of the virtual clinical trial depend strongly on peptide-MHC off-rates. There are many algorithms available to estimate the affinity of peptides for MHC-I HLA-02. Section 2.2.3 presented the predictions of two such algorithms, neither of which could reproduce the results of IMA901. We attempted to find the relative affinity of the peptides for MHC-I by combining the predictions of multiple algorithms, but the algorithms we tried agreed neither on the relative ranking of each peptide nor the order of magnitude of the difference between them. Combining these estimates hence does not increase confidence in predicted off-rates. We chose to remove this source of uncertainty by measuring the off-rates, and

Patient parameter means							
		3 trials			100 trials		
	Actual	Pred.	% diff	# σ	Pred.	% diff	# σ
T	30	50 ± 20	67%	1.00	35 ± 10	18%	0.53
P	30	27.8 ± 1.6	-7%	-1.40	28.3 ± 1.1	-6%	-1.59
p	2	1.7 ± 0.7	-14%	-0.38	1.7 ± 0.7	-14%	-0.38
A_{d}	0.2500	0.2331 ± 0.0087	- 7%	-1.94	0.2400 ± 0.0054	-4%	-1.85
Patient parameter deviations (2 × Gaussian standard deviation)							
		3 trials			100 trials		
	Actual	Pred.	% diff	# σ	Pred.	% diff	# σ
T	10	22 ± 12	119%	0.97	16 ± 10	63%	0.62
P	24	2.9 ± 1.8	-28%	0.62	2.3 ± 1.4	-42%	-1.21
p	0.75	0.7 ± 0.4	-9%	0.20	0.7 ± 0.4	-8%	-0.17
A_{d}	0.01	0.0143 ± 0.0096	43%	0.45	0.0089 ± 0.0057	-11%	-0.19

Table 2.3. Predictions by ABC of the patient parameter distributions used to generate the synthetic data from Table 2.2, extracted from the density plots shown in Figures 2.14 and 2.16. The top table shows the predictions for the means of the Gaussians for each patient parameter distribution, and the bottom table the standard deviation of the Gaussians. Each table shows the predictions with 3 or 100 repeats of the virtual clinical trial for each sampled set of patient parameter distributions, and the difference between predictions and the true values, by percentage and by the number of standard deviations. Abbreviations: T, number of T-cells. P, lymph transit time. p, lymph transit spread. A_{d}, dermal MHC-I. Pred, predicted value. % diff, percentage difference between prediction and true value. # σ, number of standard deviations between prediction and true value.

found a rank order and average off-rate very different from the predicted values.

That the algorithms' predictions of the off-rate values were so wrong shows that due caution should be exercised when key results depend strongly on a value extrapolated from machine learning analyses. Secondly, that the model could only explain the data of IMA901 when the values of the off-rates were measured lends support to the hypothesis that short peptide vaccination success depends critically on pMHC-I off-rates. However, it should be stressed that it was the relative rank of peptides that was important; a peptide that is predicted to yield no responses in this vaccine may be predicted to yield a stronger response in a formulation in which it has the highest affinity.

Our collaborators measured the off-rates with two different experiments. In the first, empty human MHC-I complexes were mixed with peptides and left to bind overnight. The peptide was washed off and then either 12nM or 24nM of fluorescent index peptide was added, and the rising fluorescence was equated with the off-rate of the peptide of interest. The experiment that used 12nM of peptide had a better dynamic range (signal-to-background ratio) and was thus favoured over the other. Experiments were performed at 25°C rather than 37°C, which may impact results, though this is not expected to cause a large change in off-rates or the peptide ranking.

The other experiment measured the degradation of pMHC-I on B-cells. This allows

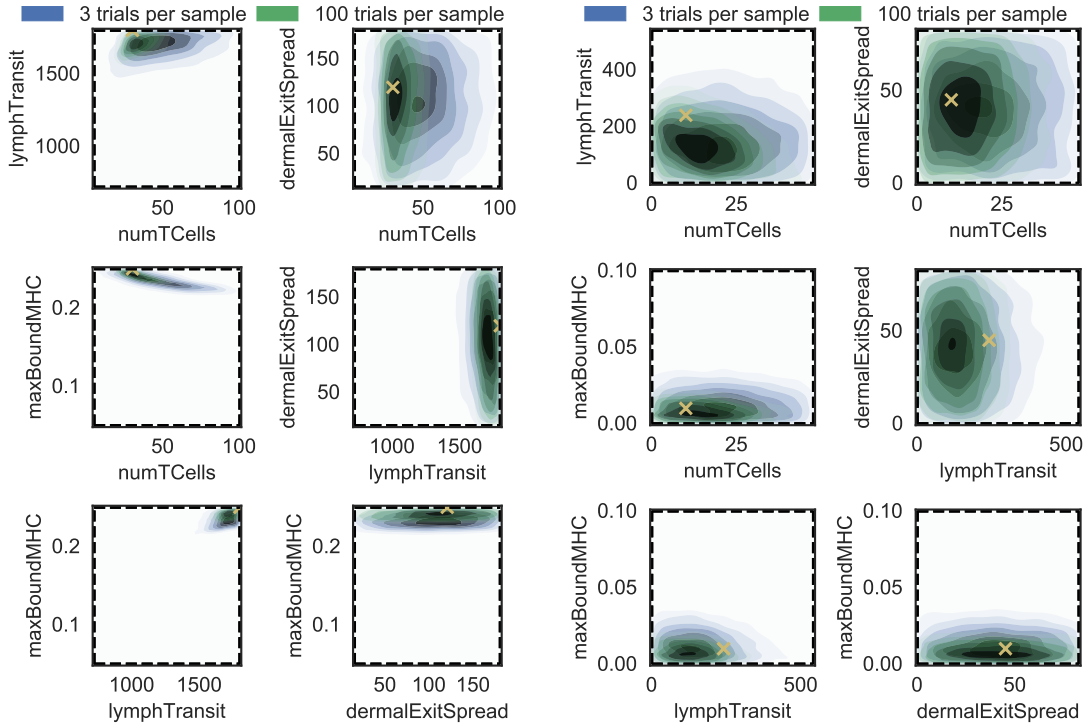


Figure 2.16. Density plots of the posterior distributions obtained through Approximate Bayesian Computation when matching to a synthetic dataset produced by the simplified model (described in Section 2.1.4), with either 3 or 100 repeats of the virtual clinical trial are made for each set of parameters sampled from the prior distributions. Other features of the plot are as in Figure 2.14.

quantification of other cellular processes that impact the lifetime of pMHC-I complexes on the cell surface, such as the recycling of the complexes in endosomes. However, the dynamic range of this experiment was also much less than that of the fluorescence experiment. This may be because the measurement was not of the rate of peptide unbinding, but rather the disappearance of the complex itself. Usually, such an experiment is used to determine whether pMHC-I binding happens at all, not to precisely measure off-rates. It is also unknown whether the rates of degradation measured on T2 B-cells are applicable to cells *in vivo*, and in particular to dendritic cells, whose function is to transport antigen long distances before presenting them, unlike B-cells. This uncertainty in the relation of the data to actual peptide off-rates, coupled with the low dynamic range, led us to favour the 12nM fluorescence experimental results for the off-rates.

We did not measure the peptide on-rates, as they are particularly difficult to measure. We have assumed that they are equal among vaccine peptides, and this has consequences for the modelling of peptide competition, as discussed in Chapter 1. As stated there, previous research has found that MHC-I affinity is determined mainly by the off-rate [142, 143], perhaps because the initial binding event is similar among peptides of equal size, or that it is diffusion-limited. However, the measurement of this quantity is key for refinement of

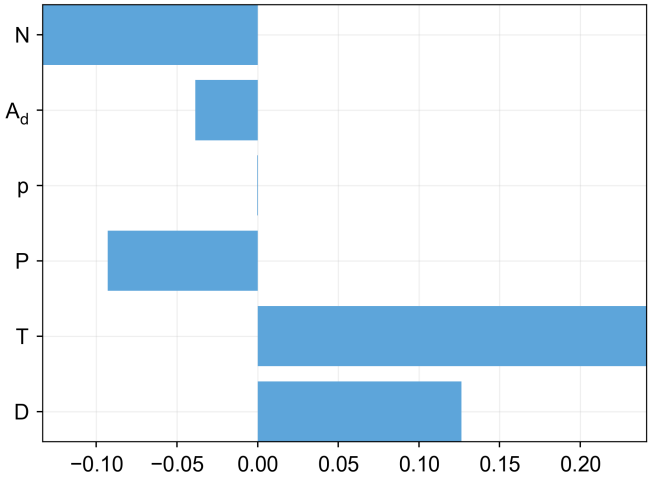


Figure 2.17 Pearson rank correlation coefficient for the difference between the simulated and simplified model predictions, against each model parameter in turn. Several parameters can be linked to the discrepancy between the models, which is contrary to what the sensitivity analysis suggests (Table 2.4). Symbols match Table 1.0 (page 46). A_{d} is A_{dermis}.

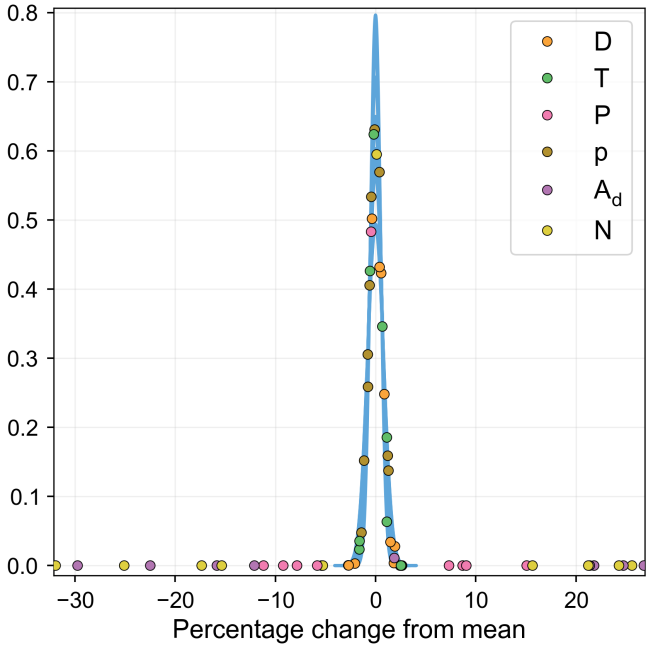


Figure 2.18 Local sensitivity analyses on 10 different solutions found by ABC for phase II (equivalent data exist for phases 1 and 3). The blue curves are Gaussians that indicate the variation of these solutions over 10,000 repeats, normalised such that they have a mean of 0. Each point, coloured by parameter, indicates how the mean score (measure of the difference between the virtual patient response distribution and IMA901's) changes after the parameter is changed by 1%. This is expressed in terms of the percentage difference from the unperturbed solution's original mean, to indicate how sensitive solutions are to small changes in parameters.

model predictions, particularly for the determination of the relative response probability for peptides with similar off-rates.

We can present competing hypotheses that could explain the results of IMA901. The first is that the dermal dendritic cells might not be adequately recruited or activated. However, prior work [104] suggests that the rate-limiting processes for T-cell activation following vaccination are diffusion, vascular clearance, binding of peptides to MHC-I receptors and the number of peptide-presenting DCs migrating to the draining LNs; in particular, the adjuvant diffuses for a greater distance and is cleared more slowly than the short peptide is. That any T-cell responses were found at all and that 'most' (14 of 27) patients responded to the marker peptide (HBV) further suggests that this is not the case.

The second is that short peptides might not lead to activation of $CD4^+$ T-cells, either

Sensitivity to: → Parameters	Small dataset Simulated	Small dataset Simple	Small dataset Diff.	Large dataset Simple
DCs migrated to lymph node	0.025	0.02	0.15	0.019
Number of antigen-specific T-cells	0.14	0.12	0.17	0.1
Lymph transit time	0.29	0.3	0.17	0.3
Spread of dermal departure times	0.019	0.021	0.16	0.021
Fraction of MHC-I bound in dermis	0.11	0.11	0.17	0.12
MHC-I in T-cell–DC contact region	0.41	0.42	0.19	0.44

Table 2.4. Sensitivity of the probability of T-cell activation predicted by the simulated model and the simplified expression to model parameters, and the sensitivity of the difference ('diff') between these predictions to parameters. The left table shows sensitivities calculated using 15625 parameter sets generated from a Latin Hypercube for both the full and simplified models. The right table represents 1,000,000 parameter sets evaluated with only the simplified model. The sensitivity of the full and simplified models are similar, and the difference between them does not appear to be particularly sensitive to any one parameter.

because the short peptide was cleared too quickly or because they are too short (9 amino acids long) to bind to helper cell MHC-II. The absence of IL-2 from $CD4^+$ cells would lead to reduced $CD8^+$ activation and proliferation. Although it is possible that $CD4^+$ T-cells are probabilistically activated with similar mechanisms to those studied in this work, the hypothesis that peptide off-rates are critical would still hold, albeit for MHC-II receptors. In IMA901, it was not the magnitude of response that was measured, but whether one had occurred at all, through the use of flow cytometry. This test would be positive with or without T helper cell support. We have also focused on whether a response occurred or not, rather than the level of the T-cell response, and hence our conclusions hold regardless of whether $CD4^+$ cells were also activated.

Finally, the observed results may have been due to varying T-cell affinities or precursor frequencies. We ignore varying TCR affinities for different antigen, by instead assuming that there is a precursor frequency of T-cells capable of recognising the antigen with approximately equal affinity. Around half (14 of 27) of the patients in Phase I responded to the HBV marker peptide, which could be reproduced in the model by assigning a higher precursor frequency of T-cells that recognise this peptide. Except for HBV, there is no reason to expect there to be a greater precursor frequency for any vaccine peptide over the others, but we cannot completely discount this possibility.

2.3.1 Parameter sensitivity and robustness

Parameter values (such as the number of recruited dendritic cells) for each patient are drawn from Gaussian distributions. Results of the virtual clinical trials are sensitive to the means of these parameter distributions, but solutions are not unique; for example, an increase in the mean lymph transit time can be offset by a reduction in the initial amount of MHC-

I available to peptides in the dermis. This is an expected result, as immune activation should be robust to different parameter values among individuals. We quantified this non-uniqueness by producing density plots of parameter values that match IMA901's results, in Section 2.2.7. For the same reason, we cannot simply predict the results of IMA901 and must instead match them by changing parameter values, which allows us only to determine if our model is consistent with results. We ensured that all parameter values took values in the range of literature estimates of their values, and we matched the migration efficiency of dermal dendritic cells to experimentally measured data [120]. This experimental data was fit to an exponential distribution and produced a reasonable R^2 value of 0.45, though several data points lie far outside of the 68.3% confidence interval. In the absence of measured data for other parameters, we chose to trust this data and fixed the parameter to reduce the number that had to be fit using Approximate Bayesian Computation (ABC).

In silico trial results are particularly sensitive to the number of pMHC-I in the T-cell–DC contact area and the maximum proportion of MHC-I bound in the dermis, as indicated by the narrow range of acceptable values in Figure 2.11 and as shown in Section 2.2.11. This sensitivity is contrary to expectations; even though a short peptide vaccination is biologically unusual, changes in the proportion of MHC-I sites bound of the order of 1% are reasonable to expect to occur in life, and thus it may not be realistic that it results in such large changes to predicted patient response probabilities. We postulate that this sensitivity is the consequence of the hypothesis that short peptides are not loaded onto MHC-I intracellularly. The high sensitivity and the fact that predicted mean values of the maximum proportion of MHC-I bound in the dermis are too high (roughly 0.25, whilst reported values indicate it should be around 0.10 to 0.15 [140, 141]) suggest that this assumption may not be correct. It could be relaxed by including a model of intracellular transport and TAP chaperoning, such as that by Dalchau and colleagues [155, 156]. However, the hypothesis generated by our work is not expected to be affected: pMHC-I distribution(s) would still depend on off-rate, and it can be seen that our results are consistent with those of IMA901 if and only if peptide off-rates follow the 'correct' distribution, supporting the hypothesis that off-rates are critical. Since the aim of this work was to propose a hypothetical mechanism that limits short peptide vaccination and not to predict the extent of response, such an extended model is not required.

Not all of the methods used to analyse model sensitivity were consistent. The global sensitivity analysis did not predict that the model is overwhelmingly more sensitive to the amount of antigen in the T-cell–dendritic cell contact area and the maximum proportion of bound MHC in the dermis than to other parameters, unlike the local sensitivity analysis. This would alleviate some of the concerns about this parameter, but the disagreement is

likely to be due to fact that the two sensitivity analyses consider different outputs. The global sensitivity analysis was performed on the predicted probability of T-cell activation across a Latin Hypercube of model-appropriate parameter sets, whereas the local sensitivity analysis was performed on the fit of a given virtual clinical trial to IMA901, which depends on particular off-rates being above or below the transition in activation probabilities. Furthermore, global sensitivity analysis of the difference between the models predicted that no one parameter was responsible, whereas Pearson rank correlation gives coefficients that agree with the results of the local sensitivity analysis.

2.3.2 Difference between the phases of IMA-901

Neither patient-level data (other than population means of certain characteristics) nor estimates of the uncertainty in the number of peptides that each patient responded to in IMA901 were reported. It was therefore not possible to calculate p-values for the difference between virtual or observed results for each phase, nor to calculate the probability that the differences in observed results were due to differences in the patient populations. We attempted to generate p-values to quantify the difference between the results of each phase by assuming that the probability of giving *e.g.* 0 responses rather than 1, 2 or 3 is a Bernoulli random variable, but the p-values were unrealistically small and hence not meaningful.

Without the ability to quantify the uncertainty in IMA901's data or to determine the difference between patient populations, we instead inferred these from the density plots produced by ABC, as shown in Sections 2.2.7 and 2.2.5. These data show that the patient populations for phases II and III seem to be roughly equivalent, but the population in phase I is predicted to be slightly different. This was always the case with a uniform prior, whether any parameter was allowed to change (Section 2.2.7) or all but two parameters were fixed (Section 2.2.5). The expectation was that the population in phase III would give the differing fit, as Sunitinib has been (controversially) reported to reduce antigen presenting cell migration [157–159] and thus may reduce peptide response probabilities. There are several possible reasons for phase I's required parameter values to be different. Firstly, it has far fewer repeat vaccinations than the other two phases; however, this differing number of visits is already explicitly simulated for each phase and so should not manifest in a requirement for different patient parameter values. Alternatively, the difference may be due to the presence of a marker peptide from Hepatitis B Virus in the vaccine in phase I. However, this difference would be expected to manifest as a requirement for more restrictive patient parameters to fit data, *e.g.* a smaller proportion of MHC-I bound as dendritic cells leave the dermis, and the predicted phase I parameter distributions are actually shifted in the opposite direction. Finally, it is possible that parameter distributions for phase I's population are different

because of the smaller sample size, which would mean that phase I's results are an outlier and would explain why virtual trial fits for phase I are worse than for phases II and III by eye. One additional method that could confirm this would be to pool data before fitting parameters to it (the most obvious previous example of which in prior work is the prediction of clinical trial efficacy [62]), and checking whether the single parameter set can produce a good fit to each phase in turn. This may, for example, produce good fits to phases II and III but not to phase I. Pooling data in this way should improve the ability of the model to predict matching patient parameter values due to reduced uncertainty, unless the trial populations *are* fundamentally different, in which case parameter estimation would be negatively impacted. We leave this analysis to future work.

2.3.3 Reliability of Approximate Bayesian Computation

ABC greatly simplifies the traversal of a parameter space, identification of regions of minimum potential and plotting of such regions. These are advantages over other techniques such as simulated annealing. In our experience, the latter tended to lead to the discovery of several local minima and would not efficiently traverse the entire parameter space. However, the points of maximum density produced by ABC cannot be easily interpreted. Likewise, it is not trivial to determine how many minima are in the system. If a multi-dimensional potential composed of the sum of four multi-dimensional Gaussians is explored with ABC to find the potential minima, then if the Euclidean distances between the locations of greatest density is calculated, the distinct minima are not resolved unless the four means are very far apart. For example, if the means are (1.0, 0.9, 1.1, 2.0) for one parameter, the two minima would not be resolved in a plot of the distance between locations of maximal density. If the last were 20.0, then it *could* be resolved. As the range of most of the parameters in the presented density plots do not vary by as much as an order of magnitude, it is unclear if there is a single or multiple minima for the patient parameter distribution parameters that best fit the results of IMA901. It is, however, possible to select parameter sets with the greatest weighting to generate 'representative' samples of fits, which led to the results presented in this chapter.

To increase confidence in the ability for this procedure to find patient parameter distributions, we tested whether it returns known patient parameter distributions used to generate synthetic data from a virtual clinical trial, as previous authors have done (*e.g.* [154]). These results are shown in Section 2.2.10. With the number of iterations, samples per iteration and virtual trial repeats per sample set to be the same as those used to fit the phases of IMA-901, the patient parameter distributions used to create the synthetic data are not quite reproduced by this procedure (Figure 2.14). Increasing the number of virtual

clinical trial repeats dramatically improves fits, however, such that the difference between predictions and true values are less than 20% for mean values (Table 2.3) and density plots visibly shift towards the true values (Figure 2.16). However, the small remaining differences – particularly for the standard deviation of patient parameter distributions, to which results are less sensitive – suggest that still more trial repeats would improve predictions and that the number of virtual clinical trial repeats used to fit the response data of IMA-901 ought to also be increased. It is notable that the lymph transit time and dermal MHC-I have smaller values than the true values, and the predicted number of T-cells has a higher value. We have already noted that multiple sets of parameters can produce the same virtual trial results and that a change in one parameter can be 'traded-off' by a change in another, but these changes would also favour the peptides with the smallest off-rates and reduce the stochasticity in model results. Repeating all results with more repeats will form part of immediate future work, though this is not expected to alter our conclusions.

A disadvantage of ABC is the number of computations required. We determined that more than 10,000 successful samples were required per iteration. With even 10,000 samples, the acceptance ratio reaches 0.05 by the 10th iteration and 0.005 by the 20th, finally requiring 5,000,000 samples for the final, 25th iteration. We produced a simplified model that is considerably faster than the simulated model, and yet even this required a long time to generate this many samples. We hence did not take more than 10,000 samples per iteration, even though analysis indicated that further accuracy could be achieved. Furthermore, we found that the predictions of the simplified model are slightly different from that of the simulated model and hence we had to perturb all sample solutions extracted from density plots, to find a local minimum for the simulated model. The origin of these differences has not been fully determined, but the most likely reasons are that,

1. some 'waves' of dendritic cells overlap (*i.e.* some vaccination times are within 48 hours of each other), causing the numbers of cells to be doubled (or more) temporarily, if the dendritic cells present antigen for long enough. However, the simplified expression assumes that waves are independent,

2. probabilities involved are small, meaning small changes due to stochasticity are relatively important.

It is possible that a density plot produced by the full model would differ from that produced by the simplified model, particularly because the differences between the models depend on parameter values (*e.g.* waves *are* independent for fast off-rates). However, the example parameter sets obtained from these density plots indeed fit the data with the full model after small perturbations of parameter values, the differences between the models are small

(Figure 2.10) and the models are highly correlated (Section 2.2.6). Additionally, the density plots produced from ABC – which uses the simplified model – to fit synthetic data from both the simplified and full models have considerable overlap (Figure 2.14; average difference in integrals over density plots for parameter distributions: 0.23 for Gaussian means, 0.12 for Gaussian standard deviations). Hence, the conclusions obtained from the density plots (that phase I is different to phases II and III and that there is a wide region of parameter space that can fit the results) should not be affected.

A relevant parameter estimation technique not considered here is mixed effects modelling, used extensively for pharmacokinetics and pharmacodynamics in industry. Various studies reviewed in Chapter 0 [62, 64, 77] use this technique. Mixed effects modelling proceeds by defining a model that links parameters and outcomes, and a statistical model of the variation of these parameters over a population (*e.g.* log-normally). Distribution parameters (*e.g.* their mean and standard deviations) are then varied to achieve the best possible fit to data. The strength of this technique lies in the inclusion of covariates (such as patient age) to determine the magnitude of their effect on outcomes, but these data were not reported for IMA-901. Mixed effect modelling still relies on a choice of suitable algorithms to choose new parameter values (by maximising a likelihood function), requiring many of the considerations made in this chapter (and thus possibly the use of ABC, regardless!).

2.3.4 Model consequences for IMA901

An advantage of virtual clinical trials is that they allow for rapid testing of alterations to the trial design. As shown in Section 2.2.4, the number of peptide responses given by patients is predicted to be greatly increased if the migration time of dendritic cells to lymph nodes is reduced (*e.g.* through intramuscular administration) or by reducing the number of administered peptides. However, using fewer peptides reduces the number of distinct T-cell populations that can be raised, and some patients may not have tumour antigen corresponding to, or may have fewer T-cells with high affinity for, the peptides still remaining in the reduced formulation. An optimal short-peptide vaccine may thus need to strike a balance between low peptide competition and a spread of possible immune responses, as indeed was the original aim of IMA901.

The framework developed within this study can be developed to predict responses or improve the design of other clinical studies of short peptide vaccinations. Several clinical trials are still on-going, and may be studied with this system to determine their results 'blind'. Some of these studies are testing new methodologies. For example, a company named IMV Inc has pioneered a technology in which antigen is delivered with oil, such that antigen ends up trapped within oil droplets. These droplets prevent the rapid clearance of

antigen by the vasculature and act as long-term antigen 'depots', and the oil irritates the skin to encourage activation of the immune system. This may provide a long-term source of antigen that improves the probability of T-cell activation, and a potential post-doctoral research project to study this system is under discussion. Another loose end identified by this work is Sunitinib. Though it is expected to impact the probability of patient response, this did not emerge in the distribution of parameters that fit the results. It may be insightful to model Sunitinib's effects explicitly to determine if they might be expected to interact with short peptide vaccinations. Finally, results indicate that the assumption that peptides taken up by dendritic cells are not processed for presentation on MHC-I may need to be relaxed, and an integrated model of TAP transport and pMHC-I processing may thus improve results.

2.4 Conclusions

The aim of this research was to hypothesise a mechanism that could explain the response distribution exhibited by the patients in IMA901. 'Virtual clinical trials' were performed using an *in silico* model of T-cell activation in the lymph node, with results that were consistent with those of the real trials. Simulated results depend strongly on the assumed values of peptide off-rates; values predicted by two MHC-I affinity algorithms do not lead to patient response distributions consistent with IMA901. That consistent results may only be obtained with accurate peptide off-rate estimates lends support to the model's proposed mechanisms and leads to a hypothesis that it is pMHC-I off-rates that are critical for short peptide vaccination success. Model results suggest that patient responses in IMA901 could have been improved by optimising these off-rates, optimising vaccine composition for more stable peptides, improving dendritic cell migration times, using fewer peptides with the same HLA subtype, or introducing peptides into different vaccine sites, so as to increase the initial number of sites bound per peptide.

The study hypothesis depends on a few key assumptions but nothing has been assumed about IMA901 in particular, so it is expected that the hypothesis holds for all short-peptide vaccinations. Whether pMHC-I off-rates and related variables are critical for short-peptide vaccination success or not may be tested by an experimental study. This highlights the power of virtual clinical trials and mechanistic modelling: it would have been difficult to carry out an experimental study to identify a mechanism that is consistent with IMA901's results. Computational studies that produce testable hypotheses or outputs analogous to experimental observables can spur experimental studies, and then in turn be utilised for further investigation as new evidence or questions arise.

The model also highlighted one of the drawbacks of pre-clinical studies: fundamental differences in physiology between pre-clinical species and humans can dramatically change the efficacy or outcome of an intervention. The model predicts that T-cell response probability depends strongly on the time spent by dendritic cells in the lymphatics in this scenario. This timescale is much faster in mice than humans, and the model would thus predict greater efficacy in a pre-clinical study. Had an in silico study been performed alongside its experimental counterpart from the beginning, however, their combined insights could have foreseen this problem and possibly resulted in an improved vaccine design. *In silico* modelling is well established within the life sciences and there have been studies of virtual clinical trials akin to this one (*e.g.* [62, 147]), but *in silico* experimentation and combined experimental-modelling studies are less common [160]. This approach holds great potential to accelerate discovery in human medicine in the coming years.

2.A Unsuccessful methods to hasten model evaluation

We used a simplified model that makes use of pre-tabulated probabilities of T-cell–DC interaction to rapidly evaluate the lymph node model for analyses such as Approximate Bayesian Computation. Before taking this solution, we attempted two other techniques that were not successful: 'emulation' of the model (replacement of the model with machine learning) and an analytical solution.

2.A.0 Emulation

Parameter	D	P	p	A_{dermis}	T	N
Min	100	960	45	0.1	10	400
Max	1000	1200	135	0.25	50	600

Table 2.A.0. Values used to define the Latin Hypercube from which random runs of the lymph node were performed, to train the machine learning algorithms within Spartan [161]. Parameter symbols are defined in Table 1.0 (page 46).

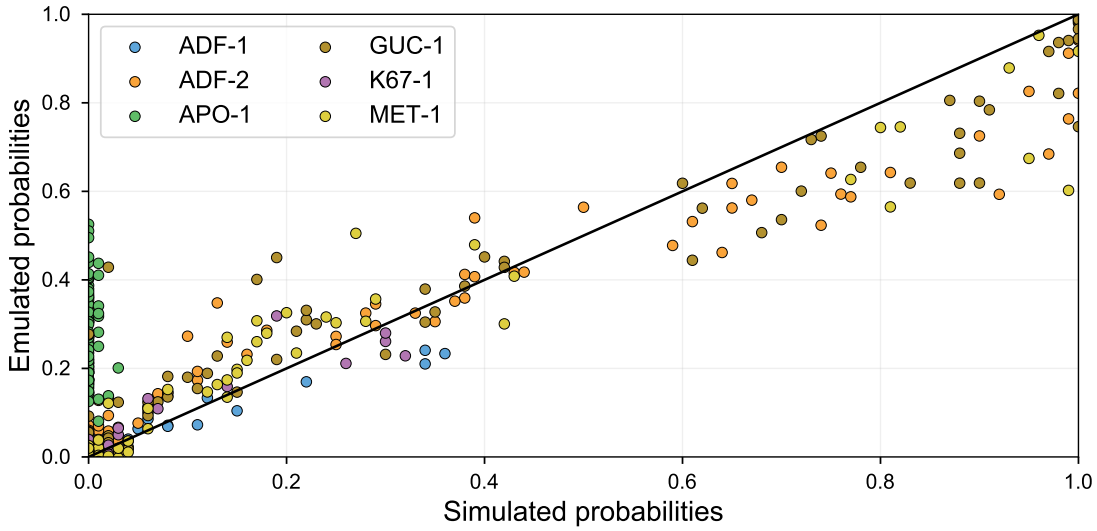


Figure 2.A.0. Performance of the emulated model against validation data, for all peptides that do not return only near-zero probabilities. The x-axis shows observed values (actual results from the lymph node model) and the y-axis shows predicted values (produced from the emulator). Much of the validation data appears to be close to the identity line for many peptides, but there are many departures and one peptide gives entirely incorrect results.

The first possible solution to evaluate the results of the model faster is to 'emulate' it using the R package named Spartan [161]. For this, a large number of runs of the model are completed with parameters drawn from a Latin Hypercube, and used to train an ensemble of machine learning algorithms (neural network, random forest, general linear model, support vector machine and Gaussian process models). Their performance is evaluated with

Parameter	Value
T-cells	30 ± 10
First DC exit	(6.00 ± 0.25) hours
Lymph transit	(21.28 ± 1.32) hours
Transit spread	(1.41 ± 0.47) hours
Dermal MHC-I	0.1950 ± 0.0083

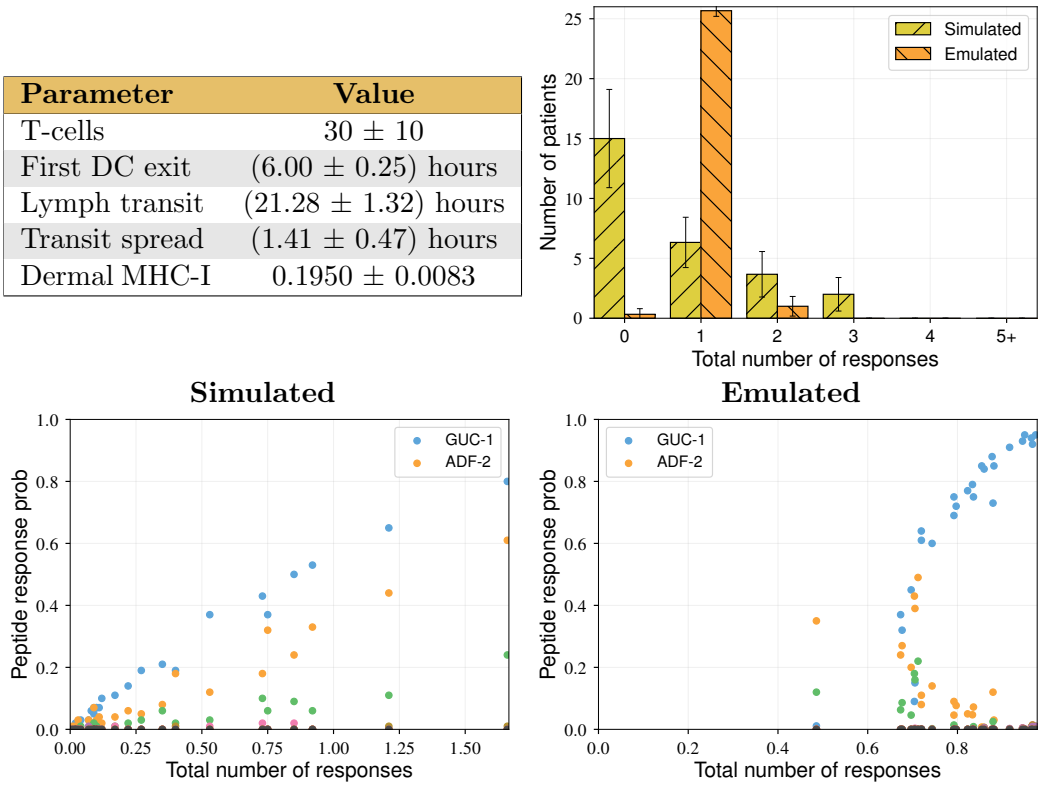


Figure 2.A.1. Performance of the emulator for the lymph node model in a virtual clinical trial. Both the lymph node model and the emulator were given the patient parameter distributions in the table and produced probabilities of response to each patient and each peptide. These values were used to calculate the number of patients responding to 0, 1, 2 or 3 peptides (bar charts), and also to plot the probability of response occurring against each peptide versus the total expected number of responses (scatter plots, below). The results do not match.

validation data, and this is used to set a weighting for each algorithm. Then, the weighted ensemble of these algorithms may be used in place of the original model to generate results. In a test of this methodology, we trained the machine algorithms with 500 or 1000 samples for six parameters between values shown in Table 2.A.0, where 500 is the value suggested by the authors for six parameters [161]. The resulting performance of the machine learning algorithms against validation data were evaluated, and then the emulator was used to run a virtual clinical trial and its results were likewise compared to the 'full' lymph node model.

The results with 1000 samples are shown in Figures 2.A.0 and 2.A.1. The former figure shows the the predictions by the trained ensemble of machine learning algorithms versus the probabilities produced by the lymph node model. There is in general good agreement but several data points depart from the identity line and APO-1 gives entirely incorrect results. In particular, the algorithm over-predicts at low probabilities and under-predicts at high probabilities for many peptides, potentially indicating problems with the transition. This imperfect fit leads to a wide departure from simulated results if the emulator is used

to run a virtual clinical trial, as shown in Figure 2.A.1. Here, the same parameters were used to run a virtual clinical trial with both the lymph node model and the emulator, but not only are the predicted patient responses very different (seen in the bar charts), but the predicted probability of response to each peptide as a function of the total number of responses exhibits anomalous behaviour; the probabilities of response to each peptide do not correlate as they should. It is not clear why the emulator failed in this way, but it could be because the regime of interest for the virtual clinical trial is when the system 'just' succeeds to activate T-cells, *i.e.* we are interested in extreme behaviour of the model. The emulator, however, captures 'average' or general behaviour of the model. A second problem is that large areas of the parameter space yield zero probability of T-cell activation, which would frustrate training of the machine learning algorithms, and effectively reduce the number of useful parameter samples taken. This latter point could possibly be remedied by more careful selection of the boundaries of the Latin Hypercube. However, we opted instead to try different techniques to speed up model evaluation.

2.A.0.0 Analytical solution

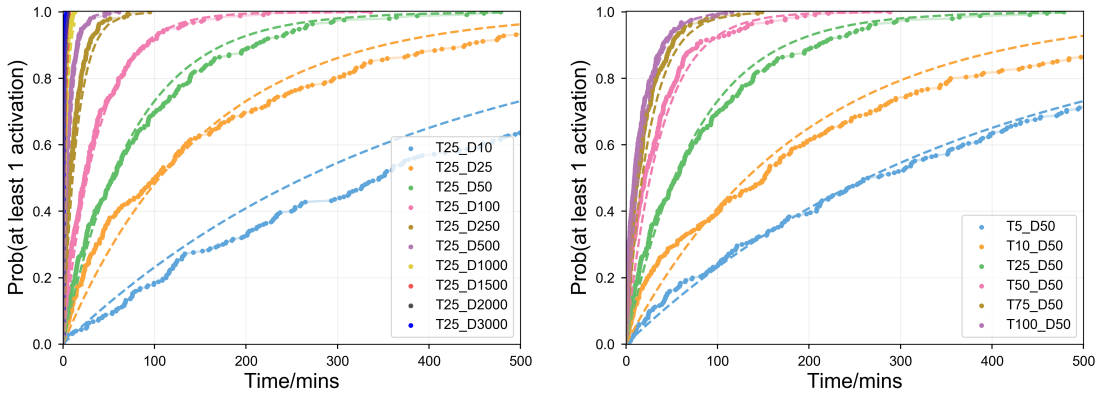


Figure 2.A.2. Comparison between the probability of at least one T-cell–DC interaction predicted by the lymph node model and an analytical approximation to it, for a variety of numbers of T-cells and DCs. The curves are split across two plots for clarity. Dotted lines indicate the analytical solution and scatter plots indicate the simulated results. Here, the mean free path of T-cell motion is 25µm, and hence the solution does not match. When it is zero, it matches (data not shown).

It would be advantageous to avoid having to run any computational model and instead use an analytical approximation for the probability of T-cell activation. As with the simplified expression, we seek both the cumulative probability of an interaction between T-cells and DCs, $p_I(t)$, and the probability of T-cell activation given this interaction, $p_{A|I}(t)$ (Equation 1.2). It is challenging to write down an expression for $p_I(t)$ when cells undergo 'run-and-tumble' with gamma-distributed velocities. If T-cells were assumed to undergo

brownian motion rather than a run-and-tumble, then an expression derived by Celli *et al* could be used:

$$\begin{aligned} p_{\text{at least one interaction}} &= 1 - e^{-\alpha DTt}, \\ \alpha &= \frac{R^3}{3Mb} - \frac{3}{5}R^2M, \end{aligned} \tag{2.A.0}$$

where M is the diffusion constant (or motility) and other symbols are as defined in Table 1.0. Other authors have derived an effective diffusion constant for run-and-tumble processes [162], or alternatively the effective diffusion constant may be 'measured' by simulating a very large lymph node with no DCs present (yielding $M = 20.3\mu\text{m}^2/\text{min}$ for parameters given in Table 1.0). Even if the latter is used, the analytical solution above does not match simulations unless the T-cell mean free path is set to zero (making them diffuse), as shown in Figure 2.A.2. As the accuracy of the fit depends on the number of cells required, and because the problem persists even with an extremely large LN radius at constant cellular density (excluding the possibility that restricted space is the problem), equation 2.A.0 cannot be correct. Attempts were made to fit a modified expression $\alpha = (\frac{R^3}{3Mb} - \frac{3}{5}R^2M)T^xD^y$ to the simulated data by varying x and y, but this was not successful. We did not pursue this problem any further; it seemed a better use of time to simply pre-calculate and tabulate the cumulative probability of interaction as a function of cell counts, as this requires only a few tens of megabytes. However, it would be an extremely interesting mathematical problem to develop an analytical expression for the probability of T-cell activation that is correct for all parameter values and relevant cellular dynamics.

Chapter 3

The circulatory system limits T-cell delivery to tumours: implications for comparative CAR T-cell dosing of mice and men

Note: much of the text in this chapter is reused from a planned publication of this work. Note also that the symbols defined in previous chapters are no longer relevant and should be considered undefined henceforth.

3.0 Introduction

Chapters 1 and 2 quantified steps 1–3 of the cancer-immunity cycle in the context of a short peptide vaccination. In this chapter, I will move onto steps 4 and 5: the trafficking of T-cells to and infiltration into tumours. One example of a bottleneck here could be poor tumour vasculature; 'leaky' vasculature has an insufficiently negative pressure gradient to force T-cells out into the tumour parenchyma, and some tumours are not even vascularised. Bioengineered T-cells such as CAR T-cells have seen success in the clinic and have been approved for both leukaemias and lymphomas (*e.g.* acute lymphoblastic leukaemia, non-Hodgkin lymphoma, diffuse large B-cell lymphomas and primary mediastinal large B-cell lymphoma [55]), in which they evoke strong anti-tumour responses [163]. However, clinical trials for solid tumours have been less successful, despite tumour regression observed in various mouse models [3, 164–166]. Notably, this is the case for all solid tumours, suggesting a more fundamental cause relating to solid tumour vasculature or biology. Studies have shown that transfused T-cells localise mainly in the liver, spleen and lymph nodes [167–169], a trait that CAR T-cells should share. This suggests that their limited efficacy could be explained in part by insufficient numbers of CAR T-cells localising to the organ bearing

solid tumours. Conversely, for blood-borne cancers, higher numbers of both CAR T-cells and their target cancerous lymphocytes localise in the blood, spleen or liver. Indeed, a relationship between dose (number of introduced cells) and response has been established in pre-clinical models for solid tumours [170] and in dose escalation studies [54, 171], but such a clear relationship has not been observed in the clinic.

Tumour clearance depends on the fine balance of tumour growth and (CAR) T-cell delivery or proliferation; if tumouricidal T-cells are not delivered rapidly enough then the tumour will continue to grow in size. Consideration of the fundamental limit of T-cell delivery due to the blood supply can thus provide insight into how efficacy may vary between organs, individuals and different species, by comparing expected delivery rates, typical dosages (cell counts) and the corresponding tumour responses observed in mice and humans.

In Chapter 0, I reviewed the use of physiologically-based pharmacokinetic modelling (PBPK) to predict drug concentrations in tissues and to study T-cell trafficking. We have not, however, seen an attempt to identify and compare maximum delivery rates in humans and pre-clinical species, nor to compare the numbers of bioengineered T-cells introduced to mice and men. The aim of this study is to compute upper bounds of clinical and pre-clinical delivery rates of CAR and other T-cells to tissues and tumours in the absence of proliferation, and to quantify how these bounds may vary. The results are used to compare the numbers of bioengineered T-cells given in mouse models and the clinic, and understand the potential efficacy of immunotherapeutics and the relevance of pre-clinical data to the clinic. To proceed, we have used published physiological reference data [172, 173] to develop dynamic models of the human, rat and mouse circulatory systems. Models were implemented as systems of ordinary differential equations and used to estimate upper physiological limits for immune cell extravasation rates per volume of tissue (cells/min/mm^3) into the major organs. We focus in particular on predicted delivery rates of an introduced population of CAR T-cells, to provide clinical context to predictions, but the results are applicable to any relevant population of lymphocytes or other immune cells. We compare physiological data from a variety of sources [172–177] to quantify how predictions may vary across individuals, experimental conditions and other factors, and finally we use published human tumour perfusion data (see Section 3.2.2) to quantify delivery rates per mm^3 of tissue for human tumours.

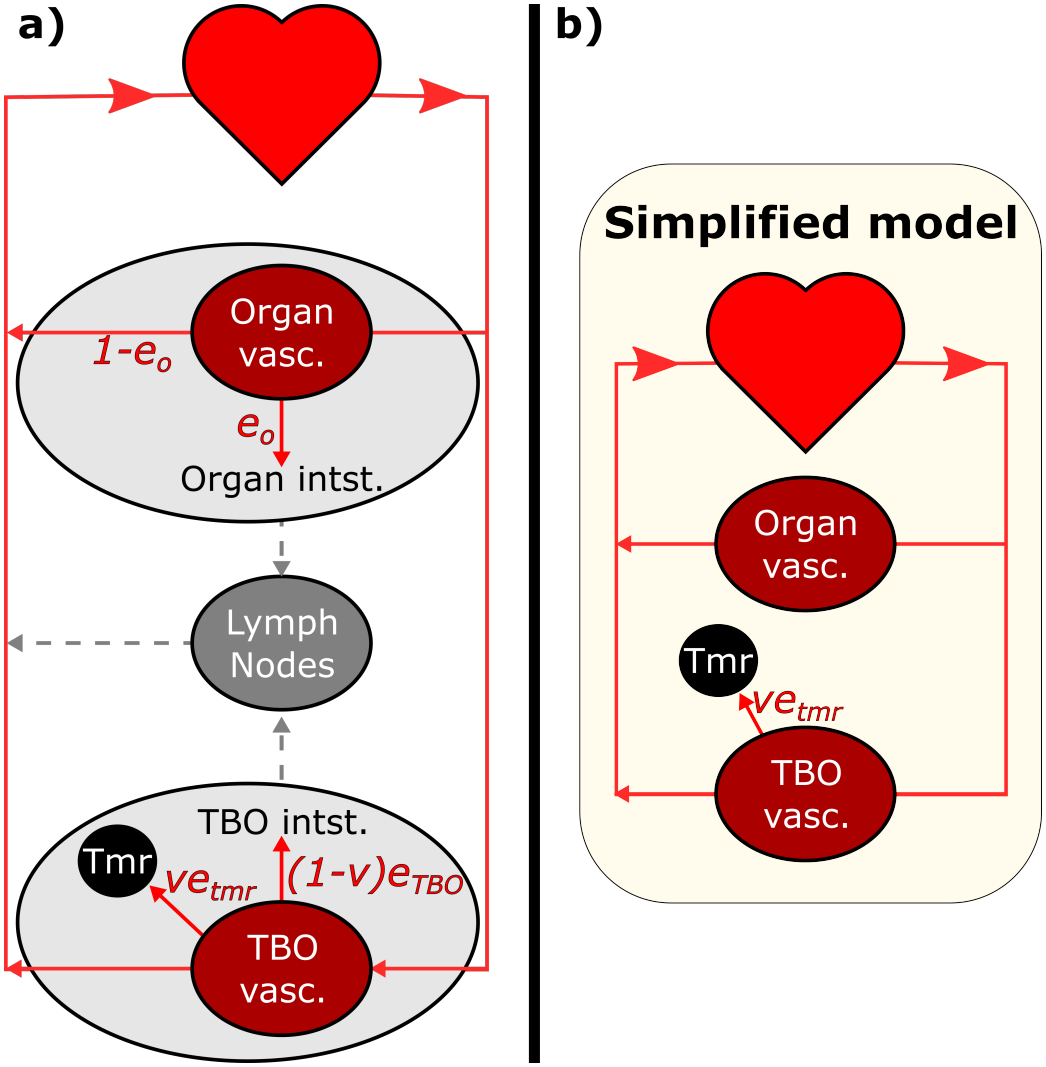


Figure 3.0. Model schematic: a) A visual summary of the presented model. Solid and dotted lines represent blood and lymph flow respectively. Cells flow from the heart to each organ o, from which a proportion e_o enters the interstitial space. Cells from the interstitium flow via the lymphatics back to the heart. One organ is designated as the tumour bearing organ (TBO), from which a proportion of cells $\frac{\tilde{V}_{\text{tmr}}}{V_{\text{TBO}}} e_{\text{tmr}}$ extravasate irreversibly to the tumour ("Tmr") instead of the organ interstitium, where $\frac{\tilde{V}_{\text{tmr}}}{V_{\text{TBO}}}$ is the proportion of the TBO volume taken up by the tumour. **b)** When calculating the maximum rate of extravasation to a tumour in each organ, a simplified version of the model that ignores the interstitial space may be used, as shown on the right. In both schematics, extra details such as the spleen or flow from the mesentery to the liver are not shown, but they are included in the model; see Section 3.1.0.

3.1 Methods

3.1.0 Model definition

3.1.0.0 Summary

A simplified schematic of the model is presented in Figure 3.0a. The model contains 20 organs (for humans), which each have a vascular and interstitial compartment, plus compartments for the heart and for tumour-invading lymphocytes. When required, the pulmonary

circuit is added, which has a vascular and 'trapped' state (functionally equivalent to interstitial space). The model's dependent variables are the vascular concentration (C) and interstitial concentration ($\tilde{C}$) of cells in each of these compartments, as a function of time t. Model parameters are the blood flow (B), vascular volume (V) and total volume ($\tilde{V}$) of each organ. T-cells flow from the heart to the vasculature of different organs and back again, with the exception of the mesentery, from which cells return via the liver. A proportion e_o of T-cells passing through each organ o extravasate into that organ's interstitial space. Cells in the interstitial space transit via lymphatic vessels to lymph nodes and then back into the bloodstream, with the exception of the spleen, from which cells return directly to blood. For the purposes of computation, one organ is defined as the tumour bearing organ, containing a 1mm^3 volume of target tissue labelled 'tmr' in the figure. A small volume is chosen for two reasons: some organs in the mouse or rat are comparable to this size, and larger tumour volumes approaching the size of the organ would introduce non-linearities. The maximum delivery per volume can be calculated from an arbitrarily small volume and so this choice is appropriate. The modelled target volume can either be healthy or tumour tissue. We have run the model targeting both healthy and cancerous tissue in turn, to find and compare the maximum delivery rates to both classes of tissue. The *maximum* delivery rate of cells to a target volume in each organ o is the case for which $e_o = 1.0$ in that region. This rate is insensitive to the values $e_{o'}$ in any other organ o' if the blood concentration of T-cells is constant (see sensitivity analysis, Section 3.2.6). Thus, the maximum delivery rate of T-cells per mm^3 is calculated for each organ in turn by setting e to zero everywhere but for a volume of 1mm^3 in the organ of interest. A schematic of the effective model is shown in Figure 3.0b. This results in a nearly constant T-cell blood concentration and thus correct calculation of the maximum rate of cell delivery to a tumour in each organ. For the sake of brevity, the remainder of the text will usually refer simply to "delivery rates", but this is more precisely "maximum delivery rates per mm^3 of tissue (or tumour)".

Predicted T-cell delivery rates are dependent on assumed anatomical parameters (blood flow, blood volume and organ volume). We collected several anatomical reference banks from the literature [172–177], each with slightly differing fractional blood flows and volumes. To remove selection bias from results, we ran our model many times ($n = 100$ per organ per animal), selecting uniformly random values for each anatomical parameter from the range of literature values, and calculated the mean and standard deviation of predicted delivery rates. For maximum delivery rates, it would also be possible to calculate the steady state of the model (which is approximately $N_o = V_o C_o N_0 = \frac{V_o}{V_h + \sum_{o'} V o'} N_0$, where N_0 is the initial and total number of lymphocytes in the system. This leads to a maximum delivery rate

to 1mm^3 of tumour of approximately $\frac{B_o}{\tilde{V}_o}\frac{N_0}{V_h+\sum_{o'} Vo'}$) and then perturb anatomical parameters to calculate this mean and standard deviation. However, maximum delivery rates are slightly different from this expression due to details such as the mesentery, and repeated calculation of ODEs with random parameters is not time-intensive, so the analytical approach was not necessary when random parameters must be computationally regardless.

3.1.0.1 ODE system

Symbol	Definition
V_o	blood volume in organ o
B_o	blood flow to organ o
C_o	(CAR) T-cell blood concentration (1/mm^3) in organ o
$\tilde{V}_o$	total volume of organ o
$\tilde{C}_o$	interstitial cell concentration in organ o
e_o	proportion of cells extravasating to interstitial space (Such that $e_oB_oC_o$ is the entry rate)
μ_o	proportion of cells leaving the interstitial space for the lymph node ($\mu_oB_o\tilde{C}_o$ is the exit rate)

Table 3.0. Definition of parameters described in Section 3.1.0. Different suffixes such as h for heart, tmr for the tumour and TBO for the tumour-bearing organ are also used.

The model has been implemented as a system of ODEs representing material exchange between the heart and organs. The change in concentration of immune cells in the vasculature of any organ C_o is given by,

$$V_o\dot{C}_o = B_o(C_\mathrm{h} - C_o), \tag{3.0}$$

where V_o and B_o are constant and are respectively the blood volume and the blood flow of the organ, whilst C_h is the immune cell concentration within the heart chambers. A proportion of cells leaving the organ e_oB_o extravasate to the interstitial space:

$$\tilde{V}_o\dot{\tilde{C}}_o = e_oB_oC_o - \mu_oB_o\tilde{C}_o \tag{3.1}$$

where $\tilde{V}_o$ is the total volume of the organ, $\tilde{C}_o$ is the interstitial concentration and μ_oB_o is the flow of immune cells into the lymph from the interstitium of organ o. Note that $\tilde{V}_o$ has not had the vascular volume subtracted from it, to avoid the combination of measurements from different experiments in a single organ's data. Both equations 3.0 and 3.1 are used for all organs, save a few exceptions. For example, both equations are modified for the liver,

which also receives blood exiting the mesenteric organs:

$$
\begin{aligned}
V_{\text{liver}}\dot{C}_{\text{liver}} &= B_{\text{liver}}(C_{\text{h}} - C_{\text{liver}}) + \mu_{\text{spleen}}B_{\text{spleen}}(\tilde{C}_{\text{spleen}} - C_{\text{liver}}) + \sum_{o' \in \text{mesentery}} B_{o'}(C_{o'} - C_{\text{liver}})(1 - e_{o'}) \\
\tilde{V}_{\text{liver}}\dot{\tilde{C}}_{\text{liver}} &= e_{\text{liver}}B_{\text{liver}}C_{\text{liver}} - \mu_{\text{liver}}B_{\text{liver}}\tilde{C}_{\text{liver}} \\
&\quad + \sum_{o' \in \text{mesentery}} (e_{\text{liver}}C_{\text{liver}} - \mu_{\text{liver}}\tilde{C}_{\text{liver}})B_{o'}(1 - e_{o'}) \\
&\quad + (e_{\text{liver}}C_{\text{liver}} - \mu_{\text{liver}}\tilde{C}_{\text{liver}})\mu_{\text{spleen}}B_{\text{spleen}},
\end{aligned}
\tag{3.2}
$$

where the summed terms are the extravasation and return of cells that did not extravasate in the mesentery and flowed through the portal vein, and the final line is cells returning from the spleen. Cells entering the lymphatics flow towards the lymph nodes:

$$
\begin{aligned}
V_{\text{LN}}\dot{\tilde{C}}_{\text{LN}} =& e_{\text{LN}}B_{\text{LN}}C_{\text{LN}} - \mu_{\text{LN}}B_{\text{LN}}\tilde{C}_{\text{LN}} \\
&+ \sum_{o' \notin \{\text{LN,spleen,PC}\}} \mu_{o'}B_{o'}(\tilde{C}_{o'} - \mu_{\text{LN}}\tilde{C}_{\text{LN}}) \\
&+ \mu_{\text{liver}} \left(\mu_{\text{spleen}}B_{\text{spleen}} + \sum_{o' \in \text{mesentery}} B_{o'}(1 - e_{o'}) \right) (\tilde{C}_{\text{liver}} - \mu_{\text{LN}}\tilde{C}_{\text{LN}}),
\end{aligned}
\tag{3.3}
$$

where the second line is the flow of lymphocytes from other organs, the third line is additional flow from the liver, and PC is the pulmonary circuit. Cells that do not extravasate from the organ vasculature return to the heart, in addition to the lymph node output:

$$
\begin{aligned}
V_{\text{h}}\dot{C}_{\text{h}} =& - C_{\text{h}} \sum_{o} B_o \\
&+ \sum_{o \notin \text{mesentery}} (1 - e_o)B_oC_o \\
&+ \left(\mu_{\text{spleen}}B_{\text{spleen}} + \sum_{o \in \text{mesentery}} (1 - e_o)(1 - e_{\text{liver}})B_{\text{o}} \right) C_{\text{liver}} \\
&+ \mu_{\text{LN}}B_{\text{LN}}\tilde{C}_{\text{LN}} + \sum_{o \notin \{\text{spleen,PC,LN}\}} \mu_{\text{LN}}\mu_oB_o\tilde{C}_{\text{LN}} \\
&+ \mu_{\text{LN}}\mu_{\text{liver}}\tilde{C}_{\text{LN}} \left(\mu_{\text{spleen}}B_{\text{spleen}} + \sum_{o \in \text{mesentery}} B_o\,(1 - e_o) \right) \\
&+ \mu_{\text{PC}}B_{\text{PC}}\tilde{C}_{\text{PC}},
\end{aligned}
\tag{3.4}
$$

where each line respectively refers to: (i) movement of cells out of the heart, (ii) return of cells from non-mesenteric organs, (iii) return of cells from the mesentery via the liver, (iv) return of cells from the lymph nodes, (v) extra return of cells from the liver via the LN, and (vi) return of cells from the pulmonary circuit.

Note that V_h is the blood volume of the heart chambers. Venous and arteriole compartments

are not modelled and are not counted in the total blood volume. We may designate an organ as a tumour-bearing organ (or any other disease of interest), and define the rate of immune cell entry to the tumour as,

$$\dot{N}_{\text{infiltrate}} = e_{\text{tmr}} \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}} B_{\text{TBO}} C_{\text{TBO}}, \tag{3.5}$$

where $\tilde{V}_{\text{tmr}}$ is the tumour volume and $\tilde{V}_{\text{TBO}}$ the organ volume of the tumour-bearing organ. Extravasation to the rest of the organ is correspondingly reduced:

$$\tilde{V}_{\text{TBO}} \dot{\tilde{C}}_{\text{TBO}} = (e_{\text{TBO}} C_{\text{TBO}} - \mu_{\text{TBO}} \tilde{C}_{\text{TBO}})(1 - \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}}) B_{\text{TBO}}. \tag{3.6}$$

The flow of cells that do not return to the heart from the TBO is subsequently changed from $B_{\text{TBO}} C_{\text{TBO}} e_{\text{TBO}}$ to $B_{\text{TBO}} C_{\text{TBO}} \left(e_{\text{TBO}} (1 - \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}}) + e_{\text{tmr}} \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}} \right)$, with changes for the liver and lymph nodes made analogously to before. Note that tumour-invading lymphocytes are assumed never to return via the lymphatics, and that these equations implicitly assume that tumour perfusion is equal to that of the rest of the organ. It may thus be understood as a target volume of healthy tissue; this formulation is used to find the maximum delivery rates to healthy tissue.

3.1.0.2 Tumour perfusion

If tumour perfusion (blood flow per volume) is expected to be different from the perfusion of the healthy organ, then we define new bloodflows B'_{tmr} and B'_{healthy} such that $B'_{\text{TBO}} = B'_{\text{tmr}} + B'_{\text{healthy}}$ and their perfusion ratio is,

$$p := \frac{B'_{\text{tmr}} / \tilde{V}_{\text{tmr}}}{(B'_{\text{TBO}} - B'_{\text{tmr}}) / (\tilde{V}_{\text{TBO}} - \tilde{V}_{\text{tmr}})}. \tag{3.7}$$

Note that $\tilde{V}_{\text{tmr}} \lesssim \tilde{V}_{\text{TBO}}$ and $B_{\text{tmr}} \lesssim B_{\text{TBO}}$ (that is, the volume and bloodflow of the tumour bearing organ explicitly include those of the tumour), so the perfusion is strictly positive. Equation 3.7 may be used to calculate an expression for the new tumour bloodflow in terms of p,

$$B'_{\text{tmr}} = B'_{\text{TBO}} \frac{p \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}}}{1 + \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}} (p - 1)}. \tag{3.8}$$

To modify tumour perfusion in this way whilst keeping the total blood flow constant, all other bloodflows must be reduced. This reduction can be derived as follows, where ΔP and $\Delta P'$ are the pressure drop across the heart with and without the tumour, $R_o = \Delta P / B_o$ is the vascular resistance to blood flow in organ o, the healthy part of the tumour bearing

organ is counted within the sum of organ bloodflows $\sum_o B_o$, and the total cardiac output B_{tot} and all vascular resistances other than that of the TBO are assumed to be constant:

$$\begin{aligned}
B_{\text{tot}} =& \sum_o B_o = \sum_o B'_o = \Delta P' \left(\sum_{o \neq \text{tmr}} \frac{1}{R_o} \right) + B'_{\text{tmr}} \\
& \therefore \Delta P' = \left(B_{\text{tot}} - B'_{\text{tmr}} \right) / \left(\sum_{o \neq \text{tmr}} \frac{1}{R_o} \right) \\
B'_o =& \frac{\Delta P'}{R_o} = B_o \frac{\Delta P'}{\Delta P} \\
=& \frac{B_o}{\Delta P} \left(B_{\text{tot}} - B'_{\text{tmr}} \right) / \left(\sum_{o \neq \text{tmr}} \frac{B_o}{\Delta P} \right) \\
=& B_o \left(B_{\text{tot}} - B'_{\text{tmr}} \right) / \left(B_{\text{tot}} - B_{\text{tmr}} \right) \\
=:& B_o \theta.
\end{aligned} \tag{3.9}$$

This allows the bloodflow to the healthy part of the tumour bearing organ to be written as,

$$B'_{\text{healthy}} = \theta B_{\text{healthy}} = \theta B_{\text{TBO}} (1 - \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}}) \tag{3.10}$$

Using the definition of $B'_{\text{TBO}} = (B'_{\text{healthy}} + B'_{\text{tmr}})$ and equation 3.8 allows B'_{TBO} to be written in terms of p and B_{TBO}:

$$B'_{\text{TBO}} = B_{\text{TBO}} \frac{1 + \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}} (p-1)}{1 + \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}} (p-1) \frac{B_{\text{TBO}}}{B_{\text{tot}}}}. \tag{3.11}$$

A similar expression can then be derived for B'_{tmr},

$$B'_{\text{tmr}} = B_{\text{TBO}} \frac{p \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}}}{1 + \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}} (p-1) \frac{B_{\text{TBO}}}{B_{\text{tot}}}} = B_{\text{tot}} \frac{p \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}}}{\frac{B_{\text{tot}}}{B_{\text{TBO}}} + \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}} (p-1)}, \tag{3.12}$$

which in turn allows θ to be written in terms of p:

$$\begin{aligned}
\theta =& \frac{B_{\text{tot}} \left(1 - \frac{p \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}}}{\frac{B_{\text{tot}}}{B_{\text{TBO}}} + \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}} (p-1)} \right)}{B_{\text{tot}} - \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}} B_{\text{TBO}}} \\
=& \frac{B_{\text{tot}}}{B_{\text{tot}} + \frac{\tilde{V}_{\text{tmr}}}{\tilde{V}_{\text{TBO}}} (p-1) B_{\text{TBO}}},
\end{aligned} \tag{3.13}$$

which finally allows the bloodflows to the tumour and all organs to be redefined in terms of p. Note that the total bloodflow to the tumour bearing organ, equation 3.11, is greater ($p > 1$) or less ($p < 1$) than the original bloodflow B_{TBO}, implying that growth of a vascular tumour increases overall bloodflow to the organ and that the growth of an avascular tumour decreases overall bloodflow to the organ.

3.1.0.3 Initial conditions

The model used in this chapter comprises of all of the equations defined above. Their initial conditions can take one of two forms. The first is a distribution of cells by blood volume,

$$\begin{aligned} V_{\mathrm{h}}C_{\mathrm{h}}(0) =& N_{\mathrm{tot}}\frac{V_{\mathrm{h}}}{V_{\mathrm{tot}}}, \\ V_oC_o(0) =& N_{\mathrm{tot}}\frac{V_o}{V_{\mathrm{tot}}}, \\ \tilde{V}_o\tilde{C}_o(0) =& 0, \\ N_{\mathrm{infiltrate}}(0) =& 0, \end{aligned} \tag{3.14}$$

where N_{tot} is the total number of lymphocytes in the system and $V_{\mathrm{tot}} = V_{\mathrm{h}} + \sum_o V_o$. The other form is a distribution of cells by blood flow,

$$\begin{aligned} V_{\mathrm{h}}C_{\mathrm{h}}(0) =& N_{\mathrm{tot}}\frac{V_{\mathrm{h}}}{V_{\mathrm{tot}}} \\ V_oC_o(0) =& N_{\mathrm{tot}}(1-\frac{V_{\mathrm{h}}}{V_{\mathrm{tot}}})\frac{B_o}{B_{\mathrm{tot}}} \\ \tilde{V}_o\tilde{C}_o(0) =& 0 \\ N_{\mathrm{infiltrate}}(0) =& 0, \end{aligned} \tag{3.15}$$

where $B_{\mathrm{tot}} = \sum_o B_o$. The relaxation time of the ODE system is fast, so results are not affected by the choice of initial conditions. Were it to matter, the latter choice of initial conditions would be used for injected T-cells (*e.g.* radiolabeled or CAR T-cells introduced to the bloodstream at time 0), and the former for endogenous immune cells (with uniform blood concentration).

3.1.0.4 Extravasation proportion values

In most results presented in this chapter, e_{tmr} is equal to 1.0 (i.e. all cells that enter the tumour vasculature extravasate) and $\tilde{V}_{\mathrm{tmr}}$ is $1\mathrm{mm}^3$, to find the *maximum* possible rate of immune cell entry to each $1\mathrm{mm}^3$ of tumour. This ignores the potentially reduced rate of extravasation in the tumour microenvironment [178], as well as the possibility for cells to enter the tumour from adjacent tissue. Though the range and rate of the latter is unknown, the defined tumour volume may equivalently be considered the effective volume from which extravasating cells may reach the tumour, and e_{tmr} the average proportion of cells that extravasate over this volume. In most presented results (Figure 3.2, Table 3.2 and Table 3.3), there is no tumour, but we are still interested in the delivery of T-cells to a $1\mathrm{mm}^3$ volume of tissue and hence use the same formulation with $p = 1$. Throughout the text, the $1\mathrm{mm}^3$ volume under consideration is referred to as a tumour, but in the cases where tumour perfusion is the same as the rest of the organ, this could be understood to

be delivery to normal (*i.e.* normally vascularised) tissue, notwithstanding that invading lymphocytes do not return after entering the tissue.

The relative rate of extravasation to the target tissue (tumour) in any organ is not sensitive to any value of e_o and μ_o except for e_{tmr}. Absolute rates presented in Table 3.2 are calculated assuming that the blood concentration of T-cells is constant, as the rates would fall with the concentration. If $e_o \neq 0$, the rate at which the simulated blood concentration falls would be fast enough to impact results. Thus, in order to calculate a realistic value for maximum extravasation rate within the tumour, e_o and μ_o are set to 0.0 for all organs but the tumour, thereby ensuring that simulated blood concentration of T-cells remains constant. Investigation of the balance of extravasation to each organ, $e_o \neq 0$, will be addressed in Chapter 4.

3.1.1 Parameter selection from literature

Organ	Blood flow/mL	Blood volume/mL	Organ volume/mL
Stomach	48.2 (30.4 – 66.5)	35.6 (8.1 – 119)	146 (143 – 154)
Small intestine	546 (344 – 753)	150 (4.0 – 533)	822 (385 – 1652)
Large intestine	314 (198 – 433)	66.2 (5.6 – 262)	423 (356 – 548)
Pancreas	83.5 (52.7 – 115)	18.0 (3.7 – 71.4)	104 (84 – 133)
Spleen	144 (90.9 – 199)	51.1 (10.8 – 167)	183 (144 – 221)
Liver	83.3 (49.5 – 113)	327 (68.6 – 1190)	1825 (1687 – 2143)
Brain	696 (447 – 1001)	110 (19.7 – 593)	1416 (1381 – 1450)
Heart b.s.	215 (139 – 326)	112 (8.5 – 734)	312 (266 – 341)
Lung b.s.	139 (96.0 – 183)	312 (35.5 – 1487)	958 (519 – 1169)
Kidneys	1119 (704 – 1464)	73.4 (11.8 – 243)	309 (295 – 332)
Fat	300 (194 – 484)	318 (94.1 – 906)	15,008 (10,010 – 19,791)
Skin	314 (199 – 454)	393 (56.5 – 2245)	4247 (2249 – 7770)
Skeletal muscle	946 (544 – 1559)	1762 (97.9 – 10,831)	28,819 (27,550 – 30,078)
Skeleton	275 (159 – 333)	463 (102 – 1767)	23,723 (7164 – 30,078)
Thyroid	78.0 (32.6 – 121)	3.2 (1.1 – 7.1)	20 (19 – 20)
Adrenals	16.7 (11.5 – 22.0)	3.2 (1.1 – 7.1)	15 (14 – 17)
Lymph nodes	141 (86.9 – 241)	11.2 (2.2 – 37.6)	564 (274 – 709)

Table 3.1. Anatomical parameter values used for humans, after compiling data from several published anatomical references [172–177] and eliminating outliers. Note that the liver row does not include the portal vein, and the lung row does not include the pulmonary circuit. *b.s.:* blood supply.

Blood flows, blood volumes and organ volumes in the model were parameterised from a variety of physiological reference data [172–177], in particular the compilations by the ICRP and Shah *et al* [172, 173]. As anatomical values vary greatly between and even within individuals, we use the variability between sources as a proxy for anatomical variability. To avoid using data from different studies for a single modelled animal, data sets that are as complete as possible were chosen. In particular, the total blood flow and blood volume, the

volume of each organ, and the fractional blood flow and blood volume of each organ were recorded from each reference. An example of such data found for humans is presented in Table 3.1. Presented results are the mean and standard deviation of predictions obtained by choosing random values from the literature. Random parameter values are selected from the range of literature values. We cannot be more confident in any one report than another, so we choose the random values for all of these parameters (and for the other organs and species) uniformly. This process is repeated 100 times to yield the presented results. When considering tumour perfusion distinct from healthy organ perfusion, we use measurements of tumour perfusion from the literature and suppose that, since these are all measurements of different tumours, the data should follow a normal distribution. Thus, we choose normally distributed random values of tumour perfusion.

3.2 Results

3.2.0 ODE results versus time

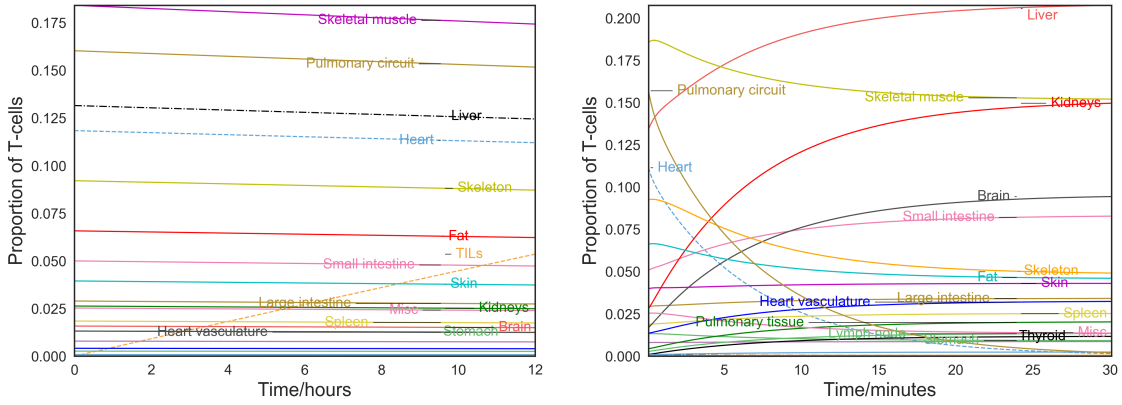


Figure 3.1. Example results from the ODE system defined in Section 3.1.0 for two different parameter sets described in the main text. The left plot shows a system in which T-cells may only extravasate into a tumour in the liver. The right plot shows a system in which there is no tumour and T-cells have an equal probability of extravasating into any organ. Each line represents the sum of vascular and interstitial localisation. The dashed line is the heart and the dash-dotted line is the tumour-bearing organ. TILs: tumour invading lymphocytes.

Before presenting results that consist of many runs of the system, it is insightful to consider the general time-dependent behaviour of a single iteration of the ODE system. Shown in Figure 3.1 are two examples of the ODE solution with human (ICRP [173]) values for the blood flow, blood volume and organ volumes, timestep $\Delta t = 0.1$ minutes and T-cells distributed according to organ blood volume as initial conditions. The first example has the extravasation probability equal to zero everywhere but for a 1mm^3 tumour in the liver, where it is 10%. The second example has no tumour, and the probability of extravasation

Organ	Human	Rat	Mouse
Lungs	162.2	394,520	2,871,160
Kidneys	97.8	28,962	239,926
Thymus	49.1	24,308	243,096
Pancreas	26.2	12,135	119,047
Spleen	25.6	12,368	119,202
Heart vasculature	20.3	28,310	436,944
Large intestine	20.9	10,538	102,045
Liver	20.7	11,234	102,993
Small intestine	28.6	15,263	147,464
Brain	13.2	5484	45,196
Lymph nodes	11.9	2797	27152
Stomach	11.2	17731	6104
Skin	3.0	768	10306
Skeletal muscle	1.0	1452	14177
Fat	0.7	1296	12593
Skeleton	0.2	560	10031

Table 3.2. Predicted absolute maximum CAR T-cell delivery rates per volume (cells/min/mm^3) by organ and species, assuming homogeneous organ perfusion and that 10^8 CAR T-cells are introduced to each species, and using previously compiled physiological parameter values [172]. The interspecies differences in absolute delivery rates per volume depend only on organ perfusion.

is 10% everywhere. The maximum time is reduced for 30 minutes for this example so that early dynamic behaviour can be seen more clearly.

3.2.1 CAR T-cell delivery to organs in humans, mice and rats

The model was used to compare the delivery rate (cells/min/mm^3) of a typical number of CAR T-cells used in the clinic (10^8 [4, 54]) to healthy tissues in different human, rat and mouse organs. The results, shown in Table 3.2, are approximately equal to the product of the organ perfusion and CAR T-cell blood concentration, with some deviation due to the details of the circulation or limited organ size. Note that the flow from both the hepatic

Organ	Equivalent dose	(continued)	
Lungs	1.8×10^{11}	Small intestine	5.2×10^{10}
Kidneys	2.5×10^{10}	Brain	3.4×10^{10}
Thymus	5.0×10^{10}	Lymph nodes	2.3×10^{10}
Pancreas	4.5×10^{10}	Stomach	5.5×10^{9}
Spleen	4.7×10^{10}	Skin	3.4×10^{10}
Heart vasculature	2.2×10^{11}	Skeletal muscle	1.4×10^{11}
Large intestine	4.9×10^{10}	Fat	1.8×10^{11}
Liver	5.0×10^{10}	Skeleton	5.0×10^{11}

Table 3.3. The dosage of CAR T-cells in humans predicted to be required to give the same absolute delivery rate per mm^3 as in a mouse given 10^7 cells. The numbers required are much larger than many clinical dosages [4, 54].

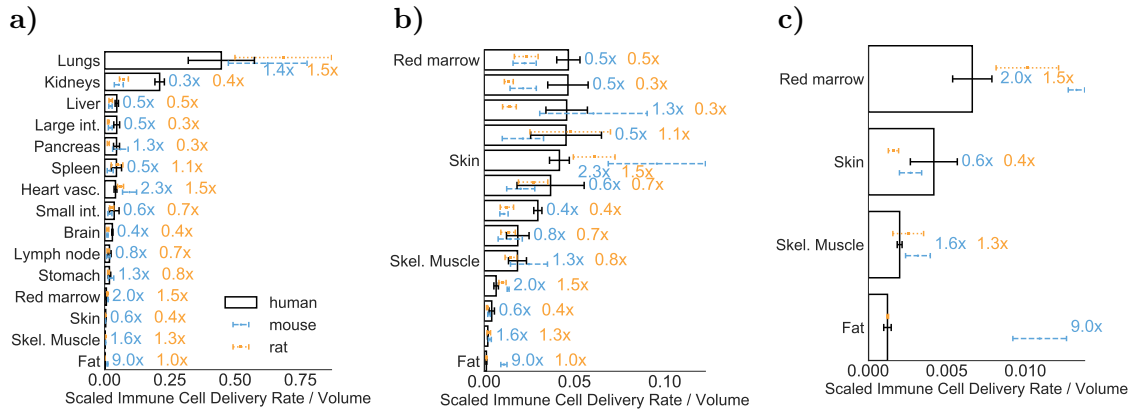


Figure 3.2. a) Relative predicted delivery rates for humans, rats and mice. Rates within each species were normalised to sum to 1.0 to give relative values for comparison. See Table 3.2 for a comparison of absolute rates. Predictions are presented as a mean and standard deviation over 100 repeats, with random fractional blood flows and volumes uniformly drawn from experimental data in the literature. Numerical values (text labels) give the mean predicted delivery rates in rats and mice relative to human values, with colours matching the legend. The relative distribution of rates across the major organs differs by species, and, consequently, inter-species scaling of delivery rates is organ-specific. **b,c)** Copies of plot **a** with successively fewer organs, to increase visibility of organs with lower delivery rates. Blue and orange colours have the same meaning as in plot **a**.

artery and portal vein are included in delivery rates to the liver, and the pulmonary circuit and lung blood supply are both included in the rates for the lung. This is the case for all presented results, unless otherwise stated. The difference in delivery rates to the same organ in different species may be extreme. For example, the predicted absolute lung delivery rates per volume in the mouse are 17,700 times higher than in humans if the same number of CAR T-cells were introduced to each species (obtained by dividing 2,871,160/162.2 from Table 3.2). Should a constant blood concentration of endogenous cells be considered instead of a constant number, then rates per volume depend on both organ perfusion and total blood volume, and the absolute delivery rates for mice are 8.4 times higher than in humans (data not shown). These data suggest a novel approach for scaling murine dosages to humans, by ensuring that the same delivery rate of cells to each organ is achieved. The results of Table 3.2 have been used to calculate the CAR T-cell doses (introduced cell numbers) required to obtain the same delivery rates in humans as in mice given a typical pre-clinical dose of 10^7 CAR T-cells. The results are shown in Table 3.3; equivalent doses are organ-specific and most are of order 10^{10} to 10^{11} cells.

To illustrate this organ-specific scaling and to allow interspecies comparison of the distribution of delivery rates across organs, predicted delivery rates in Table 3.2 were subsequently normalised such that the sum of the mean predictions within each species is 1.0. The mean and standard deviation of predicted delivery rates, normalised by species to sum to 1.0, are plotted in Figure 3.2. The distributions share similarities but otherwise the relative rates

scale differently for each organ. For each species, the lung has the highest delivery rate, followed by the kidneys.

3.2.2 Tumour perfusion compared to healthy tissue

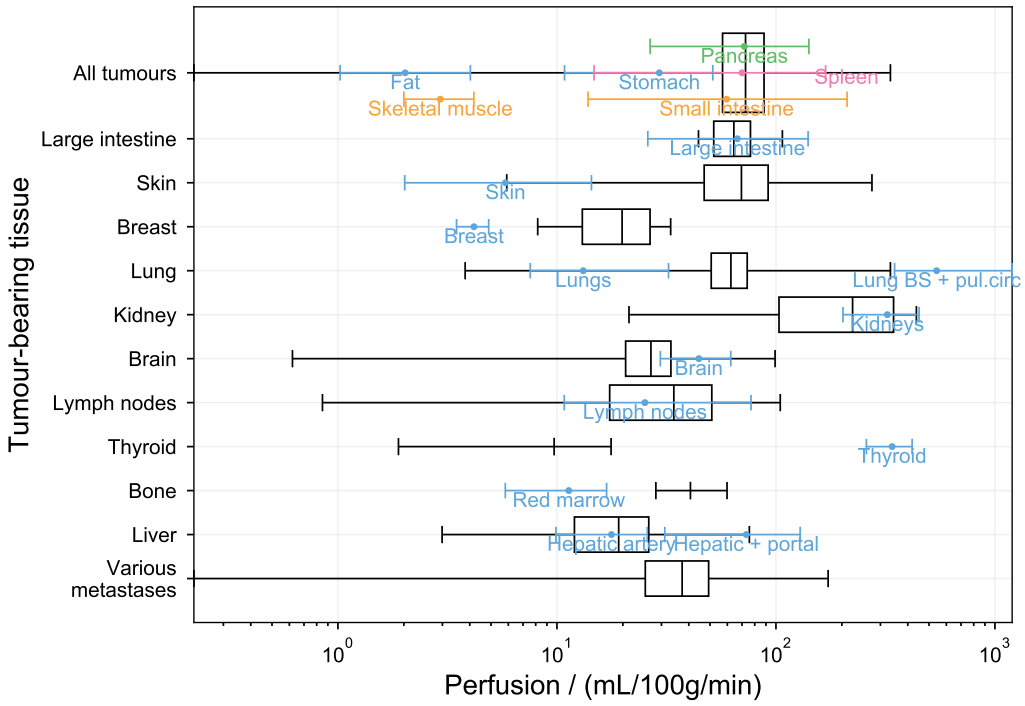


Figure 3.3. Perfusion of human tumours: Tumour perfusion sourced from literature [179–224]. Data are arranged by tumour location, with all lung-located tumours in the 'Lung' row. Black boxes represent the mean ± standard deviation of tumour perfusion and whiskers indicate the range of tumour perfusion found in the literature. Coloured error bars represent the mean and range of perfusion for healthy tissue calculated from human anatomical parameters. Note that the horizontal axis is logarithmic. "Lung BS + pul.circ" indicates the perfusion of the lung when blood from both the lung blood supply and the pulmonary are considered, whereas the "Lungs" entry indicates the perfusion when only the blood supply is considered.

Predicted maximum delivery rates per mm^3 described above assume that organ perfusion is homogeneous. However, a tumour may have very different perfusion to healthy tissues. A series of literature searches were carried out to extract a quantitative sense of the variability of reported perfusion data for various types of human tumours. The data sourced from literature [179–224] are presented in Figure 3.3. Tumour perfusion can be greater or lower than that of the source tissue, varying by many orders of magnitude. The range is likely to represent tumours that do or do not secrete angiogenic factors, or whose new vasculature has been renormalised or not. Whether tumour perfusion typically outstrips that of the bearing organ appears to be linked to the function of the organ itself, *e.g.* the kidneys have evolved to have very efficient perfusion that is unlikely to be outstripped by random,

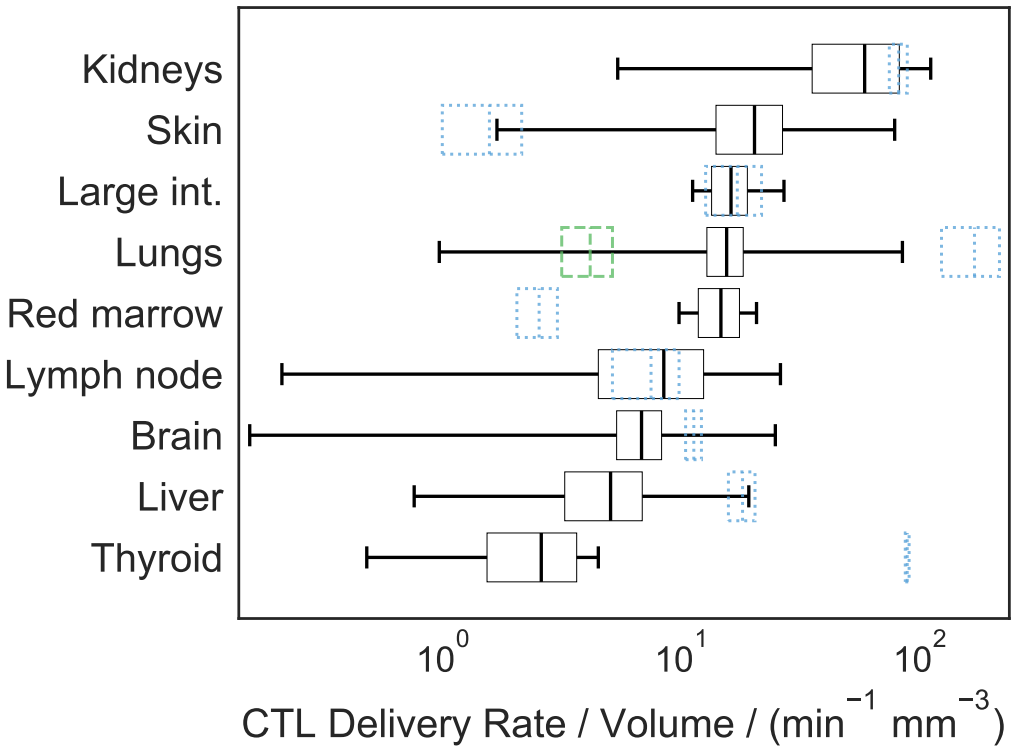


Figure 3.4. Absolute predicted delivery rates of CAR T-cells to human tumours. Predictions are presented as a mean and standard deviation over 150 repeats, with random anatomical parameters and tumour perfusion drawn from experimental data in the literature. Black boxes represent the mean and standard deviation of predicted tumour delivery rates, and whiskers indicate predictions using the extremes of possible tumour perfusion according to the literature. Blue dotted boxes indicate the mean and standard deviation of predicted delivery to healthy tissue, *i.e.* the data used to generate Figure 3.2. The green dashed box indicates delivery rates per mm^3 to healthy lung tissue when the pulmonary circuit is assumed not to contribute. Note that now the kidneys and skin have the highest predicted tumour delivery rates, and that the horizontal axis is a log scale.

disordered angiogenesis.

3.2.3 CAR T-cell delivery to human tumours

We used tumour perfusion data to implement tumour-location-specific perfusion into our model. The model was run many times ($n = 150$ per organ), drawing uniformly random values for all organ parameters and normally distributed values of tumour perfusion from the extracted literature data, to find a mean and deviation of predicted delivery rates for CAR T-cells to human tumours. As before, it would also be possible to calculate maximum delivery rates analytically and perturb anatomical parameters that way, as values are roughly proportional to $B_o/\tilde{V}_o$ for each organ (with slight deviations due to details such as the mesentery). However, tumour perfusion would complicate such calculations and repeated runs of the model are not time-intensive, so this approach was taken for simplicity. These rates are shown in Figure 3.4, along with the corresponding delivery rates under the

Cell Type	Rate Relative to CARs	Biological Variation
B Cells	240×	24× – 400×
Total T Cells	400×	64× – 560×
CD4+ T-cells (Th)	350×	130× – 540×
CD8+ T-cells (CTL)	160×	32× – 410×
NK cells	95×	0× – 350×
Monocytes	80×	0× – 290×

Table 3.4. Relative delivery rates by immune cell type: Delivery rates of immune cells to organs are proportional to blood immune cell concentrations, thus, predicted delivery rates scale trivially when a different cell concentration is considered. Published data on human blood immune cell concentrations [225, 226] were used to calculate mean delivery rates and inter-individual variation for B, T helper, cytotoxic T-lymphocyte (CTL), natural killer and monocyte cells, relative to the values given for CAR T-cells in this study. For example, assuming a human CTL concentration of $C = 3 \times 10^3$ cells/µL and a total blood volume of 5.3L [173] yields approximately 1.6×10^{10} cells, or 160 times the values quoted for CAR T-cells in this study. This, however, assumes that *all* blood CTLs are relevant to the target of interest.

assumption of homogeneous perfusion (*i.e.* to healthy tissue; blue dotted boxes). The rank order of delivery rates to tumour and healthy tissues are very different. In most cases, the average of predicted delivery rates for tumour tissue is similar to or less than that for the healthy tissue, but in some cases (*e.g.* the skin) it is considerably greater. However, their variation is considerable; extreme values (whiskers) vary over many orders of magnitude above and below that of the corresponding healthy tissue, for most organs.

3.2.4 Delivery rates of other immune populations in humans

The results of the model are relevant not just to CAR T-cells but to any immune cell population of a known blood concentration or count. The data in Table 3.2 may be used by multiplying results by total blood volume and the blood concentration, or count, of an immune cell population of interest. The mean and range of observed values [225, 226] for several immune cell blood concentrations are shown in Table 3.4.

3.2.5 Tumour control

If one assumes that T-cells arrive at a constant delivery rate per volume, kill tumour cells at a constant rate per T-cell and have a constant probability of anergisation (or death or exhaustion) after each kill, then a conservation relation can be written:

$$\text{Critical growth rate} = \frac{\text{T-cell delivery rate}}{\text{T-cell anergisation probability per kill}}, \tag{3.16}$$

where the *critical growth rate* (equal to $\ln(2)/(\text{tumour doubling time})$) is the growth rate of a tumour that can only just be eliminated by T-cells arriving at a given rate. This assumes that T-cells do not proliferate, die or compete for space with the tumour, and that the tumour contains enough cells for continuum approximation to be valid. These

critical tumour growth rates have been calculated for various human organs and are given in Table 3.5, assuming that 10^8 CAR T-cells are introduced, that delivery rates per mm^3 of tumour are as presented in Figure 3.4, and that a T-cell can kill on average 10 cancer cells before anergising [41, 42]. These correspond to the fastest growing tumours that CAR T-cells could control in each organ, if every cell that was delivered by the vasculature were to extravasate and the CAR T-cell population did not proliferate.

Tumour organ	**Maximum controllable growth rate/day^{-1}**	
Kidneys	3.47 ± 1.46	(0.324 - 6.66)
Skin	1.24 ± 0.37	(0.101 - 4.70)
Large intestine	1.02 ± 0.18	(0.668 - 1.61)
Lungs	0.96 ± 0.18	(0.058 - 5.06)
Skeleton	0.88 ± 0.16	(0.586 - 1.24)
Lymph node	0.53 ± 0.23	(0.013 - 1.56)
Brain	0.38 ± 0.10	(0.009 - 1.49)
Liver	0.28 ± 0.10	(0.045 - 1.15)
Thyroid	0.15 ± 0.07	(0.029 - 0.27)

Table 3.5. Maximum growth rates of tumours that can be eradicated by CAR T-cells delivered from the blood, assuming tumour perfusion values presented in Figure 3.4, that 10^8 CAR T-cells are introduced to the bloodstream, that *all* T-cells extravasate from tumour vasculature and do not proliferate, and that a tumour cannot persist if T-cells are delivered faster than the product of tumour growth rate and the probability of T-cell anergisation per kill. Values in the table are presented as mean ± standard deviation (range), corresponding to tumour perfusion values found in the literature, as in Figure 3.4.

3.2.6 Sensitivity analysis

In order to keep the blood concentration of simulated (CAR) T-cells constant and thus calculate a realistic maximum extravasation rate, T-cells must not be allowed to extravasate rapidly into all tissues. To ensure that this is true, extravasation probabilities for all organs are set to zero but for a unit volume in an organ of interest. To show that setting other values to zero does not affect results, we performed a local sensitivity analysis on the model with $p = 1$. Random values of the extravasation probabilities of T-cells into all 20 organs and the "tumour" ($1mm^3$ volume of interest) between 9% and 11% are chosen. Setting the tumour to be in each organ in turn, the extravasation rate of T-cells into the tumour is calculated for all organs. This process was repeated over 1000 simulations and the data was used to grow a random forest for each organ, using Scikit [135]. This process produces indices between 0 and 1 which indicate the 'importance' of each parameter for prediction of the extravasation rate of T-cells into a tumour in a given organ. On average, the importance factor of the tumour-specific extravasation proportion was 0.91. The importance factor of all other organs, including the tumour-bearing organ, was 0.004. The average Pearson rank correlation between the extravasation rate of T-cells into the tumour and the tumour-specific

	Tumour bearing organ (predicted delivery rate to organ):																			
Extrvstn prob into↓	Sto	SI	LI	Pan	Spl	Liv	Bra	Hrt	Lng	Kid	Fat	Skn	Mus	Bon	Thy	Gnd	Adr	Bla	LNs	Mis
Tumour	0.916	0.909	0.911	0.912	0.909	0.907	0.905	0.908	0.906	0.906	0.916	0.911	0.914	0.919	0.906	0.913	0.906	0.908	0.905	0.919
Stomach	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Small intestine	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Large intestine	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Pancreas	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Spleen	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Liver	0.004	0.004	0.004	0.004	0.004	0.005	0.005	0.004	0.005	0.004	0.004	0.005	0.004	0.004	0.004	0.004	0.005	0.005	0.005	0.004
Brain	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Heart vasculature	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Pulmonary tissue	0.003	0.003	0.003	0.003	0.002	0.002	0.003	0.003	0.003	0.003	0.002	0.002	0.003	0.002	0.003	0.002	0.003	0.003	0.003	0.002
Kidneys	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Fat	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.003	0.004	0.004	0.004	0.004	0.004	0.003
Skin	0.002	0.003	0.002	0.002	0.002	0.003	0.003	0.003	0.003	0.003	0.002	0.003	0.003	0.002	0.002	0.002	0.003	0.003	0.002	0.002
Skeletal muscle	0.053	0.058	0.057	0.056	0.058	0.055	0.061	0.058	0.060	0.060	0.053	0.057	0.055	0.051	0.061	0.055	0.060	0.058	0.061	0.052
Skeleton	0.003	0.003	0.003	0.003	0.003	0.004	0.003	0.003	0.004	0.003	0.003	0.003	0.003	0.003	0.004	0.003	0.003	0.003	0.004	0.003
Thyroid	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Gonads	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Adrenals	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Urinary bladder	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Lymph node	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Misc	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Table 3.6. Sensitivity of the predicted delivery rates to different tumour bearing organs to different parameters, derived from the random forest algorithm (see text). The parameters tested are the probabilities of T-cell extravasation e_o into each organ o. A sensitivity of 0 means no effect on the observable and 1 is maximal importance. Columns are abbreviations of the organs as written on the left of each row, starting from "Sto" for "Stomach".

extravasation proportion was 0.95, and the coefficient with the extravasation proportion into all other organs was -0.04. Exact data are reproduced in Table 3.6.

3.3 Discussion

This study aimed to quantify physiological constraints on the rate of CAR T-cell delivery by the blood to target tissues in different species, to better understand how to determine optimal human CAR T-cell doses from pre-clinical data. It has focused on tumour infiltration, though the methodology may also be applicable to other therapeutic areas such as autoimmune disease or immune-related adverse event prediction. Values were calculated by assuming that 10^8 CAR T-cells are introduced to the blood; however, delivery rates due to any other desired number or blood concentration of cells can be calculated by scaling – multiplying the presented results by the ratio of the desired number to 10^8, and multiplying blood concentration by the total blood volume in the target species if required. Though models to predict expansion of a T-cell population have been studied in the past [227], it is difficult to quantify how much a given population would proliferate in a given organ, and how much would return to the blood. A maximum rate of delivery due to the anatomy can be estimated with more certainty, so proliferation was not considered and is left to future work.

3.3.0 Organ-specific delivery rates and their variation

Results shown in Figure 3.2 show that the organs predicted to have the highest CAR T-cell delivery rates are those with the highest perfusion: the lungs and kidneys in humans. When tumour-specific perfusion is considered as in Figure 3.4, however, it is the kidneys, skin, large intestine and lungs that are predicted to have the highest delivery rates per mm^3. This result is interesting because some of the first successes of non-cellular (IL-2 and checkpoint blockade) immunotherapy were in the kidney, skin, colon and lung [228–233], consistent with the hypothesis that efficacy is driven in part by organ perfusion and hence delivery rates. In the case of cellular therapies such as CAR T-cells, delivery by the blood should similarly correlate with efficacy, but there is the additional complication that T-cells must extravasate into the target tissue. T-cells, particularly naïve cells, preferentially extravasate into lymph nodes, spleen and liver [167–169], which may explain why CAR T-cells have previously been highly successful for leukaemia and lymphoma but not solid tumours [4, 54, 163]. Another striking result in Figure 3.4 is that predicted mean delivery rates to tumours exceed those to healthy tissue for only a minority of organs. Predicted delivery rates to tumours in the skin vary over many orders of magnitude, but are usually greater than those for healthy tissue. Healthy skin is not usually highly perfused and it contains shunts to control bloodflow in

response to temperature. Most anatomical data for the skin describes the organ at rest and at room temperature with no inflammation, meaning most of these shunts will be open. Tumour tissue can increase its perfusion through inflammation or angiogenesis and likely subverts these shunts, which could explain both the greater mean and variation in predicted delivery rates for skin tumours. The liver and kidneys, by comparison, are highly ordered organs whose perfusion is not likely to be improved by random angiogenesis. Delivery rates to tumours in these organs are hence not predicted to exceed that of healthy tissue. The thyroid may be explained similarly, or these results may alternatively be due to the small size of the organ, given that delivery rates are calculated per mm^3. Results for the skeleton are likely erroneous, owing to the ambiguity in defining anatomical parameters for 'bone' tissue. Finally, predicted delivery rates to lung tumours may or may not exceed that of healthy tissue, depending on whether the pulmonary circuit is assumed to contribute to tumour blood supply (blue dotted box in Figure 3.4) or not (green dashed box).

Both Figures 3.2 and 3.4 also show that predicted delivery rates are highly variable. Variation in organ delivery rates may be caused by differences in experimental techniques or it may be due to individual variation: individuals with the same gender, height, weight and lifestyle may still differ in bloodflow distribution. More critically, however, the bloodflow distribution within one individual varies. For example, bloodflow to the mesentery increases after a meal, or to the muscles during exercise, or to the skin in response to temperature. This is already utilised in the clinic: hair loss in chemotherapy patients, for example, can be reduced by cooling the scalp. Such techniques could and should be utilised for CAR T-cell therapies: delivery to the mesentery could be increased by providing patients with a meal, delivery to the skin could be controlled by heating or cooling, and tumour-specific bloodflow could be increased with anti-angiogenic therapies such as Avastin (which cause vascular normalisation). This variation in anatomical parameters is, however, a double-edged sword for research. Care should be taken when interpreting results that are, or depend on, anatomical parameters from few individuals. If variability is not captured and/or care is not taken to control factors like anaesthesia or the time of day [234], then comparison of data sets may be invalid. Ideally, any study making use of bloodflows and organ volumes should try to consider multiple measurements and include 'error' bars to indicate variation.

3.3.1 Species-specific delivery rates and dosage scaling

Figure 3.2 and Table 3.2 show predicted maximum delivery rates for humans, mice and rats. Relative delivery rates are distributed among organs differently in each species, meaning that dose scaling is organ-specific. Furthermore, predicted values of absolute delivery rates of the same dose of CAR T-cells (10^8) exhibit extreme differences in different species, with

delivery to mouse lungs 17,700 times higher than to human lungs. This is because the total blood volume in a mouse (1.5mL) is considerably lower than that in humans (5L). Despite this, typical doses (cell numbers) introduced to mice are not considerably lower than those given to humans. Most patients are given CAR T-cell dosages between 10^7 and 10^9 cells [4, 54]. Examples of mouse studies for CAR T-cells featured two doses of 1 to 2.5×10^6 cells a week apart [166], two doses of 10^7 cells a week apart [164], and a single dose of 10^7 cells [165]. To illustrate how large these doses are, we calculated the equivalent human dosages that would yield the same absolute delivery rates in humans (Table 3.3) as in a mouse given 10^7 CAR T-cells. The resulting doses range between 10^{10} and 10^{11} T-cells, much higher than typical clinical doses and many dose escalation studies [54]. This may explain why pre-clinical success does not always translate to the clinic. A pre-clinical model of CEA CAR T-cells achieved tumour regression in mice given subcutaneous tumours [86] with a dose of 5×10^6 cells (equivalent of 1.7×10^{10} in humans) following a dose of sublethal radiation. Another reduced all tumours below the detection limit in mice with pancreatic tumours [235] with a dose of 10^7 cells (equivalent of 4.5×10^{10} in humans). In the clinic, a study of CEA CAR T-cells against colorectal cancer [171] escalated doses between 10^7 and 10^{10}. The authors found that the lower doses did not stop tumour progression (in 3 of 14 of presented patients) and higher doses achieved only stable disease. By comparison, our results suggest that dosages of order 10^{10} cells would be required to drive tumour regression at the primary site, and of order 10^{11} would be required for the lung metastases. Clinical studies in which Tumour Infiltrating Lymphocytes (TILs) were introduced in greater numbers (10^9 to 10^{11}) [236–239] and in which CAR T-cells were introduced regionally (bypassing trafficking) [4] are associated with greater efficacy.

In many patients, however, there may exist several metastatic sites which would complicate the regional administration of cells. To drive disease regression at all sites, T-cell dosage must be sufficient to drive tumour regression at the least perfused and/or the fastest growing site. The high dosages required may, however, raise the risk of autoimmune pathology. To avoid unnecessarily high dosages, the distribution of T-cell delivery must be optimised for the required sites, by manipulating bloodflow with agents such as Avastin or interventions such as meal consumption, as described above.

The numbers presented here compare organs like-for-like between mice and humans, but scaling is even more uncertain for xenografts. The ratio of the maximum absolute cell delivery rate to skin tissue between mice and humans is 2 if the same blood concentration of immune cells is assumed, or 3400 if the same number of immune cells is assumed. The ratio of absolute delivery rates to mouse skin versus human kidney tissue, meanwhile, is 0.05 if the same concentration of cells is used, or 100 if the same number of immune cells is assumed.

A previous study [179] has shown that small xenografts have similar local perfusion to the original tissue, but that as they grow larger, the perfusion is reduced relative to normal tissue. This non-linearity may cause preclinical results to be difficult to extrapolate or to be non-indicative of human responses, and may explain some of the historically observed difficulties with both the growth of xenograft models and the translation of their successful therapeutic results [240].

The problems with dosage and xenografts make the interpretation of pre-clinical success difficult, even before considering differences in biology such as the receptors required for T-cell extravasation. To scale dosages between mice and humans correctly, a model that considers organ-specific bloodflow and volume is required to determine equivalent dosage. However, if a theoretical model is required to interpret preclinical results and they are nonetheless still unreliable, it calls the value of pre-clinical studies into question, particularly in the light of the three Rs (replace, reduce and refine [animal studies]). The use of labels to non-invasively track T-cells in patients would allow us to better understand T-cell homing and localisation in humans, in turn allowing refinement of our predictions of the dosages required to drive efficacy in patients, when coupled with appropriate scaling of pre-clinical dosages and predictions.

CAR T-cell delivery rates per mm^3 to tumours were used to calculate the highest growth rate of tumours that they could eliminate, shown in Table 3.5. These values assume that all T-cells in the tumour vasculature will extravasate, thus these growth rates are again an upper bound. They were calculated with a simple balancing argument: the tumour size remains constant only if tumour cells replicate at a rate equal to T-cell killing. We assume that T-cells do not proliferate, die or affect the tumour carrying capacity (maximum tumour size that can be supported by the surrounding tissue), but have a finite anergisation probability per kill. This argument thus represents early invasion of the tumour. Tumour growth rates can be related to doubling times by the relation growth rate = $\ln(2)$/Tumour doubling time. Typical volume doubling times for human tumours range from 1 to 400 days, with most between 50 and 300 [241]. These correspond to growth rates of 10^{-3}/day to 10^{-1}/day (most between 10^{-3}/day to 10^{-2}/day), which are of a similar order of magnitude to the calculated maximum growth rates. This suggests that in the absence of barriers to T-cell extravasation and engagement, most tumours are controllable by the immune system. However, in many cases, tumour vasculature is very poor, limiting the arrival and extravasation of T-cells, even if they would normally preferentially extravasate in the target organ. Alternatively, the cells may extravasate but become trapped at the tumour periphery (a so-called 'immune desert'). Thus, these data might indicate that 10^8 CAR T-cells are too few to handle many solid tumours.

3.3.2 Refinement of delivery rate predictions

Model predictions include the maximum rate of CAR T-cell delivery per volume (cells/min/mm^3) to organs and tumours, based on measurements of organ blood flows and volumes, and tumour perfusion. As previously discussed, refining these predictions requires two considerations: the extent of proliferation and organ-specific homing of CAR T-cells. The probability of T-cell extravasation differs by location and cell type. For example, naïve T-cells extravasate mainly into the lymph nodes or spleen and activated cells have a higher probability of extravasating into outer tissues [167, 242], distributed according to the homing receptors they have upregulated. These probabilities may also differ from species to species, which would further limit inter-species extrapolation of pre-clinical results. Both extravasation and proliferation of cells would require either detailed mechanistic modelling or fitting to data, and are left to future work.

Another challenge for CAR T-cells in solid tumours is the identification of suitable target antigen. The ideal antigen is highly expressed on tumour cells and not expressed on healthy cells elsewhere. A typical target for B-cell malignancies is CD19 [54], as it is expressed by the entire pool of B-cells and is limited almost exclusively to B-cells. Several different antigens have been targeted for solid tumours, but only GD2 has had encouraging results [54]. Target antigen may only be expressed by a subset of tumour cells, and may not be sufficiently rare elsewhere in the body. For example, CAIX is expressed in some renal cell carcinomas, but it is also expressed in the liver bile duct and resulted in on-target, off-site toxicities in a phase III trial [243]. Tumours may also evolve to reduce expression of target antigen in response to successful T-cell killing, exacerbating existing problems. Although these considerations are a barrier to treatment success, the rate at which cells can be delivered is a parallel and important factor. CAR T-cells that are specific for an antigen that is expressed on most tumour cells will not drive tumour regression if their kill rate is lower than the tumour growth rate, given the combined rates of T-cell delivery and proliferation. On the other hand, CAR T-cells specific for a rarer antigen may drive tumour regression if they arrive in such sufficient numbers to eliminate all cells carrying that antigen and drive a secondary response against one or more other antigens (*i.e.* epitope spread). Like T-cell delivery rates and T-cell extravasation probabilities, typical tumour growth rates are species, organ and individual specific. Hence, the extension of the model to consider specific T-cell proliferation rates, homing and tumour growth rates would further develop insight into the determination of suitable clinical doses from pre-clinical immunotherapy studies.

Together, these considerations show that tumour immunotherapy is often a 'numbers game' and hence quantitative studies can be a useful tool for understanding clinical and pre-clinical success and failure.

3.4 Conclusions

ODE models of human, rat and mouse circulatory systems were implemented to predict CAR T-cell delivery to human tumours, and to human, rat and mouse organs. Models show up to 17,700-fold greater delivery of T-cells in mice than humans, and typical clinical dosages of CAR T-cells correspond to delivery rates 100-fold less than those in typical pre-clinical studies. This could explain why pre-clinical models of solid tumour eradication by CAR T-cells see much greater efficacy than in the clinic. Dosage scaling was found to be organ-specific and is particularly hard to quantify for xenografts, frustrating the interpretation of pre-clinical results and lowering their potential value, which is an important consideration in the context of the 3Rs. Control of tumour and organ-specific bloodflow through measures such as meal consumption or the use of factors that induce tumour vascular normalisation could increase delivery of cells to tumour sites without raising the chances of autoimmune pathology. These findings may allow us to better understand how pre-clinical results scale to the clinic and hence better determine appropriate clinical dosages.

This predictions of this chapter are the maximum rates of T-cell extravasation to different organs. To find these rates, the probabilities of extravasation e were set to zero everywhere but in the tumour, to keep blood concentration constant. This implies that T-cells are primarily blood-borne, which is not true: the majority of T-cells at any moment are in the interstitial space and particularly the lymphatics. A model of the localisation of T-cells must thus estimate these values of e. The following chapter will describe the use of lymphocyte localisation data to fit the model, and the modifications made to the model to facilitate this.

3.A Tumour growth model

Coupled equations for T-cell infiltration and logistic tumour growth can be defined that yield the relation,

$$\text{Critical growth rate} = \frac{\text{T-cell delivery rate}}{\text{Tcell anergisation probability per kill}},$$

given in the main text. The evolution of the number of tumour (**c**ancer) cells C and T-cells T are assumed to be given by,

$$\begin{aligned} \dot{C} &= rC(1-\frac{C}{K}) - \kappa(C,T) \\ \dot{T} &= DC - a\kappa(C,T) \\ \kappa(C,T) &= k_0 T \frac{C}{C+C_*}, \end{aligned} \tag{3.A.0}$$

where r is the tumour growth rate, K is its carrying capacity, D is the delivery rate of T-cells per tumour cell (proportional to delivery per volume), a is the probability of T-cell anergisation after each kill it makes, k_0 is the kill-rate of a T-cell when tumour cells are not limiting, and C_* is a Michaelis-Menten saturation factor preventing the kill function κ becoming infinite as $C \to \infty$. The steady states of this equation occur when $C = 0$ and any number of T-cells, or $C = K(1 - \frac{D}{ra})$ and $T = \frac{D(C+C_*)}{ak_0}$. It can be shown that the $C = 0$ steady state, C_-, is stable if $D > ra$. The eigenvalues λ of the other steady state, C_+, are given by the following equation:

$$\lambda^2 - \lambda\left(\frac{a\kappa}{T} + r(\frac{2C}{K} - 1) + \frac{DC_*}{a(C+C_*)}\right) + \frac{\kappa}{T}\left(D + ra(\frac{2C}{K} - 1)\right) \tag{3.A.1}$$

The eigenvalues are both negative and C_+ is stable if their product is positive and their sum is negative. Their product is positive if,

$$D + ra(\frac{2C}{K} - 1) > 0,$$

implying that $ra > D$. Their sum is negative if,

$$C = K\left(1 - \frac{D}{ra}\right) > C_* r\left(ak_0 + \frac{2rC_*}{K} - r\right)^{-1}\left(1 - \frac{D}{ra}\right).$$

Assuming that $K \gg C_*$ for convenience (which is always true for reasonable parameter values), this also reduces to $ra > D$ after considering all possible cases. Hence, we find that the tumour is eliminated if $D > ra$ and persists otherwise. This may be related to the delivery per unit volume (for parity with the immune trafficking model) as follows,

$$DC = D(\tilde{V}_{\text{tmr}}\eta) = (D\eta)\tilde{V}_{\text{tmr}}, \tag{3.A.2}$$

where η is the number density of tumour cells and $\tilde{V}_{\text{tmr}}$ is the total tumour volume. Defining the delivery rate per volume $D' = D\eta$ and the volume of a tumour cell V_{cell}, the condition for tumour elimination becomes,

$$\begin{aligned} D' &> ra\eta, \text{ or} \\ D' &> \frac{ra}{V_{\text{cell}}}, \end{aligned} \tag{3.A.3}$$

which is the relation given in the main text (equation 3.16) up to a volume factor, with $a = 0.1$ and $V_{\text{cell}} = \frac{4\pi}{3}(10\mu\text{m})^3$. The values chosen for k_0 and a are such that tumour cells kill at a rate of ten per two days [41, 42] and an approximate T-cell lifespan of two days.

Chapter 4

Quantification of temporal lymphocyte localisation

4.0 Introduction



Figure 4.0. Three example slides of tumour histochemistry with or without $CD8^+$ infiltration [244]. T-cells are stained in brown. The left shows a 'hot tumour' with T-cell infiltration and an apparent ongoing immune response. The middle shows a 'cold tumour', or more specifically an 'immune desert' that has $CD8^+$ T-cells present, but which are excluded and confined to the tumour periphery. The right shows a cold tumour that is 'immune ignored', because no T-cells have been delivered to the tumour at all. Figure reproduced with permission [244].

In this chapter, I will describe how the model in Chapter 3 can be extended and utilised to describe the localisation of lymphocytes in the body as a function of time. Chapter 3 considered the maximum delivery rate of T-cells per volume to different organs and tumours. This delivery rate can be considered to be the rate of tumour infiltration in the case that all delivered cells extravasate, *i.e.* where it is not T-cell extravasation (step 5 of the cancer-immunity cycle) but delivery (step 4) that is the bottleneck. The opposite situation would be one in which T-cells arrive in the tumour bearing organ and are attracted to the tumour periphery, but are unable to extravasate into the tumour parenchyma – a so-called 'immune desert'. This example is presented in Figure 4.0. By finding how the probability of T-cell extravasation differs with properties of T-cells or organs, we aim to better understand

the situations in which T-cell extravasation impedes tumour clearance, rather than T-cell delivery. To proceed, we aim to estimate approximate values of the rates of T-cell extravasation and return to and from different organs by fitting to T-cell localisation data. Previous authors have fit T-cell localisation data to ODE systems [81–83], but they did not fit to lymph node data, which is one of the most important organs for T-cells. Furthermore, several of these published models were fit to data over much longer timescales, which have very different dynamics. Finally, some of these data have transient features not captured by the ODE fits. We have acquired lymphocyte localisation data in rats from the literature [37] that includes lymph node localisation data. I will describe a suitable methodology to fit the model to these data and demonstrate that fits with the ODE system are inadequate. After considering various techniques (Section 4.2.2), I will argue the case for extending the ODEs into age-structured PDEs, also known as 'lifespan' models. There is extensive previous work that applies age-structured PDEs to biological contexts [87, 88], but we have not seen one used to address the organ-specific localisation of cells in the context of immune trafficking. I will demonstrate that such a model is fit for purpose, discuss the difficulties in its use and outline potential and planned future work.

4.1 Methods

4.1.0 Lymphocyte localisation data

We acquired lymphocyte localisation data from the literature, originally published by Smith *et al* (1980) [37]. The authors introduced radiolabeled lymphocytes into rats and recorded the radioactivity in different organs as a function of time up to 24 hours, which is used as a proxy for lymphocyte localisation. The aim of the study was to determine which factors influence lymphocyte migration patterns. They considered the type/source of antigen that the cells are specific for, the places the cells were activated, the lymph nodes they were extracted from and the size of the cells (whether they were effector cells or not). There are five data sets, labelled:

- ABA (**A**llogen-stimulated **B**rachial and **A**xillary lymph node cells),
- CFA (**C**omplete **F**reund's **A**djuvant-stimulated brachial/axillary lymph node cells),
- MLN (**M**esenteric **L**ymph **N**ode cells),
- TDL (**T**horacic **D**uct **L**ymphoblasts) and
- TDLn (**T**horacic **D**uct **L**ymphocytes – **n**aïve cells).

Note that the authors refer to "lymphoblasts" (large, dividing cells) as distinct from small lymphocytes. The lymphoblasts are a mix of all effector lymphocytes: $CD4^+$ and $CD8^+$

T-cells and B-cells, whilst the small lymphocytes (in the TDLn dataset) are a mix of T- and B- memory and naïve cells. In the following section, I will describe how this data can be used to fix values of the extravasation probabilities e and return rates μ.

4.1.1 Methods to fit model to data

To fit the ODE model defined in Chapter 3 or the PDE model developed in Section 4.1.2 to the data, we fitted values of the extravasation probabilities e and return rates μ. We ran the model(s) with rat parameters up to 24 hours and compared the percentage of simulated lymphocytes in each organ to the percentage of radiation measured in each organ at each time point. We assigned a score to the fit by calculating the sum of the squared differences between model predictions $N(t)$ and the data $\tilde{N}(t)$:

$$S = \sum_i \left(\tilde{N}(t_i) - N(t_i)\right)^2, \tag{4.0}$$

where the t_i are the time points at which the data exists. We varied values of e and μ for each of the organs present in the data to minimise the score. We set all other values of e to zero to represent the fact that no significant amount of radiation was recovered from other organs. Note that the value of μ_o is irrelevant if $e_o = 0$ and organ o is not the lymph nodes.

To find the best values of e and μ, we first attempted to use the *simulated annealing* algorithm in Scipy [118]. This algorithm attempts to find the global minimum of a system by repeatedly jumping to a new 'basin' (random position in parameter space) and finding a local minimum. The score of the new and old local minima are compared. The new one is accepted if it is a lower score, otherwise it is accepted with probability equal to a Boltzmann factor that depends on a user-defined temperature T. After multiple iterations, the algorithm, if successful, finds the global minimum. However, simulated annealing did not produce satisfactory results: many parameter values had not differed from their original values, several biological expectations were not satisfied and manual alterations to parameters could produce better fits, indicating that simulated annealing had only found local minima. With up to 9 organs plus the pulmonary circuit in each dataset, our model requires $10 \times 2 = 20$ parameters. This was the likely source of the difficulties, and indicated that a more powerful approach could be useful, particularly one that can simply quantify the quality of fit of the entire parameter space.

As in Chapter 2, we utilised a Population Monte Carlo (PMC) variant of Approximate Bayesian Computation (ABC) to find the minima of the system. There is an immediate problem with the use of ABC to fit the trafficking model, even with the use of PMC. ABC requires a very large number of simulations to adequately determine a posterior distribution and the systems of equations are somewhat slow to solve. One possible solution, as in

Chapter 2, would be to 'emulate' the system with a trained ensemble of machine learning algorithms. However, even with 9 (out of 21 possible) organs plus the pulmonary circuit, the number of parameters is unmanageably large to adequately traverse a Latin Hypercube. For example, taking 3 different values for each parameter would require $3^{((9+1)\times 2)} \approx 3 \times 10^9$ runs. If each run took one second, this would require 111 years to complete, and 3 values per parameter is not adequate to describe the behaviour of the system. 5 values would require one 1000^{th} of the age of the universe to complete. For the same reason, regardless of the method used to produce model predictions, ABC is not expected to be able to adequately traverse the parameter space without constraints to assist it. To reduce the number of parameters, it would be possible to define two organ compartments: the "target" organ and "other" organs. However, non-linearities would prevent the mapping of predicted parameter values for the target organ in the simplified system to the full system, whether for the ODE system or the PDE system described in Section 4.1.2. The equation for the number of cells in an organ o in the full system is

$$V_o \frac{\mathrm{d}C_o}{\mathrm{d}t} = \sum_p M_{op} C_p,$$

where M is a matrix encoding the system of ODEs. The equation for the concentration in a combined organ o' in the simple system, $c_{o'}$, is

$$v_{o'} \frac{\mathrm{d}c_{o'}}{\mathrm{d}t} = \sum_{p'} m_{o'p'} c_{p'},$$

where

$$v_{o'} c_{o'} = \sum_{i=\sigma_{o'0}}^{\sigma_{o'N}} V_i C_i,$$

and $\sigma_{o'0}, ..., \sigma_{o'N}$ are the indices of the full system organs that make up simplified organ o', which form a sequential list (*e.g.* $\sigma_0 = 0, 1$ and $\sigma_1 = 2, 3, 4$). $v_{o'}$ is an undetermined combination of organ volumes, though a desirable choice would be $v_{o'} = \sum_{i=\sigma_{o'0}}^{\sigma_{o'N}} V_i$. To have

$$v_{o'} \frac{\mathrm{d}(c_{o'})}{\mathrm{d}t} = \sum_{i=\sigma_{o'0}}^{\sigma_{o'N}} V_i \frac{\mathrm{d}(C_i)}{\mathrm{d}t},$$

we require

$$\sum_{p'} m_{o'p'} c_{p'} = \sum_{i=\sigma_{o'0}}^{\sigma_{o'N}} \sum_p M_{ip} C_p$$

$$\sum_{p'} m_{o'p'} \sum_{j=\sigma_{p'0}}^{\sigma_{p'N}} \frac{V_j C_j}{v_{p'}} = \sum_{i=\sigma_{o'0}}^{\sigma_{o'N}} \sum_p M_{ip} C_p$$

By comparing coefficients, one derives conditions where the sum over sections of adjacent rows of the matrix M must have a ratio that depends only on organ volumes, which is not consistent with the form of M. For example, if the second simplified compartment contains organs 2, 3 and 4 from the full system, then we find that $\sum_{j=2}^{4} M_{j0} = \frac{V_0}{V_1} \sum_{j=2}^{4} M_{j1}$. Hence, the system cannot be simplified in such a way that factors of e estimated in one system can be sensibly mapped to the other.

A similar strategy would be to ignore certain parameters, if they can be shown to have no effect on model predictions in a particular organ. We performed a sensitivity analysis on the ODE and PDE models to determine which parameters most affect the localisation of lymphocytes in each organ. The results, shown in Section 4.2.0, indicate that each organ o is sensitive almost solely to its own values of e_o and μ_o, except the lymph node, which is insensitive to its own value of e_{LN}, due to arrival of cells from other organs via the lymphatics. This suggests an alternative strategy for fitting the data. ABC can be used to fit the predicted localisation of lymphocytes to the data of a single organ o at a time, by varying a single pair of parameters e_o and μ_o. Since lymph node localisation depends more on other values of e than its own, e_{LN} and μ_{LN} are fit to the LN data as e_o and μ_o are fit to their data simultaneously, except when the spleen is being considered. When the lung data is being fitted, then e_{BS}, μ_{BS}, e_{PC}, μ_{PC}, e_{LN} and μ_{LN} are all fit to the lymph node and lung data at the same time, where BS is the lung blood supply and PC is the pulmonary circuit. We define a 'pass' as a run through every organ, fitting each pair of parameters e_o and μ_o to the measured localisation in each organ in turn. We repeat this process for several more passes until a final posterior distribution is reached. For the first pass, a uniform prior is taken between $0 \leq e \leq 1$ and $0 \leq \mu \leq 1$ for the ODEs, or $0 \leq \mu \leq 24$ hours for the PDEs (note that μ has a slightly different definition in the PDE model, see Section 4.1.2). For all subsequent passes, Gaussian priors defined by the mean and standard deviation of the previous predicted values are used.

4.1.2 PDE model definition

Fits of the ODE model to the data (see Section 4.2.1) indicate that an extension of the model is required, for example with delay differential equations, multi-compartment ODEs, or PDEs (*e.g.* [245]). We opted to use PDEs for our purposes (see Section 4.2.2). A set of PDEs may be defined with equations analogous to those defined for the ODEs in Section 3.1.0. Obscuring the complexities due to the lymph nodes, liver and spleen, the

basic equations for organ o are as follows:

$$\begin{aligned} V_o \frac{\mathrm{d}C_o}{\mathrm{d}t} =& B_o(C_h - C_o), \\ \tilde{V}_o \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) \tilde{C}_o =& -\gamma_o(a)\tilde{C}_o(a,t), \end{aligned} \tag{4.1}$$

where V_o, B_o and C_o are the blood volume, blood flow and lymphocyte concentration in the vasculature of organ o, and $\tilde{V}_o$ and $\tilde{C}_o$ are the organ volume and interstitial concentration, as before, and once more using the total organ volume for $\tilde{V}_o$ rather than the explicit interstitial volume $\tilde{V}_o - V_o$. The new independent variable a is the age of a cell within an interstitial compartment, *i.e.* the amount of time they have spent there. The function $\gamma_o(a)$ is a new term for the propensity of lymphocytes of age a to leave the interstitial space of organ o, defined as,

$$\gamma_o(a) = \tilde{V}_o \frac{f_o(a)}{1 - \int_0^a \mathrm{d}a' f_o(a')}, \tag{4.2}$$

where $f_o(a)$ is the probability density for a lymphocyte of age a to leave organ o. A function of this form is well-studied and is utilised in previous age-structured models (*e.g.* [89]). We define $f_o(a)$ as a (truncated) Gaussian with mean μ_o and standard deviation $\frac{\mu_o}{2}$, where μ_o is the mean residence time of a lymphocyte within organ o, distinct from its similar definition in the ODE model. The Gaussian is truncated at zero age, which introduces exponentially small effects on the probability densities elsewhere. The standard deviation governs the rate of decay of oscillations in lymphocyte concentration in each organ; if it is set to zero, then a bolus of lymphocytes that enters a compartment together will also exit together as a group. As we do not have highly resolved temporal data to quantify the decay of these oscillations, we fixed the value to be proportional to the mean, to reduce the number of parameters that must be fitted. The boundary conditions for the PDE are,

$$\begin{aligned} \tilde{V}_o \tilde{C}_o(0,t) =& e_o B_o C_o(t) \\ \tilde{V}_o \tilde{C}_o(a,0) =& \begin{cases} e_o B_o C_o(0), & a = 0 \\ 0, & a \neq 0 \end{cases}, \end{aligned} \tag{4.3}$$

where e_o is the extravasation probability for lymphocytes passing through the vasculature of organ o. The ODE for the heart (without complexities due to liver, spleen, lung or lymph nodes shown) is,

$$V_h \frac{\mathrm{d}C_h}{\mathrm{d}t} = \sum_o B_o(C_o(1 - e_o) - C_h) + \int_0^\infty \mathrm{d}a\, \gamma_o(a)\tilde{C}_o(a,t). \tag{4.4}$$

The extension of these equations to include the details of the liver, lymph nodes, pulmonary circuit and spleen is analogous to that for the ODE system.

4.1.3 Organ exit propensity γ

The definition of $\gamma_o(a)$ can be understood as follows. The leading $\tilde{V}_o$ is required for dimensional consistency. Dropping $\tilde{V}_o$ and the subscript o for convenience, the remaining fraction represents the probability that a cell of age a leaves the organ given that is has not left already:

$$\begin{aligned} p\Big(\text{Leave in } [a, a+\mathrm{d}a) \,\Big|\, \text{Not left by } a\Big) =& \frac{p\Big(\text{Leave in } [a, a+\mathrm{d}a)\Big)}{p(\text{Not left by } a)} \\ =& \frac{f(a)\mathrm{d}a}{1-\int_0^a \mathrm{d}a' f(a')} \end{aligned} \tag{4.5}$$

Dividing by da gives the 'probability' density that defines γ. Since $f(a)$ is a normalised probability density function, its integral approaches one as age tends to infinity. The integral of $\gamma(a)$ can be calculated by defining $I(a) = 1 - \int_0^a \mathrm{d}a' f(a')$ such that,

$$\int_0^\infty \gamma(a)\mathrm{d}a = -\int_0^\infty \mathrm{d}a \frac{\mathrm{d}I/\mathrm{d}a}{I} = -\big[\ln(I)\big]_{a=0}^{a=\infty} = \infty. \tag{4.6}$$

This shows that γ is itself *not* a probability density but a propensity (density), and that the longer that cells remain in the organ, the more likely they are to leave in the next computational timestep. Note that this result is independent of the form of $f(a)$, as long as it is a normalised probability density function.

4.1.4 Numerical algorithms to solve PDEs

Much of the discussion in this chapter will focus upon numerical issues, so we will begin by outlining potential numerical algorithms. There are a number of schemes that could be used to solve this PDE system. To simplify this section, I will illustrate them with the advection equation, $\partial_t u = -s\partial_x u$, where s is the velocity of advection, following arguments by LeVeque [246]. The simplest of schemes is the forward-time Euler scheme, with equation $u_j^{n+1} = u_j^n - \frac{s\Delta t}{2\Delta x}(u_{j+1}^n - u_{j-1}^n)$, where subscripts j indicate discretised steps in space for the solution, and superscripts n indicate discretised steps in time, Δt is the timestep and Δx is the space step. This can be expressed as a matrix update step,

$$u^{n+1} = u^n + \Delta t \mathbf{A} u^n, \tag{4.7}$$

by defining the matrix $\mathbf{A}$ with the stencil $\frac{s}{2\Delta x}(1, 0, -1)$. A stencil of length three indicates the values to be placed into the matrix left of, on and right of the main diagonal. This stencil results in,

$$\mathbf{A} = \frac{s}{2\Delta x}\begin{bmatrix} 0 & -1 & \dots & & 1 \\ 1 & 0 & -1 & \dots & \\ 0 & 1 & 0 & -1 & \dots \\ & & & & \ddots \end{bmatrix}. \tag{4.8}$$

Scheme	Update equation, $u_j^{n+1} =$	Stencil
Forward Euler	$u_j^n - \frac{s\Delta t}{2\Delta x}(u_{j+1}^n - u_{j-1}^n)$	$\frac{s}{2\Delta x}(1, 0, -1)$
Leapfrog	$u_j^{n-1} - \frac{s\Delta t}{\Delta x}(u_{j+1}^n - u_{j-1}^n)$	$\frac{s}{\Delta x}(1, 0, -1)$
Lax-Friedrichs	$u_j^n + \frac{1}{2}(u_{j-1}^n - 2u_j^n + u_{j+1}^n)$ $-\frac{s\Delta t}{2\Delta x}(u_{j+1}^n - u_{j-1}^n)$	$\frac{1}{2\Delta t}(1, -2, 1)$ and $\frac{s}{2\Delta x}(1, 0, -1)$
Lax-Wendroff	$u_j^n + \frac{s^2\Delta t}{2\Delta x}(u_{j-1}^n - 2u_j^n + u_{j+1}^n)$ $-\frac{s\Delta t}{2\Delta x}(u_{j+1}^n - u_{j-1}^n)$	$\frac{s^2\Delta t}{2\Delta x^2}(1, -2, 1)$ and $\frac{s}{2\Delta x}(1, 0, -1)$
Upwind	$u_j^n - \frac{s\Delta t}{\Delta x}(u_j^n - u_{j-1}^n)$	$\frac{s}{\Delta x}(1, -1, 0)$
Beam-warming	$u_j^n - \frac{s\Delta t}{2\Delta x}(3u_j^n - 4u_{j-1}^n + u_{j-2}^n)$ $+\frac{s^2\Delta t^2}{2\Delta x^2}(u_j^n - 2u_{j-1}^n + u_{j-2}^n)$	$\frac{-s}{2\Delta x}(1, -4, 3, 0, 0)$ and $\frac{s^2\Delta t}{2\Delta x^2}(1, -2, 1, 0, 0)$

Table 4.0. A list of several schemes for the advection equation, $\partial_t u = -s\partial_x u$. Both the update equation for the discretised solution with timestep Δt and space-step Δx and the 'stencil' (see main text) used to convert the equation into matrix form are shown. Lower and upper indices j and t are respectively for space and time. The upwind scheme is given assuming $\frac{\Delta t}{\Delta x} > 0$.

Scheme	Stencil	
Forward Euler	$\frac{1}{2\Delta a}(1, 0, -1)$	and $\frac{\gamma}{\tilde{V}}(0, -1, 0)$
Leapfrog	$\frac{1}{\Delta a}(1, 0, -1)$	and $\frac{\gamma}{\tilde{V}}(0, -1, 0)$
Lax-Friedrichs	$\frac{1}{2\Delta a}(1, -2, 1)$	and $\frac{1}{2\Delta a}(1, 0, -1)$ and $\frac{\gamma}{\tilde{V}}(0, -1, 0)$
Upwind	$\frac{1}{\Delta a}(1, -1, 0)$	and $\frac{\gamma}{\tilde{V}}(0, -1, 0)$

Table 4.1. Update stencils for the PDE organ system, equation 4.9, analogous to those defined for the advection equation in Table 4.0. Note that $\Delta a = \Delta t$.

Various other solutions can be defined analogously; their update equations and corresponding stencils are given in Table 4.0. Applied to the equation $\tilde{V}\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right)\tilde{C} = -\gamma(a)\tilde{C}(a,t)$, where the organ subscript o has been dropped for convenience, we can define the forward-Euler scheme:

$$\frac{\tilde{C}_j^{n+1} - \tilde{C}_j^n}{\Delta t} + \frac{\tilde{C}_{j+1}^n - \tilde{C}_{j-1}^n}{2\Delta a} = \frac{1}{\tilde{V}}(eBC_*^n - \gamma_j \tilde{C}_j^n), \tag{4.9}$$

where C_*^n is C^n if $j = 0$ and 0 otherwise, to provide the initial condition $\tilde{C}_0^n = eBC^n$. C^n has no spatial (age) subscript because it is the vascular concentration of lymphocytes, $C(t)$. Rearranging the forward-Euler update equation,

$$\tilde{C}_j^{n+1} = \tilde{C}_j^n - \frac{\Delta t}{2\Delta a}(\tilde{C}_{j+1}^n - \tilde{C}_{j-1}^n) + \frac{\Delta t}{\tilde{V}}(eBC_*^n - \gamma_j \tilde{C}_j^n), \tag{4.10}$$

we can define stencils $\frac{-1}{2\Delta a}(-1, 0, 1)$ and $\frac{-\gamma}{\tilde{V}}(0, 1, 0)$. Stencils for the Leapfrog, Lax-Friedrich and Upwind schemes are displayed in Table 4.1, but Lax-Wendroff and Beam-warming are omitted, because they are complicated expressions that require derivatives of the boundary

conditions and multiple stencils of length greater than 3. Note that the timestep and "age-step" are the same, $\Delta a = \Delta t$, because age is simply time. Hence, the Lax-Friedrichs scheme is identical to Upwind in this case (after the two distinct stencils for Lax-Friedrichs are added together).

4.1.5 Method of characteristics

The simple form of the interstitial PDE and its boundary condition allows its solution to be expressed as a characteristic integral. Clearly, if $\gamma(a) = 0$ for all a, then the number of cells in the interstitial space of an organ is the integral over the boundary condition since the beginning of time:

$$\tilde{V}\tilde{C}(t) = \int_0^t \mathrm{d}t'\, eBC(t').$$

With $\gamma(a) \neq 0$, we can see from Equation 4.1 that characterising data-curves by s yields $\frac{\mathrm{d}t}{\mathrm{d}s} = 1$, $\frac{\mathrm{d}a}{\mathrm{d}s} = 1$ and $\frac{\mathrm{d}z}{\mathrm{d}s} = -\frac{\gamma(a)z(a,t)}{\tilde{V}}$, where z is some function describing the initial data (boundary conditions). We characterise with $s = a$ and let $t = t_0 + a$. The boundary condition hence propagates along age-lines, degraded by $\gamma(a)$. Integrating the equation for z along characteristic curves yields,

$$z(a,t) = \tilde{C}(a,t) = \frac{eBC(t_0)}{\tilde{V}} e^{-\int_0^a ds\gamma(s)/\tilde{V}}$$

where $t_0 = t - a > 0$ and $\tilde{C}(a,t) = 0$ for $t < a$. To calculate the total number of cells at time t in the interstitial space, we define $t' = \min(a_{\max}, t)$, where $a_{\max}$ is a large age at which the Gaussian is truncated again, so that the PDE solution can be numerically defined with a finite amount of memory in a non-sparse array. The total number of cells is,

$$\tilde{V}\tilde{C}(t) \;= \int_0^{t'} \mathrm{d}a\, \tilde{V}\tilde{C}(a,t) \;= \int_0^{t'} \mathrm{d}a\, eBC(t_0) e^{-\int_0^a \mathrm{d}a'\, \gamma(a')/\tilde{V}}.$$

This formulation simplifies the calculation of the number of cells in the interstitial compartment at times t not equal to the current computational timestep. The form of $\gamma(a)$ makes this calculation simpler than it appears at first sight, as the exponential can be pre-calculated and tabulated using error functions. From the fact that

$$\gamma(a)/\tilde{V} = \frac{-F'(a)}{F(a)},$$

where $F(a) = 1 - \int_0^a \mathrm{d}a' f(a')$ for some probability distribution $f(a)$, it follows that

$$e^{-\int_0^a \mathrm{d}a'\gamma(a)/\tilde{V}} = F(a).$$

If $f(a)$ is assumed to be a Gaussian with standard deviation $\sigma = \mu/2$, then

$$F(a) = 1 - \frac{1}{\eta}\int_0^a \mathrm{d}a' e^{-\frac{2(a'-\mu)^2}{\mu^2}},$$

where η is the normalisation of the Gaussian. Assuming the Gaussian is truncated at zero and $a_{\max}$ and renormalised, simple algebraic manipulation then yields a result in terms of the error function, erf:

$$F(a) = e^{-\int_0^a \mathrm{d}a' \gamma(a)/\tilde{V}} = \frac{\mathrm{erf}(\sqrt{2}\frac{a_{\max}-\mu}{\mu}) - \mathrm{erf}(\sqrt{2}\frac{a-\mu}{\mu})}{\mathrm{erf}(\sqrt{2}\frac{a_{\max}-\mu}{\mu}) + \mathrm{erf}(\sqrt{2})}, \tag{4.11}$$

This expression can be pre-calculated for a desired $a_{\max}$ and tabulated. Thus, the equation for $\tilde{C}(t)$ requires only a single integral of the product of the boundary condition history and terms from the error function table.

4.1.6 Combined ODE and characteristic solver

	$a_{*,1}$	$a_{*,2}$	$a_{*,3}$	$a_{*,4}$	$a_{*,5}$	$a_{*,6}$	$a_{*,7}$
k_1							
k_2	1/5						
k_3	3/40	9/40					
k_4	44/45	-56/15	32/9				
k_5	19372/6561	-25360/2187	64448/6561	-212/729			
k_6	9017/3168	-355/33	46732/5247	49/176	-5103/18656		
k_7	35/384	0	500/1113	125/192	-2187/6784	11/84	
N^{i+1}	35/384	0	500/1113	125/192	-2187/6784	11/84	
$\hat{N}^{i+1}$	5179/57600	0	7571/16695	393/640	−92097/339200	187/2100	1/40

Table 4.2. Butcher Table for the Dormand-Prince method [247]. The row is the term to be calculated and the column gives the coefficient of previously calculated k terms that must be used in its calculation. For example, for a system in which $\frac{\mathrm{d}N}{\mathrm{d}t} = f(N_i)$,
$k_4 = f\left(N_i + \Delta t\left(a_{41}k_1 + a_{42}k_2 + a_{43}k_3\right)\right) = f\left(N_i + \Delta t\left(44/45\ k_1 - 56/15\ k_2 + 32/9\ k_3\right)\right)$. Note that the time of evaluation has been omitted because our equations are not a function of this variable. k_1 has no entries in the table because it is simply $f(N_i)$. N^{i+1} and $\hat{N}^{i+1}$ are two estimates of the model's solution at timestep t_{i+1}. The first may be used alone for a fifth order solution, and the two may be used together to estimate the current error in the solution.

The solution of both the ODEs and the PDEs defined in Section 4.1.2 require integrals. Specifically, calculating the content of the interstitial compartments requires an integral over the content of the vascular compartment at previous times,

$$\tilde{C}(t) \propto \int_0^{\min(t,a_{\max})} \mathrm{d}a\, C(t-a) e^{-\int_0^a \mathrm{d}a'\, \gamma(a')/\tilde{V}},$$

and this integral is needed again to calculate the number of cells flowing from the PDEs to the heart, liver or lymph node ODEs. This presents several problems. It would be inefficient and undesirable for this integral to be re-calculated at every intermediate step of the Runge-Kutta algorithm, particularly if the intermediate steps are at a finer resolution than the values of $\gamma(a)$ are tabulated at. The values of the blood concentration within an organ, C, that are required within the integrand present a further problem. If the ODEs are

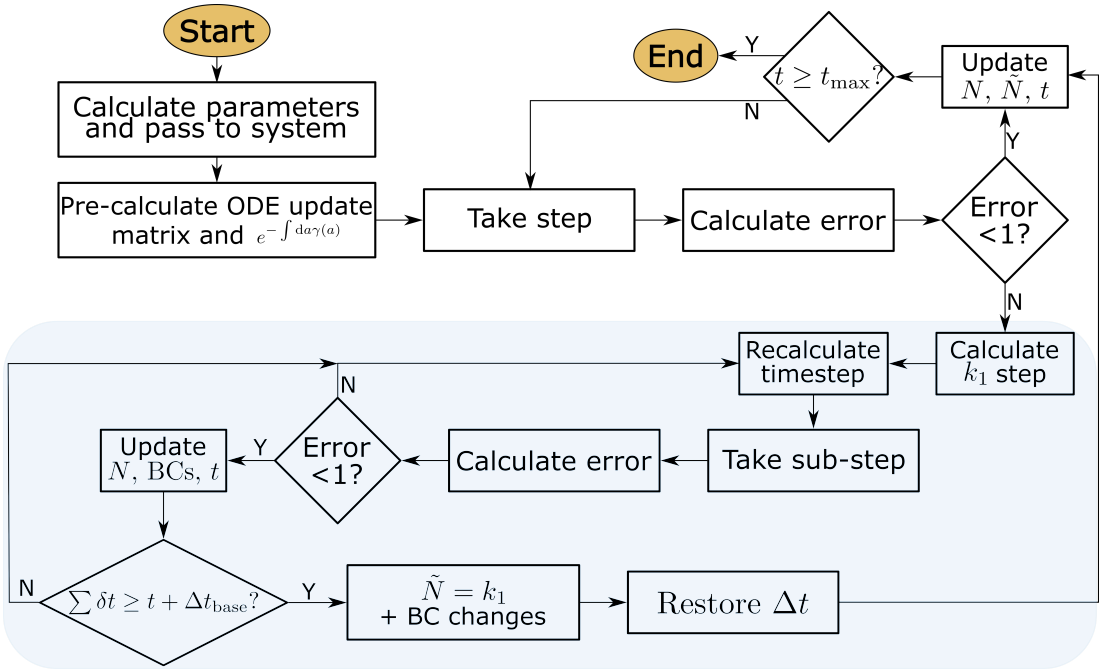


Figure 4.1. A flowchart describing the program flow of the custom ODE/PDE solver, explained in Section 4.1.6. The blue box is a visual aid for the part of the algorithm invoked when the error from the current timestep is large enough that a dynamic time-stepper is required.

calculated over multiple intermediate or even reduced timesteps (due to a dynamic time-stepper), then the number of cells leaving the ODE compartment between t and $t + \Delta t$ is not equal to $eBC(t)\Delta t$, but depends on smaller timesteps and intermediate terms from the Runge-Kutta algorithm. In this case, choosing linearly distributed coarse values of C for the boundary conditions of the PDEs and the related integral would numerically result in a loss of conservation of mass. These two problems together mean that off-the-shelf solvers such as Python's *dopri5* implementation or MATLAB's equivalent *ode45* are not suitable. Instead, we had to develop a custom ODE/PDE solver, whose calculation steps are presented in Figure 4.1. An explanation for each step in the figure is given below.

Calculate parameters and pass to system: Blood flow and volume parameters, values of e, μ, initial values of t, *etc* are set before the solver is launched.

Pre-calculate ODE update matrix and $e^{-\int_0^a da'\gamma(a')/\tilde{V}}$: The core calculation in the vascular ODE update step is

$$N(t + \Delta t) = MN(t), \tag{4.12}$$

where M is a matrix that encodes all of the ODEs and can be simply calculated. The exponential $e^{-\int_0^a da'\gamma(a')/\tilde{V}}$ can be calculated using error functions, as described in Section 4.1.5.

Take step: A single step consists of a Dormand-Prince update step [247], using 7 evaluations of the core step function. The core step function updates the ODEs using

equation 4.12 and calculates the current PDE solution. The number of cells ('NOC') in the interstitial space at timestep t^i is given by

$$(NOC)^i = \sum_{a=0}^{\min(t^i, a_{\max})} eBC(t^i - a)e^{-\int_0^a \mathrm{d}a'\, \gamma(a')/\tilde{V}},$$

where $C(t)$ here is calculated from the 'history' of the ODE solution, $H(t) := VC(t)$. The number of cells leaving the PDE compartments during the i^{th} timestep is thus

$$(CL)^i = (NOC)^i - \left((NOC)^{i-1} + eBC(t_i)\Delta t\right).$$

The number of cells in the interstitial space and the number that leave in the current timestep are calculated by integration *only for the first Runge-Kutta step* k_1. The number of cells leaving the PDE compartments during the i^{th} timestep, $(CL)^i$, is recorded and cells that flow to the LN PDE are added only after this first evaluation. This is because newly extravasated cells cannot leave on a timescale less than Δa (since numerically, $\gamma(a) = 0$ for $0 \leq a < \Delta a$), and all evaluations after k_1 are at intervals smaller than $\Delta t = \Delta a$ by definition.

In every following intermediate (Dormand-Prince) Runge-Kutta step, the cells leaving the LN, pulmonary circuit and spleen are added to the liver (for the spleen) or heart (for the LN and pulmonary circuit) ODEs. The ODE system is updated and $(NOC)^i$ is updated using the relevant Dormand-Prince update equation. For example, after step 1, the intermediate solution value for the interstitial (PDE) compartments $\tilde{k}_1 = (NOC)^i + \frac{eB}{V}\Delta t(a_{21}k_1)$, where $\tilde{k_1}$ denotes only the interstitial part of the solution, k_1 is the vascular (ODE) solution calculated from the matrix of ODEs described earlier, and the a_{xx} terms are the coefficients of the intermediate steps of the Dormand-Prince steps, given in Table 4.2. After step 2, $\tilde{k}_2 = (NOC)^i + \frac{eB}{V}\Delta t(a_{31}k_1 + a_{32}k_2)$. However, note again that $(CL)^i$ is *not* updated. After all seven evaluations, the ODE solution is set to $N + \Delta t \sum_{j=1}^{6} a_{7j}k_j$, the PDE solution is set to $(NOC)^i + \frac{eB}{V}\Delta t \sum_{j=1}^{6} a_{7j}k_j$, and the "history" of boundary conditions to be used when calculating the number of cells in the interstitial space NOC is set to $H^i = \sum_{j=1}^{6} a_{7j}\left(N(t^i) + \Delta t \sum_{l=1}^{j-1} a_{jl}k_l\right)$, to ensure that the mass leaving the ODE is equal to the mass entering the PDE.

Calculate error: Dormand and Prince's algorithm includes a seventh evaluation of the step function which is not necessary for the calculation of N^{i+1}. This can be used to calculate a second estimate of the solution at the next timestep, $\hat{N}^{i+1}$. These two can be used to estimate the scalar error on the new solution estimate. Following previous analysis [247–249], we define $\epsilon_o^i = N_o^i - \hat{N}_o^i$ and $Sc_o^i = \text{Atol} + \text{Rtol} \times \max(|N_o^i|, |\hat{N}_o^i|)$, where Atol and Rtol are the absolute and relative solution tolerances and the subscript o indicates each

equation in the ODE/PDE system (or equivalently, each organ compartment). Then, the scalar error (relative to the tolerances) at timestep i is $ScE^i \approx \sqrt{\sum_o (\frac{\epsilon_o^i}{Sc_o^i})^2 / n_d}$, where n_d is the number of equations (compartments) in the ODE/PDE system. If the error is less than 1, indicating that the error is less than both the absolute and relative tolerances, then the solution proceeds to the next timestep. If not, then the timestep is adjusted adaptively, as follows:

Calculate k_1 step: The PDE solution for the timestep requires an integral that can be no finer than Δa. When finer timesteps are nonetheless required, the k_1 step is performed with the base Δt to obtain the number of cells leaving the PDE over that timestep and the number of cells in the PDE at the start of the timestep.

Recalculate timestep: In order to keep the scalar error in the solution below 1 (*i.e.* smaller than the allowed tolerances) and the solution within allowed values (*i.e.* positive numbers of cells), finer timesteps can be chosen. The new sub-timestep is $\delta t = \Delta t S (ScE)^E$, where $S = 0.9$ is a safety factor and $E = 0.2$ is an arbitrary exponent. After this calculation, two additional checks are performed: if $\delta_t > \Delta t$, then the sub-timestep is set equal to Δt. If the sum of all previous substeps plus the proposed step $\sum \delta t_{\text{prev}} + \delta t_{\text{new}} > \Delta t$, then δt_{new} is set to $\Delta t - \sum \delta t_{\text{prev}}$, because the solution must be evaluated at intervals of the original timestep.

Take sub-step: This is the same as a normal step, but the integral for the PDE is not calculated and the solution from k_1 is used. After each sub-timestep, the change to N is used to alter the boundary condition history,

$$H^i \mathrel{+}= \frac{\delta t}{\Delta t} \left(\left(\sum_{j=1}^{6} a_{7j} \left(N + \Delta t \sum_{l=1}^{j-1} a_{jl} k_l \right) \right) [t' + \delta t] - \left(\sum_{j=1}^{6} a_{7j} \left(N + \Delta t \sum_{l=1}^{j-1} a_{jl} k_l \right) \right) [t'] \right),$$

where t' is the time at the start of the sub-timestep and $+=$ denotes accumulation of H^i.

$\tilde{N} = \tilde{k}_1$ + BC changes: After all sub-timesteps have been calculated, the ODE solution and the boundary condition (BC) history have been updated, but the PDEs have not been. Their values are set to the number of cells present in the PDEs at the start of the timestep according to $\tilde{k}_1$, plus small changes due to changes in the boundary conditions, *i.e.* $\tilde{N}^{i+1} = \tilde{k}_1 + \frac{eB}{V} \Delta t (H^i - N^i)$.

4.2 Results

4.2.0 Sensitivity of steady state of ODE system

To find a strategy to utilise Approximate Bayesian Computation (ABC) to fit our models to data, we performed a sensitivity analysis of the steady state of the ODE system to values

	Stom	SInt	Sple	Live	LgBS	Skin	Skel	LNs	Thym	LgPC
Self-e	0.37	0.43	0.36	0.32	0.05	0.35	0.38	0.00	0.32	0.37
Self-μ	0.39	0.43	0.40	0.39	0.08	0.34	0.38	0.54	0.29	0.39
Other-e	0.13	0.07	0.13	0.15	0.43	0.16	0.13	0.28	0.20	0.13
Other-μ	0.10	0.07	0.10	0.13	0.44	0.14	0.11	0.18	0.19	0.11
Self-sum	0.76	0.86	0.77	0.71	0.13	0.70	0.76	0.54	0.61	0.76
Other-sum	0.24	0.14	0.23	0.29	0.87	0.30	0.24	0.46	0.39	0.24

Table 4.3. Sensitivity analysis for the steady state of some of the organs in the ODE system with respect to changing the values of the extravasation probabilities e and return rates μ, as described in Section 4.2.0. Importance indices obtained from the Random Forest technique for the organ's own values of e and μ and the sum of the indices for all other organs' values of e and μ are shown in the top part of the table. The sum of the self-indices and other-indices are shown in the bottom part of the table. For most organs, they are considerably more sensitive to their own parameter values than any other, except the lymph nodes and lung blood supply.

of e and μ. A set of values of e and μ that yield a reasonable fit to the ABA dataset were chosen, and all values of $e = 0.0$ were changed to 0.1. Then, 1700 random sets of parameter values were generated, with all parameter values varying between 90% and 110% of their default values. The random forest technique was used to determine the sensitivity of the steady-state localisation in each organ to all other organ parameters. A summary of a subset of the results is presented in Table 4.3. The results show that organs other than the lymph nodes and lung blood supply are more sensitive to their own values of e and μ than to others. Results for the PDE system are similar (not shown). We hence fit these parameters by considering the localisation data for a single organ at a time. Localisation in the lymph node is more sensitive to other organs' values of e than to e_{LN}, so e_o, μ_o, e_{LN} and μ_{LN} are fit to the data for organ o and the lymph node together in each case. Cells draining the spleen return to the blood rather than the lymph nodes, so this is the only organ for which e_o and μ_o are fit alone. The lung has two sets of parameters – the blood supply and pulmonary circuit – and all four are fit simultaneously to the lung localisation data, along with the lymph node parameters and data.

4.2.1 ODE fit to localisation data

We attempted to fit the ODE model to Smith *et al*'s data using Approximate Bayesian Computation (ABC), using the procedure described in Section 4.1.1 and scoring fits by the sum of the squared differences between the data and model predictions. This yields both the density plots from ABC and parameter values predicted to yield a good fit (minimum score). An example of the density plots obtained is shown in Figure 4.2, for the TDL dataset. For this dataset, ABC indicates that only particular values of e and μ for the spleen and lymph nodes provide good fits to the data, there is a line of appropriate values of e and μ for the liver, skin and skeleton that provide good fits, any value of e but a constrained value

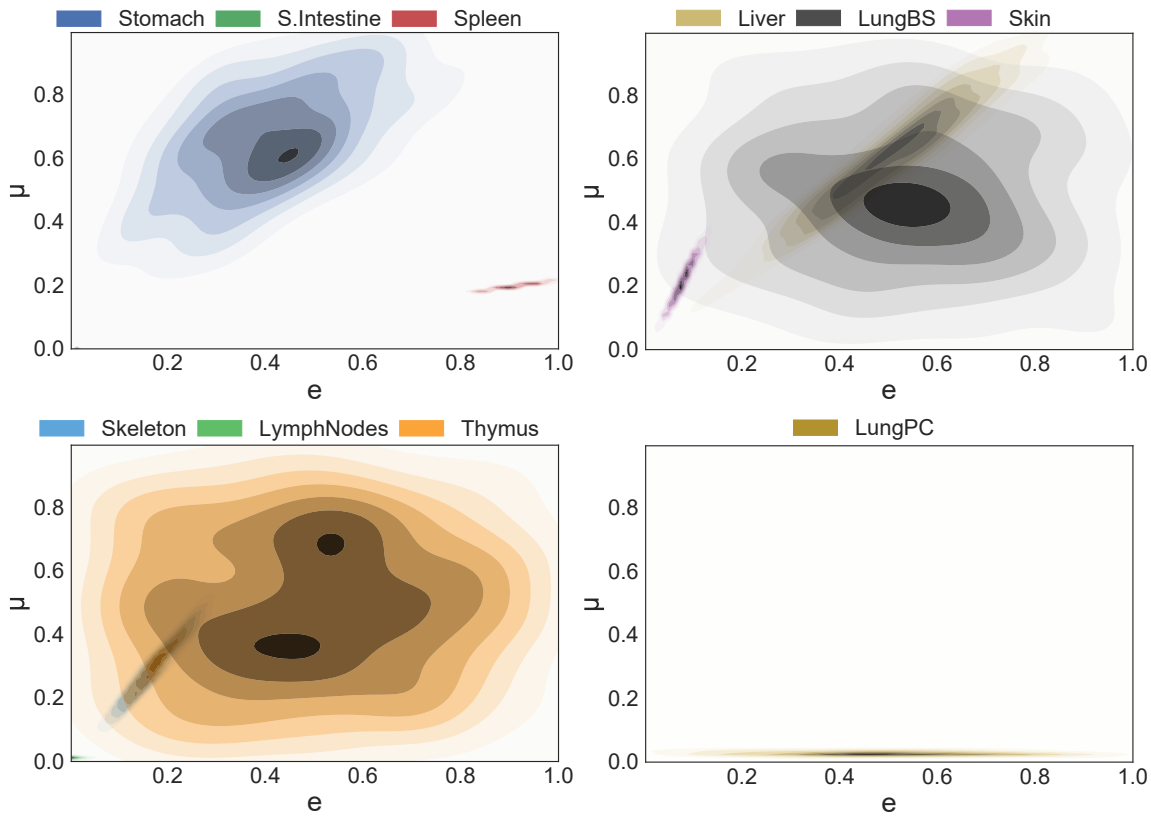


Figure 4.2. Posterior distribution for parameter values that cause the ODE model to fit the TDL dataset, obtained from Approximate Bayesian Computation (ABC). Each panel shows the return rate μ versus the extravasation probability e for up to three parameters marked in the legend. Parameters such as the spleen have very constrained parameter values for the fit, but others are much more uncertain.

of μ provides a good fit for the lung pulmonary circuit, a large region of parameter space provides an equally good fit for the stomach, and the values chosen for the thymus and lung blood supply are irrelevant. Lung blood supply parameters are undetermined because the parameters for the pulmonary circuit are more important, whilst thymus parameter values are undetermined because the bloodflow and the percentage of radiation recovered in the thymus are very low, so the probability of extravasation to the thymus makes no difference to the fit compared to other parameters. The best fit from the density plot for each data set, selected either from the point of maximum density or the mean parameter values (whichever was a better fit), are presented in Figure 4.3 for the ODE model. The TDL dataset is one of only two with a reasonable fit, though it fails to capture the initial transient in the lung. Other datasets are much poorer: for example, the fit to TDLn does not capture the transient localisation in the spleen, nor the gradual build-up in the lymph nodes. This suggests that an extension to the ODE model is needed, to take into account delays or persistence of cells within organs.

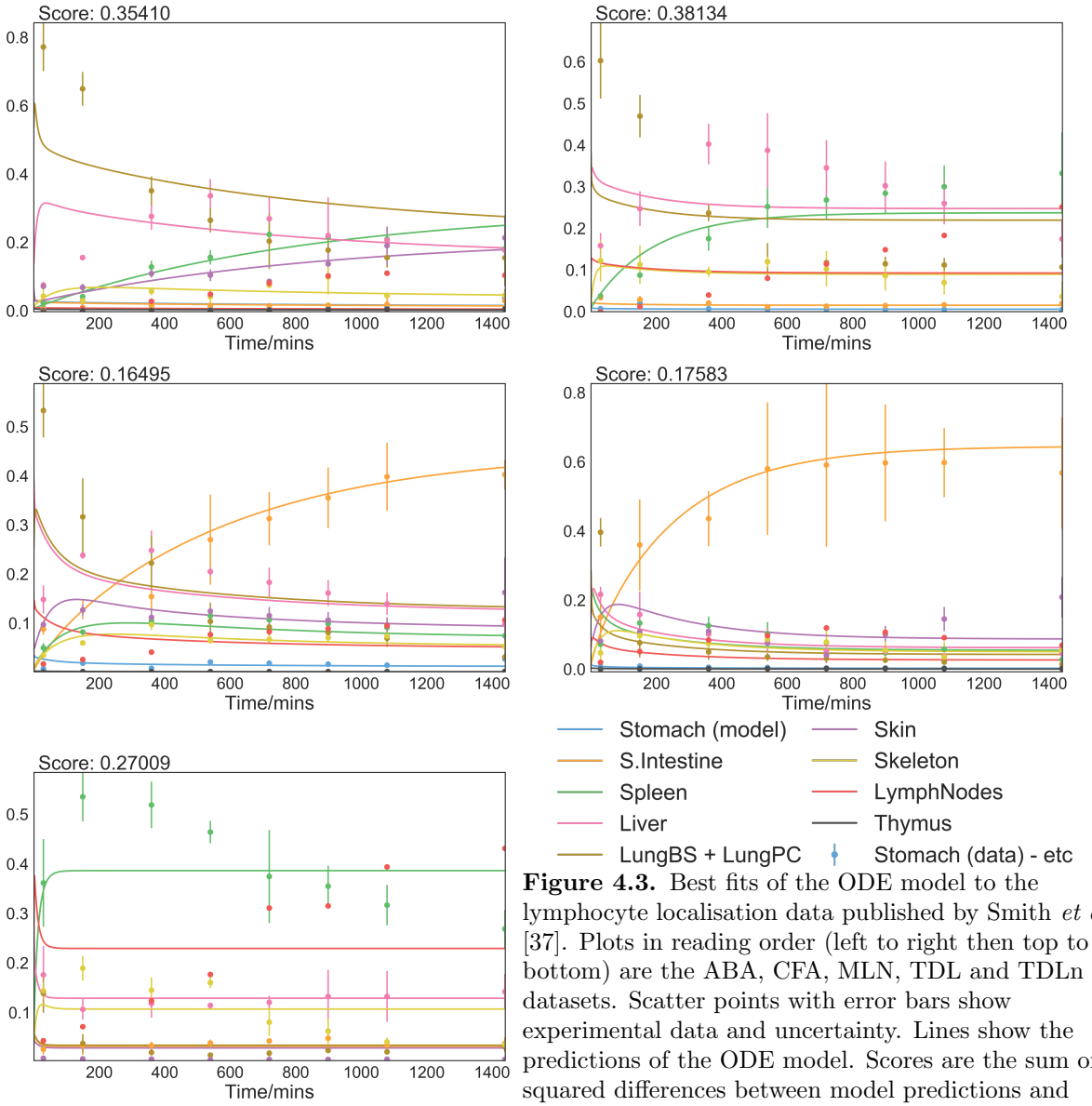


Figure 4.3. Best fits of the ODE model to the lymphocyte localisation data published by Smith *et al* [37]. Plots in reading order (left to right then top to bottom) are the ABA, CFA, MLN, TDL and TDLn datasets. Scatter points with error bars show experimental data and uncertainty. Lines show the predictions of the ODE model. Scores are the sum of squared differences between model predictions and the data at each time-point.

4.2.2 Motivation for PDEs

Fits of the ODE model to the data in Section 4.2.1 were poor; the initial transient and the steady state could not be simultaneously fit. This is likely to be because the ODEs cannot model persistence of cells within an organ: the use of ODEs implicitly assumes that each cell has an equal probability $\frac{\mu B\tilde{C}}{\tilde{N}}\Delta t$ of leaving during each timestep (for small enough Δt), and so a cell that has just arrived in the organ has an equal probability of leaving as one that has been there for an arbitrarily long time. Organ localisation can only be controlled with parameters that describe input and output flow, and the input flow can be no higher than the rate of delivery. Cells cannot be made to leave with a delay rather

than exponential decay, and thus transient behaviour such as an initial build-up of cells cannot be described. This likely also causes issues with lymph node localisation, which can experience temporarily large input flows from other organs, but in which a build-up of localisation can only be achieved by assigning slow output flow. In the best fit to the TDLn data set, this is visible with a large initial spike of cells in the lymph nodes which decays to a steady state, rather than the desired behaviour of a gradual build-up. Though previous authors have fit T-cell localisation data to ODE systems [81–83], the timescales were much longer (hundreds of hours) and thus exhibited different behaviour, or they had similar issues with transients, and/or they did not fit to lymph node data.

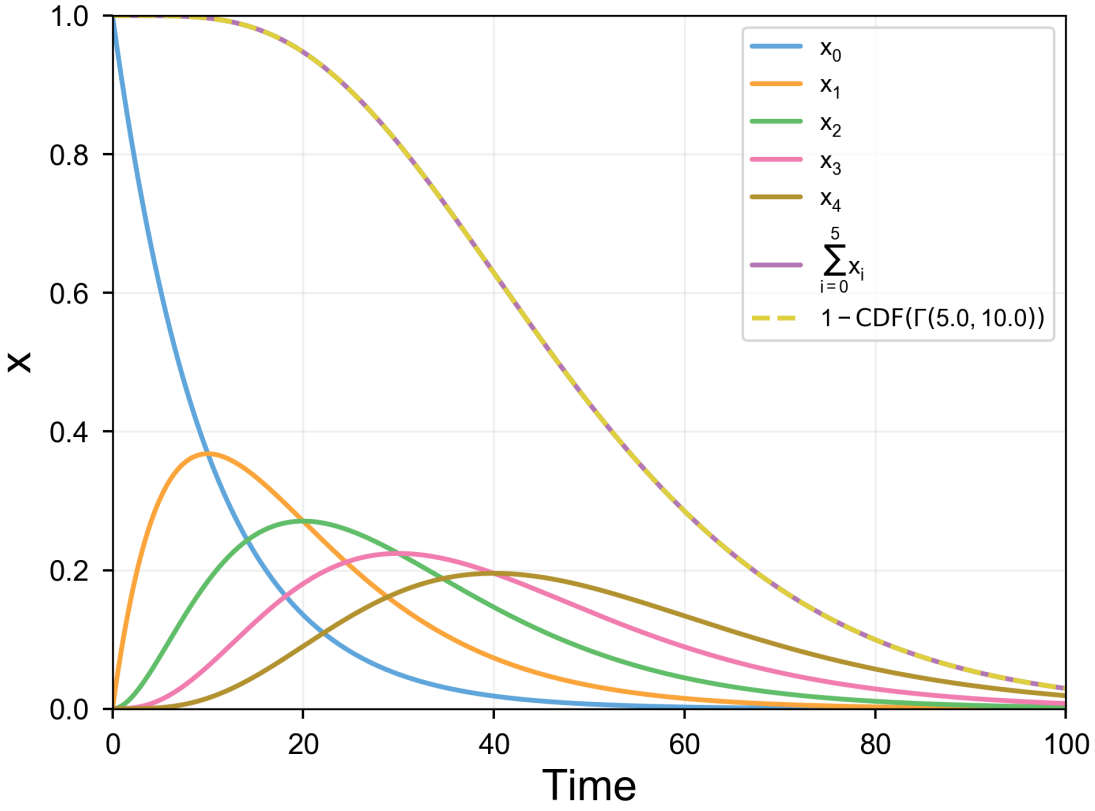


Figure 4.4. The solution to Equation 4.13 with five compartments and $k_{\text{in}} = 0$, Equation 4.14. Initial conditions are zero for all compartments but the zeroth. The amount of material in each of the five compartments against time is shown alongside the total amount of material in all compartments. Lines are shown for the contents of each compartment x_i and for the total number in all compartments. One minus the CDF of a gamma distribution (Γ) with shape and scale respectively equal to the number of compartments and $1/k$ is shown for comparison.

To improve the fit to the data, we desire a model in which the rate of exit of cells from each organ depends on how long they have been present. There are several methodologies to achieve this, such as multiple-compartment ODEs for each organ (*e.g.* [245]), delay differential equations (DDEs) or age-structured PDEs. However, the use of multi-

compartment ODEs would complicate control of the distribution of leaving times from each organ. Consider a set of ODEs that represent flow of cells x through a single organ:

$$\begin{aligned} \dot{x}_0 =& k_{\text{in}} - kx_0, & i = 0 \\ \dot{x}_i =& k(x_{i-1} - x_i), & i > 0 \end{aligned} \tag{4.13}$$

If k_{in} is zero and only the zeroth compartment initially has cells in it, then the solution to each equation is,

$$x_i(t) = \frac{(kt)^i}{i!} e^{-kt}. \tag{4.14}$$

For $i = 0$, this is exponential decay, but as i is increased, the distribution spreads out across time (see Figure 4.4). The sum of the cells in all compartments $\sum_i x_i(t)$ is equal to one minus the CDF of an Erlang distribution (gamma distribution with integer shape), with shape equal to the number of compartments N_{comp} and scale equal to $1/k$. As the number of compartments tends to infinity, then this total is of course equal to $x_0(0)$ for all time. Though the mean and variance of the distribution can be simply described in terms of distribution parameters, it is difficult to biologically justify the use of an Erlang/gamma distribution (or a sum of exponential distributions) to describe the persistence time of cells within an organ. We would not be able to explain what the number of compartments corresponds to, nor whether the distribution's heavy tail is real. In the face of uncertainty, the choice of a Gaussian(-based) distribution for the probability distribution of organ departure times is more parsimonious. A further problem is that control of the gamma distribution's mean and variance would require precise handling of the number of compartments in each organ, leading either to coarse-grained control of persistence times or to a large number of compartments per organ. The computational complexity here would not offset the performance advantage of using ODEs over PDEs or DDEs, if any remains with a large number of compartments. On the other hand, DDEs enforce no variation in the residence time, leading to degenerate dynamics such as persistent oscillations that do not decay. They can be as difficult to analyse and numerically solve as PDEs, but do not offer the same flexibility. Age-structured PDEs allow for the precise definition of the distribution of departure times of cells from an organ, which is a practical advantage over both multi-compartment ODEs and DDEs. For these reasons, we model the persistence of lymphocytes within organs with age-structured PDEs, as defined in Section 4.1.2.

4.2.3 Advection equation

To numerically solve unmodified partial differential equations, an appropriate discretised form of the equations must be selected. When $\gamma = 0$, the equation for the interstitial space takes the form of the advection equation, $\partial_t u = -s\partial_x u$. Thus, we begin with this equation

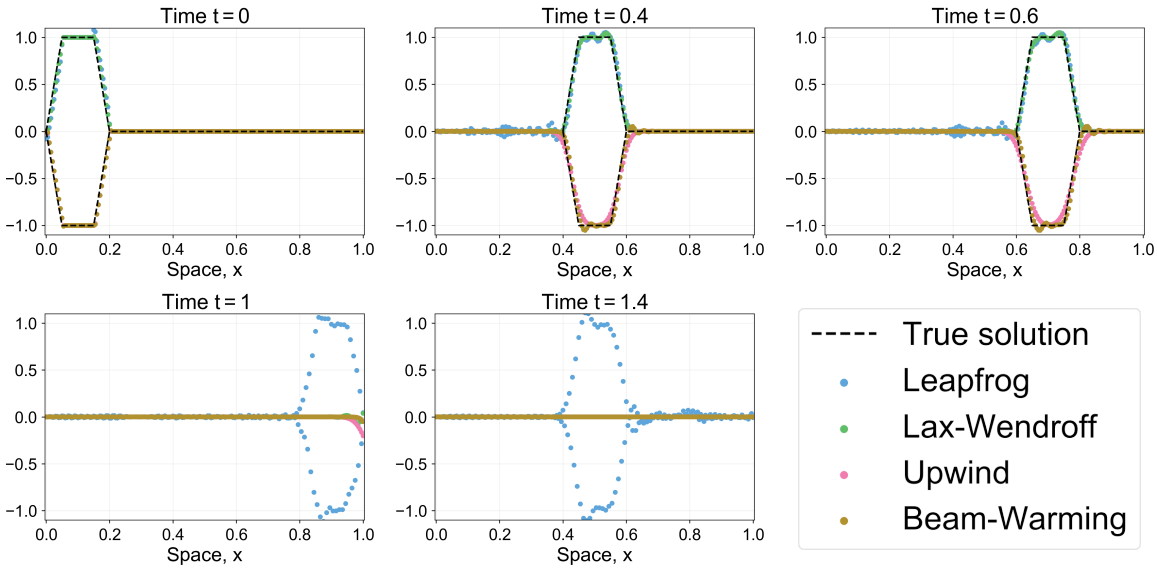


Figure 4.5. Numerical solutions to the advection equation, $\partial_t u = -s\partial_x u$, plotted against the true solution *without* periodic boundary conditions and with $s = 1$. Each panel shows a different timepoint; the initial profile can be seen to numerically evolve over time. The solution is symmetric about zero, so the Leapfrog and Lax-Wendroff solutions are shown for positive values, the Upwind and Beam-Warming solutions are shown for negative values and the true solution is shown for both. The dispersive nature of Leapfrog, Lax-Wendroff and Beam-Warming and the dissipative nature of Upwind can be seen. Without periodic boundary conditions, the leapfrog method generates a reflection of the solution at the boundary.

for simplicity and to understand the flaws in different techniques. We set $s = 1.0, \Delta x = 0.005, \Delta t = 0.8\Delta x$ and choose an upside-down paper cup as initial conditions. We solved the equation with each of the numerical update schemes listed in Table 4.0, with the exception of the Euler method, as it is unstable for this system. The solution due to each of these update equations is plotted alongside the analytical solution in Figure 4.5 (non-periodic boundary conditions) and Figure 4.6 (periodic boundary conditions). Generally, each solution gives a close approximation to the true solution, but each has some small problems. The Lax-Wendroff, Beam-Warming and especially Leapfrog methods are highly dispersive around the corners of the 'cup', and the Upwind method is dissipative. When periodic boundary conditions are not used, the Leapfrog additionally generates an anomalous reflection at the boundary. The integrals of each numerical solution in Figures 4.5 and 4.6 across all space as a function of time are shown in Figure 4.7. The dispersion from several numerical update schemes, particularly Leapfrog, can be clearly seen.

If Δt is set equal to Δx, then the Leapfrog method becomes unstable and the Lax-Friedrichs method becomes equivalent to Upwind. Furthermore, the solutions given by Upwind and Lax-Friedrichs become equal to the analytical solution to within machine precision (results not shown). This regime is important because in an age-structured PDE, both independent variables are an expression of time and thus $\Delta a = \Delta t$.

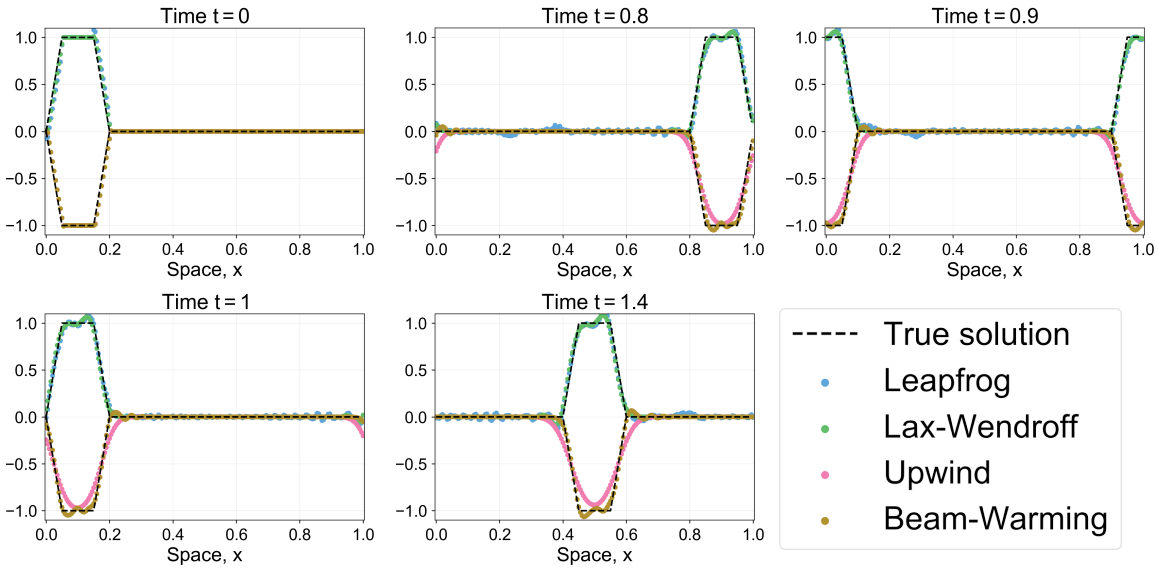


Figure 4.6. Numerical solution for the advection equation, $\partial_t u = -s\partial_x u$, plotted against the true solution *with* periodic boundary conditions and with $s = 1$. Unlike Figure 4.5, the Leapfrog method does not generate a reflection at the boundary.

4.2.4 PDE solution to single organ system

In an advective system where $\Delta a = \Delta t$, Euler and Leapfrog are both unsuitable. In the trafficking PDE system described in Section 4.1.2, the update equations for Lax-Wendroff and Beam-Warming contain a very large number of terms, some of which require a 5-point stencil, and hence we do not consider them here. That leaves the Upwind and Lax-Friedrichs methods, which are identical for $\Delta a = \Delta t$. The results of using these procedures to update a single organ system are shown in Figure 4.8. In this system, $\Delta t = \Delta a = 0.005$, $e = 0.5$ and $B = V = \tilde{V} = 1.0$. The left panel shows the instability of the leapfrog method in this regime. The middle panel shows the results of the heart and organ vasculature ODEs against time, the sum of the Upwind solution to the PDE against time and a snapshot of the Upwind solution against age for $t = 3.00$ minutes. The right panel shows the total number of cells in the system against time for the solutions shown in the middle panel, which is not conserved. The mass loss stems from the boundary condition of the PDE, $\tilde{N}(0,t) = eBC(t)$. The ODEs were solved by Scipy's well-documented *dopri5* solution (implemented in MATLAB as *ode45*), originally by Dormand and Prince [247]. The PDE, however, was solved with an effective RK2 solution. Hence, the mass leaving the ODEs each timestep is proportional to

$$\sum_{j=1}^{6} a_{7j} \left(N(t^i) + \Delta t \sum_{l=1}^{j-1} a_{jl} k_l \right)$$

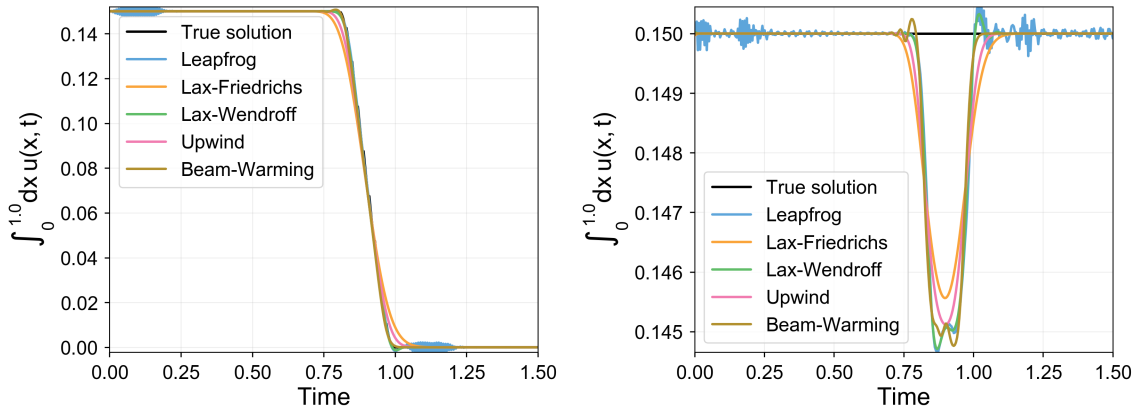


Figure 4.7. Integral of each of the numerical solutions and the true solution to the advection equation, for the time-courses shown in Figures 4.5 and 4.6. The left panel shows the integral for non-periodic boundary conditions and the right shows periodic boundary conditions. The loss of mass seen at $t \approx 0.9$ on the right plot is an artefact from the solution passing through the boundary.

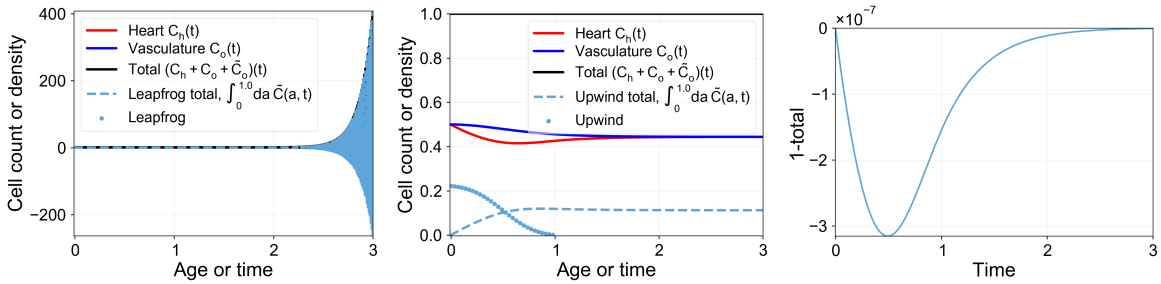


Figure 4.8. Numerical solutions to the PDE system described in Section 4.1.2 for a *single* organ as a function of time. Solid lines show ODE solutions (heart and organ vasculature) and the dotted line shows the integral of the PDE solution (interstitial space) over age. On the same axis, the scatter plot shows the solution of the PDE as a function of age at time $t = 3.00$. The left plot shows the Leapfrog method, which is unstable for this system. The middle plot shows both the Upwind and Lax-Friedrich methods, which are identical for this system. In each system, the initial number of cells is 1.0; as a numerical test, this total should always remain the same. The right plot shows the difference of the total number of cells from one (note axis is $\times 10^{-7}$) as a function of time for the middle plot, showing that this initial implementation of the PDE system did not conserve mass.

where the various a_{jl} terms are the coefficients of the intermediate steps of Dormand-Prince, given in Table 4.2. On the other hand, the mass entering the PDEs through its boundary condition is proportional to

$$\frac{1}{2}(N(t^i) + N(t^i) + \frac{\Delta t}{2}k_1),$$

causing a mismatch. Remedying this requires the values of every intermediate value used in the Runge-Kutta update step for use in the PDE integral for all previous times, which are not retrievable from *dopri5*. This hence required the creation of a custom solver, as described in Section 4.1.6.

One feature of the PDE system simplifies this task, however. The content of the inter-

stitial compartment is an integral over age,

$$\tilde{C}(t) = \int_0^{a_{\max}} \mathrm{d}a \tilde{C}(a,t),$$

which, using the boundary condition $\tilde{N}(0,t) = eBC(t)$ and the PDE to define a characteristic integral (see Section 4.1.5), can be restated:

$$\tilde{C}(t) = \int_0^{t'} \mathrm{d}a \frac{eB}{\tilde{V}} C(t-a) e^{-\int_0^a \mathrm{d}a' \frac{\gamma(a')}{\tilde{V}}}, \quad t' = \begin{cases} t, & t \leq a_{\max} \\ a_{\max}, & \text{otherwise.} \end{cases}$$

The exponent can be pre-calculated, because γ, $a_{\max}$ and $\tilde{V}$ are all known quantities. Hence, the value of $\tilde{C}(t)$ may be calculated every timestep using only the solution of C from previous timesteps. However, as Figure 4.8 indicates, mass will not be conserved if point estimates are used for C at every interval Δt to calculate $\tilde{C}$: the integral must include each intermediate step calculated by the Runge-Kutta algorithm, including smaller timesteps required to keep the error in the calculated solution within tolerance. This critical requirement further motivated the development of the custom solver described in Section 4.1.6.

4.2.5 Method of characteristics solution

4.2.5.0 Simple system

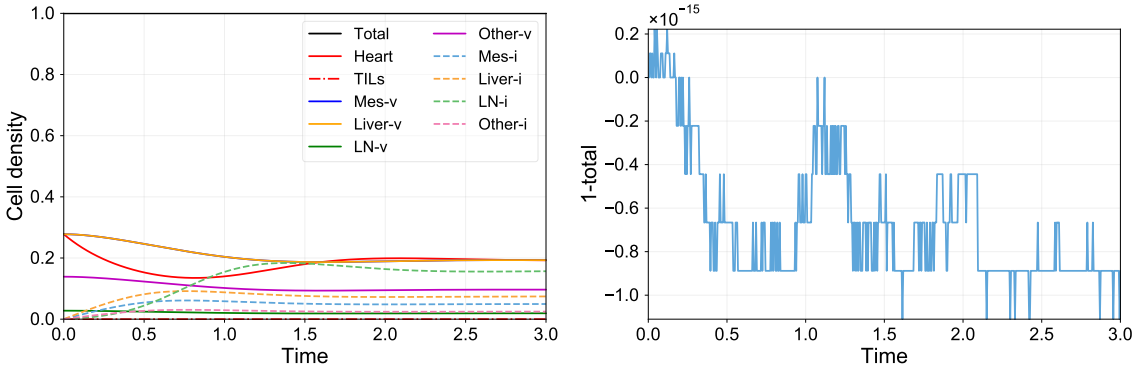


Figure 4.9. Numerical solution to the PDE system described in Section 4.1.5, where the PDEs have been replaced by equivalent characteristic integrals, for a simplified system. The system contains the heart, a single compartment for the mesentery containing the tumour, the liver, the lymph nodes and an "other" compartment. Solid lines show ODE solutions (heart and vasculature) and dotted lines show the PDE solutions (characteristic integrals; interstitial space). The total number of cells in the system at initiation is 1.0 and the right panel shows the difference in the total number of cells from 1.0 as a function of time.

The custom algorithm was first used on a test system which contains the heart, the mesentery, the liver, the lymph nodes, 'other' organs and a tumour in the mesentery, with parameters as follows: Δt=0.005 min, e_{tmr}=0.0, $e \in \{0.0, 0.5, 0.5, 0.99, 0.5\}$, $\mu \in \{0.5, 0.5, 0.5, 0.5\}$ min, $B \in \{1.0, 1.0, 0.1, 0.5\}$ L/min, V_{heart}=1 L, $\tilde{V}_{\text{tmr}}$=0.4$\tilde{V}_{\text{mesentery}}$, $V \in \{1.0, 1.0, 0.1, 0.5\}$ L, $\tilde{V} \in \{1.0, 1.0, 0.1, 0.5\}$ L. The results are shown in Figure 4.9, which appears to show that the system conserves mass to machine precision.

4.2.5.1 Full system

Organ	Sto.	S.int	L.int	Pan.	Spl.	Liv.	Bra.	Hrt.	Lung
e	0.0	1.0	0.008	0.008	0.197	0.037	0.008	0.008	0.008
μ	50	480	236	236	50	180	236	236	120
Organ	Kid.	Fat	Skin	Mus.	Ske.	LNs	Thy.	Misc	
e	0.008	0.008	0.5	0.008	0.759	0.833	0.687	0.008	
μ	236	236	480	236	50	240	480	236	

Table 4.4. Values of e and μ used in the first two panels of the test of the full PDE system shown in Figure 4.10. Most organ names are abbreviated, *e.g.* "Sto." is Stomach.

After the model satisfied expectations in the simple system and appeared to conserve mass to machine precision, it was tested on the full system with rat parameters, Δt=0.1, and e and μ as given in Table 4.4. The results are shown in the top panel of Figure 4.10, which again appear to show that mass is conserved to within machine precision. To test the dynamic time-stepper and see if the system is still stable with large requested timesteps, the model was run with the same parameters but Δt=5. Again, the system appeared to conserve mass to machine precision, and the solution given was the same as for Δt=0.1 to the same precision (latter data not shown). Another set of values of e and μ with Δt=0.1 were implemented, with results as shown in the bottom panel and with the same conclusions. Finally, a number of simulations with different timesteps were compared, to confirm that the system converges to a single solution for a single parameter set. These results give us confidence that the ODE/PDE solver is behaving correctly and is stable.

4.2.6 PDE fit to localisation data

We attempted to fit the PDE model to Smith *et al*'s data with the same method as for the ODEs (Section 4.2.1). An example of the posterior distribution from (PMC-)ABC is shown in Figure 4.11 and the best fits to each data set are shown in Figure 4.12. In this case, the fits of the model to the data were very good, with scores (equal to the sum of the squared differences from the data, Equation 4.0) improved from those of the ODE model. Furthermore, the posterior distribution obtained from ABC for most parameters had centred onto a very small region of parameter space, indicating that a single best value (or very confined region) of the parameters had been found. The exceptions were the lymph nodes and lung pulmonary circuit in most cases, which each had a constrained value of the mean return time but not the extravasation probability. The fits produced from the mean parameter values or maximum densities (which were very similar) captured the features of the data qualitatively well, with some small issues such as oscillations in the number of lymphocytes localising in the lymph nodes not seen in the data. This oscillation perhaps

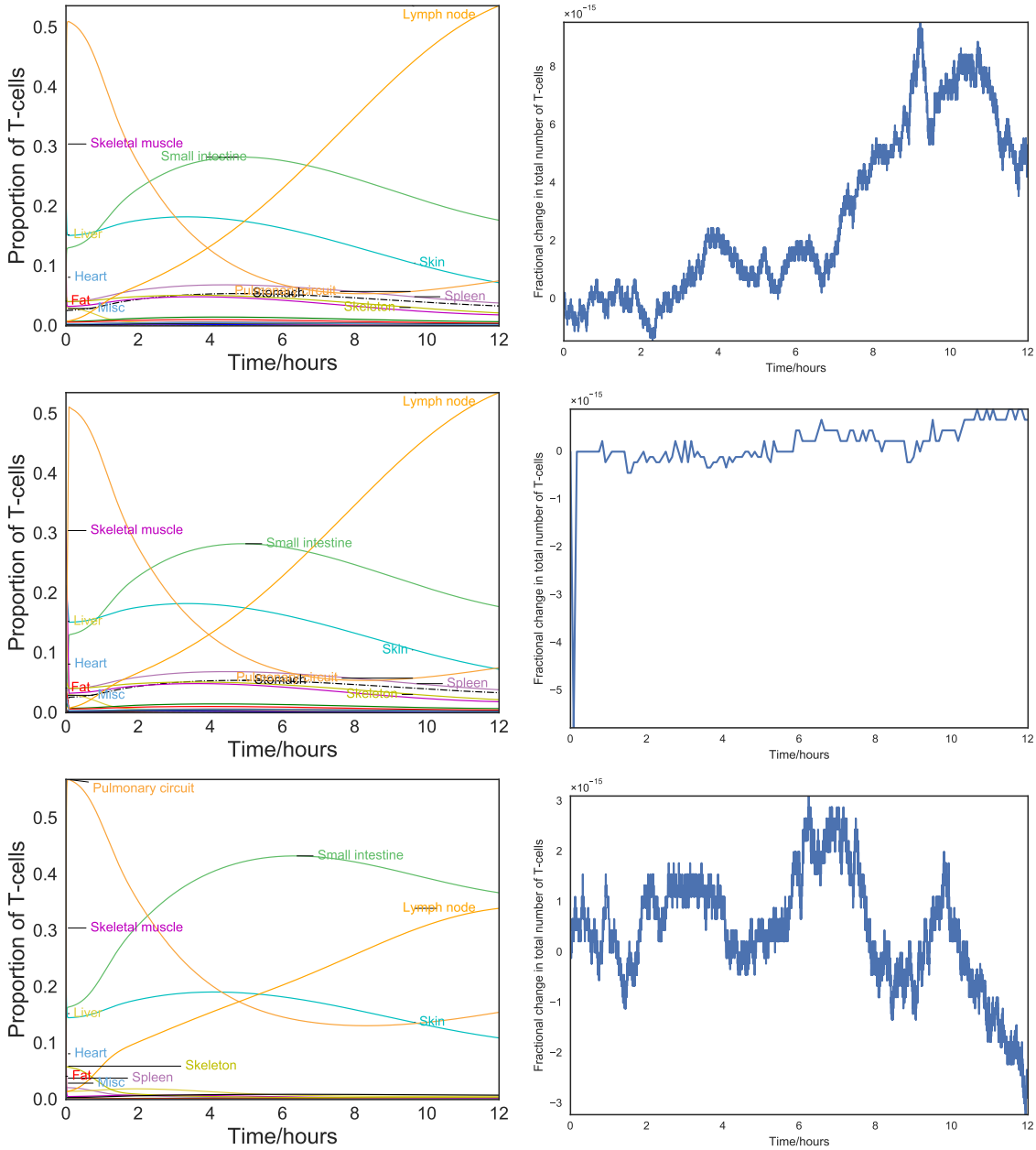


Figure 4.10. Numerical solution to the PDE system described in Section 4.1.5, where the PDEs have been replaced by equivalent characteristic integrals, for the full rat-parameterised system. Lines in the left panels represent the sum of the number of cells in vascular and interstitial compartments for each organ. Panels on the right hand side show the fractional change in the total number of lymphocytes in the system. The top two rows have identical parameters, but the top has $\Delta t = 0.1$ and the middle has $\Delta t = 5$. The bottom row uses a different set of values for the extravasation probabilities e and the mean return times μ.

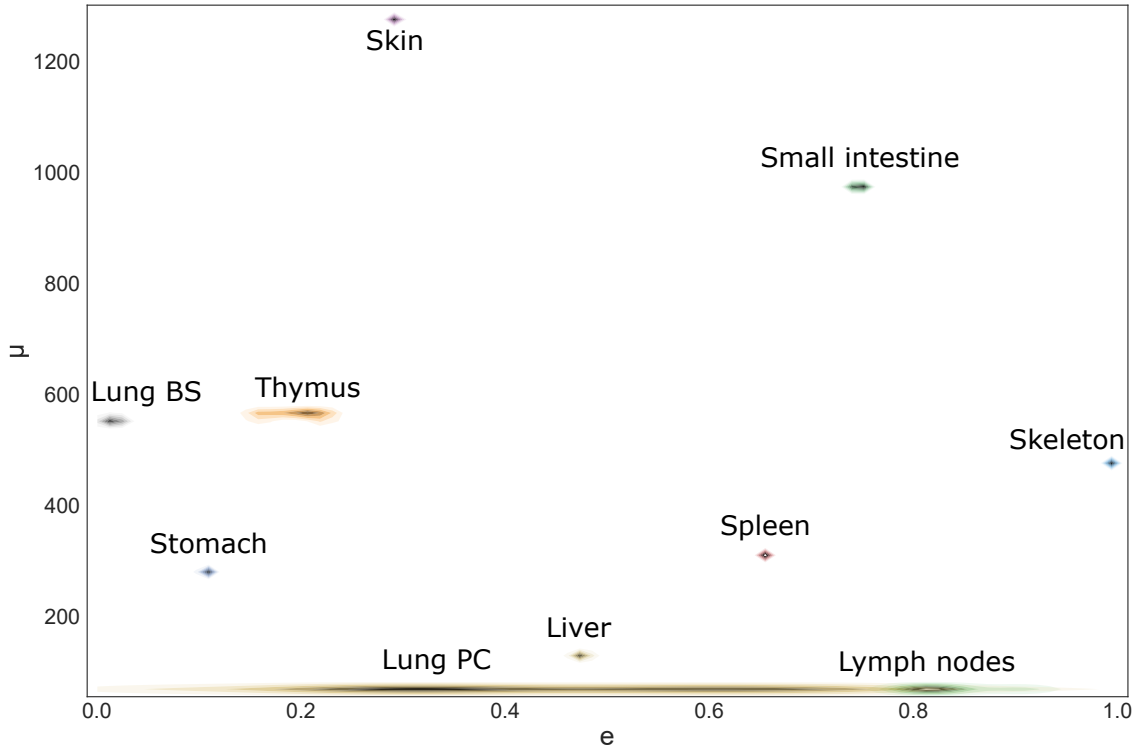


Figure 4.11. Posterior distribution for parameter values that cause the PDE model to fit the TDL dataset, obtained from Approximate Bayesian Computation (ABC). The mean return time μ is plotted against the extravasation probability e for all parameters marked in the legend. Every parameter is very constrained, except for the lung pulmonary circuit and the lymph nodes, for which the extravasation probability e is relatively unimportant.

indicates that the standard deviation of the Gaussian contained within $\gamma(a)$ was too high for the lymph node return time distribution, and that the constraint of $\sigma_{\mathrm{LN}} = \mu_{\mathrm{LN}}/2$ would need to be relaxed to improve the fit. The parameter values chosen to produce each plot are presented in Table 4.5, with which we may perform a more quantitative comparison. There are noticeable differences between each parameter set; for example, the probability of extravasation for small lymphocytes into non-lymphoid tissues is very low. Further comparison is difficult to make with unequal means, particularly as the mean value of e for the CFA dataset is much lower than any other. Normalising e and μ by the sum of values of e and μ in each data set allows for quantification of the difference between the five parameter sets. The root of the sum of the squared difference between parameter values in each set is shown in Table 4.6. It can be seen that the naïve (small) lymphocytes are consistently different to the activated (large) cells, as expected. The most similar datasets are ABA and CFA, followed by MLN and TDL.

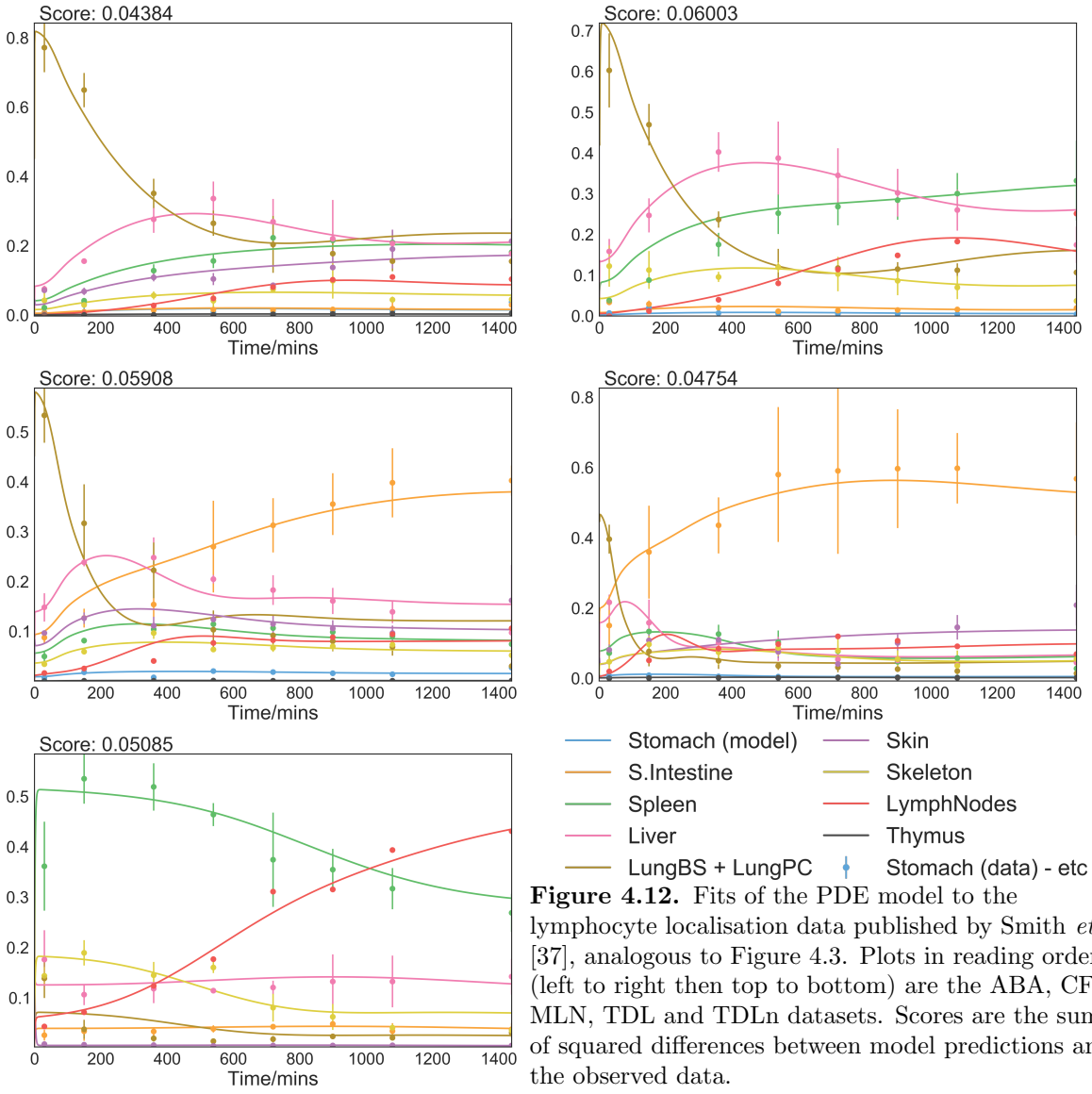


Figure 4.12. Fits of the PDE model to the lymphocyte localisation data published by Smith *et al* [37], analogous to Figure 4.3. Plots in reading order (left to right then top to bottom) are the ABA, CFA, MLN, TDL and TDLn datasets. Scores are the sum of squared differences between model predictions and the observed data.

4.3 Discussion

In this chapter, I described the development of a PDE model of lymphocyte trafficking that is able to quantify finite residence times of cells within each organ. The aim of this work was to produce a proof-of-concept of the model and to acquire first approximations of unknown biological parameters. Our future objective is to quantify the situations in which not delivery but extravasation of T-cells impedes tumour clearance. Existing differential equation solvers in Python, MATLAB, C and so on cannot solve arbitrary PDEs like they can ODEs, which necessitated the creation of a bespoke solver. This has been checked thoroughly for simple cases and, though it does not perfectly conserve mass, the fluctuations in the total mass seem to be due to machine precision (and are no greater than the fluctuations that the

Organ	ABA		CFA		MLN		TDL		TDLn	
	e	μ	e	μ	e	μ	e	μ	e	μ
Stomach	0.084	719	0.011	729	0.120	531	0.104	284		
Small intest.	0.023	723	0.007	708	0.319	1212	0.748	974	0.023	1130
Spleen	0.413	1077	0.163	1413	0.426	450	0.649	305	0.678	748
Liver	0.180	555	0.053	737	0.267	310	0.476	129	0.038	1011
Lung b.s.	0.023	486	0.046	620	0.035	473	0.015	559	0.026	658
Skin	0.262	1342			0.481	450	0.296	1280	0.006	1022
Skeleton	0.459	823	0.245	688	0.804	512	0.997	479	0.703	504
Lymph nodes	0.074	141	0.121	255	0.961	67	0.501	73	0.867	1247
Thymus	0.105	856			0.079	536	0.187	561		
Lung p.c.	0.514	66	0.088	68	0.273	64	0.244	38	0.005	438
Mean value	0.214	679	0.092	652	0.377	460	0.422	468	0.293	845

Table 4.5. Parameter values used in each panel of Figure 4.12. The labels ABA, CFA, MLN, TDL and TDLn refer to the names of each of the datasets used; see text for an explanation of each. e is the extravasation probability of lymphocytes into each listed organ, and μ is the mean residence time of lymphocytes within each organ. Abbreviations: Lung b.s., Lung blood supply. Lung p.c., Lung pulmonary circuit.

ABA-CFA	MLN-TDL	CFA-MLN	ABA-TDL	CFA-TDLn
0.267	0.274	0.316	0.326	0.344
CFA-TDL	ABA-MLN	MLN-TDLn	TDL-TDLn	ABA-TDLn
0.349	0.360	0.364	0.443	0.500

Table 4.6. The square root of the sum of the square of differences between parameter values in each pair of datasets in turn, after values of e and μ are each respectively normalised by the sum of values in their dataset and parameters missing from either dataset are excluded. The table has been arranged in ascending order. Please see the Discussion section for further details.

ODE system produces). The solutions produced for several test cases match those produced from Scipy's dopri5 algorithm for the ODEs and an RK2 update step for the PDEs by eye, notwithstanding that the Scipy implementation results in a loss of mass. The developed solver aborts if the total change in mass or change in one timestep is large, and it produces a plot of total mass against time when used for single parameter sets. These tests and checks help to ensure that the algorithm is reliable before pursuing further investigations.

The model itself, like the ODEs in Chapter 3, depends on anatomical data for blood flows, blood volumes and organ volumes, as well as two sets of unknown quantities e and μ, the extravasation probability and the mean return time of lymphocytes in and out of organs. Results thus depend strongly on the data used for parameterisation. Temporal data describing localisation of lymphocytes within the body are required to fix the values of e and μ, but such data is difficult to find and often sparse. For this reason, Gaussian distributions with standard deviations proportional to their means are used to describe the return times of lymphocytes from each organ, to reduce the number of parameters required to be fitted. With richer data, however, it would be worthwhile to investigate

whether other distributions, such as the Gamma distribution, may provide a better fit to lymphocyte return times.

The data presented in this Chapter were reported in 1980 [37] and did not control for different cell types: effector cells (B-cells, $CD4^+$ and $CD8^+$ T-cells), and small lymphocytes (naïve and memory cells) were all mixed together. CD4 and CD8, for example, had been discovered less than 5 years prior to Smith *et al*'s publication. Large, dividing (activated effector) cells are referred to as "lymphoblasts" in the study, and small (memory or naïve) cells are referred to simply as "small lymphocytes". We are interested primarily in the localisation of $CD8^+$ T-cells and the migration of these different populations is expected to differ. However, these data suffice for our aim to create a proof of concept and gain order-of-magnitude estimates for parameter values. The original study aimed to investigate the factors that determine migratory patterns of lymphocytes, hence there are datasets for lymphoblasts harvested from a variety of lymph nodes after a variety of stimulating antigen had been introduced. We can inherit this aim and determine how estimated parameter values vary between the different datasets and whether they are consistent with our knowledge. The ABA and CFA data sets were intended to determine whether the antigen that lymphocytes carry determines their homing patterns: both datasets concern cells extracted from axillary and brachial lymph nodes, but the ABA dataset considers cells that have been extracted several days after a skin graft was sewed to the centre of the rat's chest, whilst the CFA dataset is for cells that have been extracted after ground human *M. tuberculosis* bacilli were introduced into the rat's flank, near the target lymph nodes. The MLN and TDL data sets were used to determine whether the location or method of cell extraction correlated with observed homing patterns: both contain cells from rats that had been exposed only to environmental antigen, but the MLN cells were extracted from mesenteric lymph nodes and the TDL cells were extracted from the thoracic duct. The TDLn data set, finally, was used to compare the homing pattern of naïve cells: these were (mostly) "small" cells extracted from the thoracic duct, studied in the same rats as in the TDL dataset.

In 2018, we now know that lymphocyte homing patterns are determined by their class (*e.g.* $CD8^+$ T-cell versus memory B-cell) and the location that the antigen was detected by the body (and, in turn, the markers on the antigen presenting cell that activated the lymphocyte). We thus expect the parameter values for ABA and CFA datasets to be similar, as the antigen was introduced in similar locations each time. We also expect the MLN and TDL datasets to yield somewhat similar parameters to each other, as they should each contain a mix of groups of lymphocytes with distinct homing patterns, though the former may favour lymphocytes that prefer the mesenteric LN. We expect the parameters for the TDLn dataset to be completely distinct from the rest, as it contains primarily naïve

and memory cells, which localise primarily in secondary lymph tissue and, in the case of peripheral memory cells, some outer tissues (*e.g.* the dermis or bone). We will not analyse the results of the data, as this was discussed by the original authors; instead, we will focus on how the model fits the data and how this manifests in the parameter values. Nonetheless, their key observations were:

- **ABA:** A high proportion of ABA lymphoblasts were found in the lung early on and drained very slowly, which is probably the cause of the slow build-up in the spleen.
- ABA localisation in the gut was very low, much lower than the small intestine.
- Localisation in mesenteric and caecal LNs was low, but high in superficial cervical, deep cervical, axillary, para-aortic and parathymic LNs, and very high in the coeliac LNs.

- **TDL:** A high proportion of TDL lymphoblasts were found in the lung too, but left faster than did ABA lymphoblasts.
- Most TDL lymphoblasts localised in the small intestine, too.
- Localisation was high in the mesenteric and caecal LNs, 4-5 times lower in the superficial cervical, deep cervical, axillary, para-aortic and parathymic LNs, and intermediate in the coeliac LNs.

- **CFA:** The localisation of CFA lymphoblasts was very similar to that of ABA lymphoblasts.

- **MLN:** The localisation of MLN lymphoblasts was closer to TDL than ABA, but otherwise behaved as an intermediate population.

- **TDLn:** About 3–6% of TDL from unstimulated rats are lymphoblasts. Intestinal localisation of TDLn cells was about 11% of that of TDL lymphoblasts, so the TDLn group should be expected to be at least 89% small lymphocytes.
- Unlike the lymphoblasts, localisation of the small lymphocytes was found to be uniform among all LNs.

Most qualitative features, general trends and major sites of localisation are consistent between Smith *et al*'s data are the model's predictions (Section 4.2.6). Compared to effector cells, small lymphocytes have much higher extravasation probabilities to the lymph nodes

and spleen than for outer organs (as shown in Table 4.5). Likewise, the mean lymph node residence time is predicted to be higher for small lymphocytes than lymphoblasts. The mean value of e and μ across each dataset is not the same, so we normalised each set of values of e and μ by their respective sums in each dataset, to compare relative values (data not shown). Taking the sum of the square of the difference between these values in different datasets also allows us to quantify how similar they are (Table 4.6). As expected from the biology, parameters leading to a fit to the ABA and CFA datasets are the more similar to each other than any other, followed by MLN and TDL. The TDLn dataset, consisting of small/naïve lymphocytes, is most dissimilar to the rest. Though the relative values are similar between datasets, it is possibly anomalous that the mean values are so different. In particular, we would expect the mean value of e to be similar between ABA and CFA, but CFA's is much lower. This is likely to be due to the much greater localisations seen transiently in the liver and finally in the spleen in the CFA dataset than in ABA. Though the organs in which cells localised are similar between the two datasets, the exact quantitation of this localisation is different and hence the two fits require different parameter values. It is harder to explain this biologically, however. Had the ABA and CFA datasets each consisted of a single rat, it would have been likely that this difference was due to small, random differences between the rats. However, both datasets consisted of multiple rats of the same strain, and experimental handling was identical but for the differences described above.

Smith *et al* also observed that the localisation of MLN lymphoblasts was closer to TDL than ABA, which is consistent with parameter values that fit each dataset (see Table 4.6). Many of Smith *et al*'s other observations concerned the difference between different groups of lymph nodes. The parameterisation of individual lymph nodes is limited by a lack of data and so our model combines lymph nodes into one compartment, so these observations cannot be directly tested. However, fits to the model indicate that the probability of extravasation to outer tissues is considerably lower for naïve cells than activated ones, and activated T-cells preferentially home to a subset of organs. Activated cells that have extravasated into an organ return to the blood through the lymphatics, via the lymph nodes that happen to drain those organs. Because activated cells preferentially home to particular organs, this would reproduce the result that they localise preferentially in particular groups of lymph nodes. Naïve cells, however, do not generally enter activated tissues and instead reach lymph nodes by extravasation through high endothelial venules. Assuming that e and μ are identical for high endothelial venules in different lymph nodes would lead to the result that only naïve cells localise uniformly among all lymph nodes.

Fits to Smith *et al*'s data indicate that a separate compartment is required for the pulmonary circuit. If a single organ were defined for the pulmonary tissue which receives at least 50% of the bloodflow, and that compartment allows extravasation to the lung's interstitial space, then shortly after the introduction of radiolabeled lymphocytes, a proportion $\frac{(eB)_{\text{lung}}}{\sum_o B_o}$ of them extravasates into the lung's interstitium. In order to return to the blood, these cells must drain through the lymphatics into lymph nodes. In all of the data but the TDLn data, a large number of lymphocytes is present in the lung shortly after injection, as might be expected. However, they persist in the lung for several hours and there is *not* a subsequent large increase in localisation in the lymph nodes. This suggests that the cells are trapped in the lung vasculature rather than the interstitial space. Creating a second 'lung' compartment for the pulmonary circuit to which cells can 'extravasate' into a trapped state, but from which they return to the heart, greatly improves the fit to the data. There are two potential explanations for this trapping mechanism, the first of which is due to the anatomy and physiology of the lung. The vascular bed is large, branches considerably and periodically drains of blood. The blood vessels in the lower part of the lung are generally always filled, whereas the vessels in the top are generally collapsed [250]. The higher the interstitial pressure, the more of these vessels become perfused – for example, during exercise. This means that although total blood flow through the pulmonary circuit is always equal to the cardiac output, the pressure gradient and extent of perfusion are not constant. Thus, we hypothesise that there is a possibility that the transient lung localisation is due to cells that become stuck or are left behind. The second potential reason for the observed persistence stems from the lymphocytes themselves, which may be actively delaying their departure from the lungs, rather than simply being swept along by the blood. Persistence in the lung was observed in the lymphoblast localisation data, but not in the small lymphocyte (TDLn) data, which supports this possibility. Smith *et al* themselves suggest that the observation may be explained by T-cells encountering viruses in the lungs.

The existence of two compartments for the lung means that there are two pairs of values of e and μ for the same dataset. This results in much greater uncertainty in the value of e for the lung pulmonary circuit than in other parameters (visible in the density plots, Figure 4.11). Though the pulmonary circuit is required for the initial high lung localisation, late time behaviour is possibly over-specified by four parameters. Another issue with parameter values that is not related to Smith *et al*'s observations is that the values of e and μ for the skeleton are much higher than perhaps should be expected. This may, as also noted in Chapter 3, be due to unreliable values of parameters, in turn due to difficulties defining a perfused volume for the 'skeleton' to which cells can extravasate.

The greatest weakness of the PDE model is that it is slow. Whether it is solved explicitly or with a characteristic integral, every timestep must involve an integration over the entire history of values of C to obtain the number of cells leaving the interstitial space(s) in each time interval. Since the values of γ differ by age, repeating this integral is unavoidable and represents a significant bottleneck. Careful development of the ODE/PDE system and its associated Dormand-Prince solver in C alleviates most of these issues and achieves performance comparable to the ODE system, but with a key disadvantage: when parameter values lead to a stiff system, a Dormand-Prince solver is inappropriate. This may be remedied in the future by solving the PDE system with the method of lines, to produce a system of ODEs discretised in age, and then expressing the system as a matrix. This set of equations could then be solved with off-the-shelf ODEs solvers, stiff or otherwise, and should have performance comparable to the current implementation.

The efficiency of the ODE and PDE systems' current implementations are more than adequate for simple runs of the model, but not for extended analyses such as Approximate Bayesian Computation (ABC) and a sensitivity analysis. We used a very small sensitivity analysis for the number of parameters, with approximately 1000 runs among 36 parameters. This numbers of runs is very likely to be inadequate for an accurate estimate of the global sensitivity of the system to different parameters. However, a highly accurate analysis was not required for our aim to determine if organ parameters could be fit to data by considering one organ at a time. The analysis predicted that this was the case and we indeed acquired good fits to the data, which indicates that no further runs of the sensitivity analysis were required. A similar statement can be said of ABC: it is not clear whether the number of samples, iterations or passes taken were enough for ABC to converge to a final posterior distribution for either the ODE or PDE fits. However, ABC was used as a tool to gain reasonable fits of the models to the data. In the case of the PDEs, this was achieved, and so the attributes used for fitting were more than sufficient for this proof-of-principle study. For the ODEs, additional passes did not improve the density plots by eye and the equations do not seem to produce the features of the data, so it is unlikely that significantly better fits could be achieved.

The PDE model has produced satisfactory fits to the lymphocyte localisation data, thus fulfilling its purpose as a proof-of-concept. The model has demonstrated that it is possible to qualitatively match persistence of lymphocytes in different organs and to generate hypotheses about lymphocyte behaviour (on the pulmonary circuit, for example), and leaves ample opportunity to use both the PDE and ODE models for further study.

4.4 Future work

The immediate next steps from this work would be to fit other data sources. However, data that contains enough time points in enough organs is very rare. In particular, high spatial resolution requires the animal to sacrificed, which limits the experimenter to a single time point or to averaging over multiple animals (as Smith *et al* did). Attempting to track radiation in a living creature lowers resolution – below that of the level required to image lymph nodes in mice, for example. It may also require an animal to be anaesthetised, which alters the distribution of blood flow. Other issues include the fact that cells may shed their label over time, which may confound results. We have recently acquired data from a collaborator that tracks radiolabeled T-cells within living mouse cancer models. This dataset contains T-cells only and uses more modern technology, which should result in more reliable data. That these are tumour models allows quantification of tumour localisation or proliferation, in addition to organ homing. This data may allow us to determine how e and μ vary between mice and rats (and between cell types) and in the presence of a tumour, if at all.

We would also like to fit the model to other data, particularly human data, if available. In the presence or absence of data, it may also be insightful to model the process of T-cell rolling and stopping explicitly instead of assigning a probability e to this process, reducing the dependence of model conclusions on reliable data. However, this would be contingent on the additional modelling complexity and uncertainty not being prohibitive, and may also require validation data. Fitting to human, rat and mouse data would also allow for the comparison of the extravasation probabilities e and mean return times μ in each species. Though delivery rates were found to differ greatly between the species in Chapter 3, e and μ could conceivably be similar between them, as they depend on binding of selectins and integrins rather than anatomical parameters. The parameter regimes in which step 5 of the cancer-immunity cycle impedes tumour clearance could hence be closer between pre-clinical models and humans than are the regimes in which step 4 contains the bottleneck.

In keeping with our discussion of the effect of interventions such as fasting or Avastin, we could explicitly model the expected effect of such interventions on delivery rates, or we could model the vasculature of each organ and how it differs (*i.e.* the degree of fenestration or shunting). We could further develop the T-cell proliferation and tumour growth model that was introduced in Chapter 3 and extend it by modelling the presence and/or formation of tertiary lymph organs in the tumour periphery.

There is also room for further technical development of the model. It would be beneficial to re-develop the model with the method of lines, to allow for flexible use of other solvers.

Another interesting question is when, in a biological context, the PDE model is required. Other authors have achieved reasonable fits of ODE systems to data [81–83], but the experimental systems were very different. In particular, several of these models addressed data at much longer timescales (hundreds of hours). It is possible that PDEs are only required at short time-scales. An interesting question is whether these short timescales are clinically relevant, or conversely, which timescales are clinically relevant? This question, and much of the work described above, will form the basis of my post-doctoral research, following the completion of this DPhil.

Chapter 5

Thesis discussion

5.0 Research questions and hypotheses

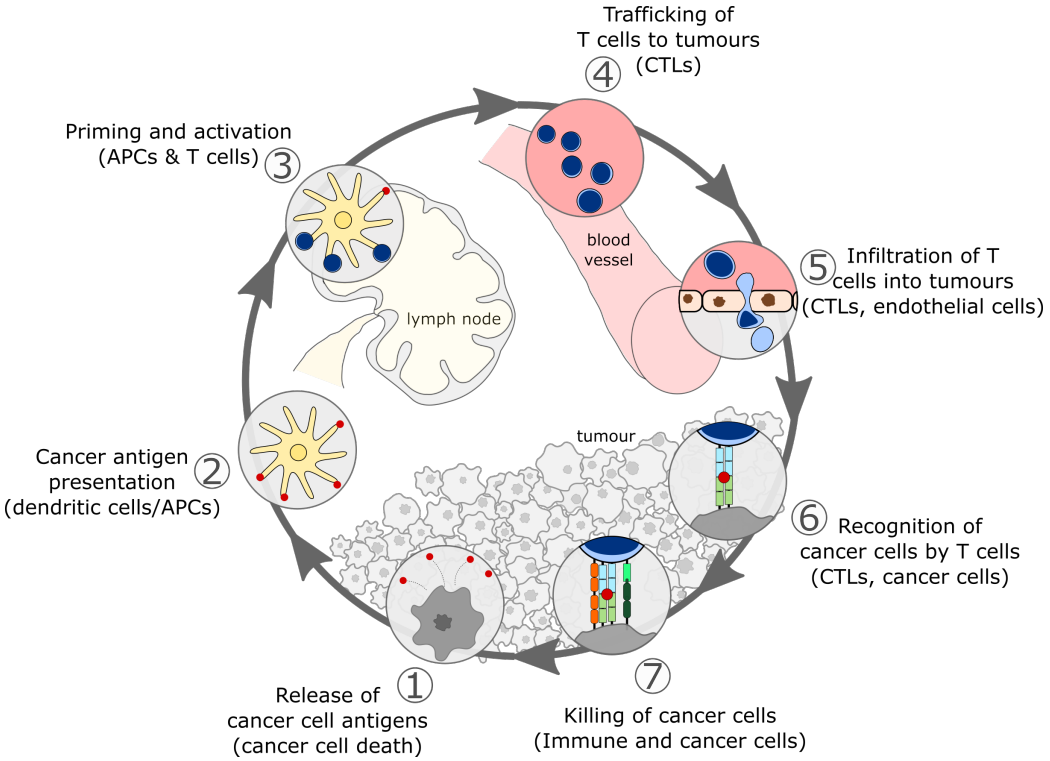


Figure 0.4. Repeated from page 13. The cancer-immunity cycle [5], reproduced with permission. It comprises seven major steps beginning with release of cancer cell antigen and ending with the killing of cancer cells by the immune system. 'APC' is 'antigen presenting cell' and 'CTL' is cytotoxic T lymphocyte. Reproduced from a published figure [5].

In the work presented in this thesis, we aimed to produce mechanistic models that could quantify the (limits of) efficacy of immunotherapeutics against cancers, and to propose mechanisms for when such immunotherapies break down. I discussed each chapter in the context of the cancer-immunity cycle [5], Figure 0.4, the proposed set of mechanisms through which the adaptive immune system eradicates tumours under normal conditions. I focused

on two examples of immunotherapies that represent different parts of the cycle – short peptide vaccinations, to consider steps 1–3, and chimeric antigen receptor (CAR) T-cells, to consider steps 4 and 5.

Chapters 1 and 2 follow the hypothesis that the results of IMA901, a clinical trial of short-peptide vaccination of renal cell carcinoma [52, 53], can be explained by properties of the vaccine that vary between patients. In Chapter 1, I presented a model of T-cell activation in the lymph node following the introduction of short-peptide in the dermis, generating the more specific hypothesis that the off-rate of short peptides from pMHC-I critically determines the resulting probability of successful T-cell activation. Chapter 2 detailed how this model has been used to reproduce the results of IMA901 and hence that IMA901's results are consistent with our hypothesis.
Chapters 3 and 4 focused on T-cell trafficking and particularly CAR T-cells. The work of Chapter 3 follows the hypothesis that the efficacy of CAR T-cells in pre-clinical models of solid tumours has failed to translate to the clinic due to differences between the rates of delivery of cells from the blood in humans and mice. I presented an ODE model of T-cell trafficking across the body that has been used to determine upper limits for the delivery of CAR T-cells into different organs, with findings that were consistent with our hypothesis. This work was extended in Chapter 4 to consider organ-specific localisation and to estimate values of the probabilities of T-cell extravasation into different organs.

In this chapter, I will summarise the findings of each chapter in turn, before discussing the work as a whole in the context of our current knowledge of the biology.

5.1 Research summary

In Chapter 1, I described an agent-based model of T-cell activation in the lymph node, building on work by Celli *et al* [19]. Many agent-based models assume diffusion of cells by default and we investigated the validity of this assumption, finding that matching experimentally observed dynamics ('run-and-tumble' motion) has a pronounced impact on predicted probabilities of T-cell activation. We tested various other assumptions common in published work, finding for example that the motion of dendritic cells does not significantly impact predictions. Previous studies have simultaneously considered cellular dynamics and antigen [70–72] and the work by Celli *et al* considered how many cells are required for successful activation, but we were the first to investigate the 'trade-off' between properties of the antigen or cell counts, and to quantify how both may lead to the failure of immune response. Unlike previous authors, we did not aim to produce a predictive model, but to propose a mechanism for clinical failure. Our model suggested that the off-rate of short peptides from pMHC-I complexes is a critical variable for short peptide vaccination success. In Chapter 2,

I presented the use of the model to run virtual clinical trials according to IMA901's schedule. We found that the virtual trial only matched IMA901's results in each phase if the values used for the peptide off-rates were measured values, as opposed to values obtained from machine learning algorithms such as NetMHC [153]. This supports the hypotheses generated in Chapter 1, that clinical failure may be caused by the rapid unbinding and clearance of short peptides, and that their off-rates from pMHC-I complexes are critical in a vaccination. We were able to use our framework to suggest alterations to the clinical trial that may have resulted in improved patient response. Virtual clinical trials have been used in the past to make predictions of clinical efficacy [62, 147], but we have not found any such studies that aim to test mechanisms of clinical failure. Had an *in silico* study akin to this been developed simultaneously with the clinical study of IMA901, then such trial modifications could have been tested or implemented and resulted in improved patient response rates.

In Chapter 3, I introduced an ODE model of T-cell trafficking across the body, akin to existing pharmacokinetics and pharmacodynamics (PKPD) and cell trafficking studies. We quantified the maximum rate that T-cells are delivered to different organs by the blood, in humans, rats and mice. We considered how delivery rates to the same organ scale between the species, showing that delivery rates of T-cells per mm^3 are 10,000-fold higher in mice than humans, but typical doses of CAR T-cells in the clinic are only 100-fold less. This contrasts with the oft-given explanation that the lack of efficacy observed is due to impaired extravasation or target-engagement of T-cells. If these mechanisms entirely explained the problems in all solid tumours, it would not explain why CAR T-cells have been observed to be effective in mouse models. Instead, it may be that tumour clearance is a 'numbers game', dependent on the number of T-cells that are delivered, extravasate and successfully engage targets, for which our model predicts the maximum rate. Our model was also able to quantify the variability of organ-specific delivery rates within an organism, finding in particular that perfusion-dependent tumour delivery rates are highly variable. The weakness of the model, however, is that these maximum delivery rates do not consider T-cell propensity to extravasate into different organs. The work in Chapter 4 seeks to address this by extending our trafficking model into age-structured PDEs. We found that the transient localisation of T-cells (particularly within lymph nodes) cannot be described by the ODEs, as ODEs implicitly assume that cells are equally likely to leave the compartment regardless of how long they have been there. Previous authors have fit ODE models of T-cell localisation to ODE systems [81–83], but the problems with transient localisation are apparent and they did not measure or model lymph node localisation. The PDE model fitted the localisation data well, and we obtained estimates for the extravasation

probabilities and mean return times of lymphocytes to/from various organs that match biological expectations. This work remains in progress, but will be a subject of future investigations (see Section 5.3).

5.2 Context

I began the discussion of this thesis with the cancer-immunity cycle, Figure 0.4. If the cycle represents the mechanism through which the immune system clears tumours, then it follows that a patient with an observable tumour has a bottleneck somewhere in the cycle. In Chapters 1 and 2, we hypothesised that the bottleneck may be in stages 2 and 3 for a short-peptide vaccination, as we found that certain conditions lead to insufficient presentation of antigen to activate T-cells. We showed that this hypothesis is consistent with observations and the data reported by IMA901. In Chapter 3, we hypothesised that the bottleneck for CAR T-cells targeting solid tumours may actually be stage 4 (insufficient number of cells reaching the tumour) in some cases, and not stages 5-7 (extravasation of T-cells or engagement of tumour cells) as is commonly believed. In Chapter 4, we extended the ODE model into PDEs and used rat lymphocyte localisation data to quantify extravasation probabilities and mean return times for different organs. Use of mouse data with a tumour model, or human data, will allow us to determine how these variables scale between species as we did for Chapter 3, and to quantify when the extravasation of T-cells is the rate-limiting step in the cancer immunity cycle (*i.e.* when the tumour is an 'immune desert', as in Figure 4.0, page 137), rather than the delivery of T-cells (when the tumour is 'immune ignored').

Both halves of the thesis came to a conclusion that problems may have been caused by insufficient numbers of a particular cell type or molecule. Expressed another way, both problems amounted to a question of dosage. Had the vaccination in IMA901 contained fewer peptides, for example, the proportion of dendritic cell MHC-I bound to each peptide would be greater, increasing the 'dose'. Use of the virtual clinical trial framework predicted that this would have greatly improved patient response to those particular peptides. The need to raise responses against multiple peptides could have been addressed by injecting the peptides into different vaccination sites, or by injecting a different subset of peptides each visit. We also found that human trials of CAR T-cells for solid tumours are designed with dosages much less than the equivalents seen in mouse models, and clinical studies with doses equivalent to pre-clinical studies saw greater success. In many scenarios in which dosage is important, clinicians must also balance other concerns such as safety or response rates. For example, too large a number CAR T-cells can result in severe side effects, as described in Chapter 0. However, mathematical modelling allows us to quantify the expected risk

of autoimmunity in different organs (as it would be proportional to delivery rate) or how specific the cells are expected to be for the target organ. Another example is that reducing the number of short peptides in a multi-peptide vaccination reduces the number of possible immune responses, and we cannot predict which peptides (if any) a patient's tumour will present. However, we could use our model to predict changes to vaccine administration (such as the use of multiple vaccine sites) that would reduce peptide competition without reducing the number of peptides given to patients.

A point raised in the discussion of Chapter 2 was that short peptide vaccinations may also fail to raise a population of $CD4^+$ T-cells. If this were the case, then even if a $CD8^+$ response were raised against the peptide, the absence of IL-2 production by $CD4^+$ cells would result in reduced proliferation of the $CD8^+$ cells. This would mean that delivery rates of the cells to the tumour would be lower, and thus may not be enough to overcome the tumour growth rate, as discussed in Section 3.2.5. This is an example of how the various components of the cancer-immune response are linked. Just as we gained insight by considering how T-cell activation in the lymph node depends on the balance of antigenic properties *and* immune cell counts, a combined quantitative analysis of multiple steps of the cancer-immunity cycle would likely yield greater insight than treatment of the steps in isolation. For this reason, further work on the trafficking model will consider both the delivery and extravasation of T-cells (steps 4 and 5) simultaneously. This, too, is an advantage of mathematical modelling: it is possible to develop mechanistic models across multiple scales and locations in the body, supported by experimental study. It would be more difficult to investigate if and why the use of short peptides over other approaches to cancer vaccination would influence the size of the generated T-cell response through experimental study alone.

One of the motivations behind this DPhil was to address the question, "why do immunotherapies fail in the clinic?" There are at least two ways to consider this question. One is that clinical outcomes are often highly heterogeneous, both between patients – where a given therapy may fail to achieve complete remission in most patients – and within patients. A single patient may have multiple tumour sites, every one of which the immune system must eradicate before the patient is tumour-free. This means that therapy must be enough to drive tumour regression at the site with the lowest delivery rate, the poorest interstitial pressure, the lowest probability of extravasation and possibly with tumour antigen that differ from cells at another site.

Another way to consider this question is "why do immunotherapies fail in the clinic, after they have succeeded in mice?" Mouse models often lack the heterogeneity present in

patients. Experimental mice come from highly inbred strains, and those in a single experiment may be nearly identical. Their tumours are grown quickly by known experimental protocols, rather than very slowly over years in patients due to a variety of mechanisms. They thus have not only much less intra-tumour genetic diversity than tumours in patients, but will be qualitatively similar among experimental mice. Mouse tumours range from a few millimetres to a few tens of millimetres, whereas human tumours they range from 1–200mm in diameter. Though these are similar sizes relative to the organism, this coupled with the differences in growth times means that mouse tumours will be structured fundamentally differently. Furthermore, mouse tumours are often grown subcutaneously and/or as xenografts, a location that would exert much less pressure on a tumour than would a dense, structured organ. Under these conditions, it is unlikely that immune cells would be excluded from the core of a small, murine tumour unless there is a mechanism actively preventing immune infiltration, whereas in humans, the size of a tumour and the interstitial pressure can prevent infiltration.

Each study of this thesis has revealed problems or outcomes that would differ between mice and humans. In Chapter 1, it was found that the probability of T-cell activation depends on pMHC-I off-rates, but that the 'transition' in the probability from one to zero can be shifted by changing the amount of time it takes dendritic cells to migrate from the vaccination site to the lymph nodes. It typically takes 20–30 hours for dendritic cells to migrate from the human dermis to a draining lymph node (Section 1.0). In mice, in which cells are expected to crawl at a similar velocity as in humans but in which the distance from dermis to lymph node is much shorter, this time would be 12–18 [12] hours or even as low as four hours [17]. The results of Chapter 2 showed that a modest improvement in migration time greatly improves predicted patient responses in a virtual clinical trial. We would hence predict that a pre-clinical trial of IMA901 would have been highly successful based on reduced timescales alone. In addition, mice tend to receive much larger injection volumes relative to those given to humans, likely increasing the numbers of DCs recruited, and experimental mice and their tumours are, as discussed above, much less heterogeneous than renal cell carcinoma would be in patients. In Chapter 3, it was found that typical clinical dosages of CAR T-cells are much less than the equivalent doses used in pre-clinical studies. The lower incidence of problems such as poor interstitial pressure and the homogeneity among individual mice and their tumours would compound to further improve the efficacy of pre-clinical studies of CAR T-cells over clinical trials.

These problems show that great care is needed to understand, predict and handle the differences between mouse models and humans. The differences between the predicted mouse and human responses to the short peptide vaccines and CAR T-cells discussed in this thesis

could explain the limited efficacy seen in clinical studies. These differing responses were not due to differences in human and mouse biochemistry or immunology, but rather due to a difference of scale. Consideration of different scales with *in silico* models like those presented in this thesis would facilitate the translation and scaling of pre-clinical results to the clinic. Likewise, quantification of the expected variation of outcomes among different patients can be easily achieved *in silico*, as shown in Chapters 2 and 3. The relative efficacy of a treatment at different tumour sites based on the anatomy can also be considered easily *in silico*, as shown in Chapter 3.

The work of this thesis is an example of how quantification of processes that we already understand can yield further insight, and how mathematical biology can contribute to experimental and clinical studies. There are many examples of models in the published literature that have generated biological insight [57]. However, most models receive little attention and very few clinicians would use models that make clinical predictions in decision making. Why are so few published models used, or conversely, why are so few models useful? Mechanistic models are often produced by researchers who are not, or do not have strong collaborative links with, experimentalists and clinicians. Often, such models are developed in isolation from the clinic and do not address specific questions, in which case they cannot be verified by experimental study. Conversely, data reported by experimental studies are often not easily applicable to models, due to a lack of quantitative information or data spanning short and long timescales, and due to the reporting of multiple patients' results with uncontrolled variability. Similarly, many human parameters required by models are not reported in experimental literature, so either parameter values that are estimated (educated guess) or based on values from mouse experiments (which are also models) are used, leading to predictions with unknown relevance to human disease.

To remedy this, we have previously argued [57] that mechanistic modelling must be more deeply integrated into experimental workflows. Models should be question-driven, taking the form of a verifiable *in silico* experiment, and their inputs and outputs must align with clinically or experimentally measurable assessments and endpoints, so that they generate biologically meaningful predictions. In turn, human parameters need to be measured and experimentalists should endeavour to report quantitative information that aids mechanistic analysis by a model. Experimentalists and modellers should endeavour to design complementary experimental and modelling studies, where data from each study can support the other and produce something greater than the sum of its parts. We endeavoured to integrate the work in this thesis with known literature and to generate biologically meaningful hypotheses, but to go further, we should develop our models in tandem with experimental work.

5.3 Future work

All of the chapters of this thesis may be further developed, and all can be done so in collaboration with experimental or clinical studies. The analysis in Chapters 1 and 2 revealed that the assumption that short peptides are not taken up and processed by dendritic cells intracellularly might be postulated to be false, as the model is too sensitive to related parameters. Neil Dalchau (Microsoft Research), Tim Elliott (University of Southampton) and colleagues have previously produced a model of intracellular TAP and MHC-I processing which may be incorporated within our model to address these problems and improve predictions. Tim Elliott's lab has the facilities to perform further experiments to test *in silico* hypotheses or to generate data for an *in silico* study to test; it is this lab that measured the peptide off-rates and MHC-I turnover for the work of Chapter 2. This work is being considered for a post-doctoral project in collaboration with IMV Inc, to consider how an oil-based vaccination affects the probability of T-cell activation. This project would benefit from experimental and clinical data from the company, for example concerning the persistence of antigen and antigen presenting cells within the dermis. This too allows 'cross-talk' between experimental and theoretical study, to further the insight generated by both. The lymph node model could also be adapted to facilitate prediction of the lymph node output for alignment with measurable biomarkers, by including effects such as inflammation or ingress and egress of cells.

There is a plethora of possible next steps for the work presented in Chapters 3 and 4. From a technical viewpoint, we wish to re-implement the system using the method of lines, so that the system is amenable to a variety of off-the-shelf ODE solvers. We wish to determine the timescales or parameter regimes in which an ODE can acceptably describe the dynamics of T-cell localisation, so that the full PDE model is not required. Having fit the ODE and PDE models to the lymphocyte localisation data in rats published by Smith *et al* [37], we aim to match to T-cell specific data that has been previously published [81–83] (with fewer organs) or to data of T-cell localisation within a mouse tumour model produced by Professor Ann Ager (Cardiff University). We have worked closely with Ann Ager throughout the DPhil and further work with this data will facilitate co-development of the experimental and theoretical studies. A mechanistic model of T-cell extravasation, rather than quantification by a probability e, could also be incorporated into the trafficking model to generate further insight. A model of tumour growth could be included to refine the investigation of the balance between T-cell extravasation, proliferation, tumour killing and tumour growth, and would allow us to determine when steps 5-7 of the cancer immunity cycle are the bottleneck, rather than step 4. The extension of Chapters 3 and 4 will form

the basis of my post-doctoral fellowship over the next three years, in partnership with a company that produces CAR T-cells, Celgene. Such a relationship will provide clinical insight, as my relationship with Roche has provided over the course of the DPhil, and may also provide clinical and experimental data to further model development. We will continue to work with collaborators such as Ann Ager and Tim Elliott, which could facilitate the generation of experimental work in tandem with our modelling work. We will attempt to produce models of CAR T-cell clearance of lymphomas, leukaemias and solid tumours in mice and humans, in an attempt to elucidate both the mechanisms impeding tumour clearance and possible therapeutic windows for CAR T-cells.

In this thesis, I have presented mechanistic models of T-cell activation in the lymph node and of T-cell trafficking across the body. In both cases, I have related the model to immunotherapies, quantified the affect of model parameters on therapeutic outcome, and generated hypotheses for the mechanisms that limit efficacy. I believe that such an approach is key for presenting mathematical biology not just as a tool to analyse discarded experimental data, but as a methodology that can generate novel hypotheses in its own right, as many mathematical biology studies have done before. Such an approach can be deeply integrated within experimental and clinical studies, and there are already many research facilities that have theoretical groups incorporated within them (using oncology as an example: The Moffitt Cancer Center, Florida, and the Centre for Cancer Immunology, Southampton). With data-driven models already accepted as supporting information by the US Food and Drug Administration and European Medicines Agency (for example), we may soon begin to see mechanistic modelling at the forefront of hypothesis-driven immunological and oncological research.

Bibliography

[1] O.J. Finn. Immuno-oncology: understanding the function and dysfunction of the immune system in cancer. *Annals of Oncology*, 23(Suppl 8):viii6–viii9, 2012. ISSN 0092-8674. doi: 10.1093/annonc/mds256.

[2] Douglas Hanahan and Robert A. Weinberg. Hallmarks of Cancer: The Next Generation. *Cell*, 144(5):646 – 674, 2011. ISSN 0092-8674. doi: 10.1016/j.cell.2011.02.013.

[3] Kheng Newick, Edmund Moon, and Steven M. Albelda. Chimeric antigen receptor T-cell therapy for solid tumors. *Molecular Therapy - Oncolytics*, 3(October 2015): 16006, 2016. ISSN 23727705. doi: 10.1038/mto.2016.6.

[4] Praveen Sridhar and Fabio Petrocca. Regional delivery of chimeric antigen receptor (CAR) T-cells for cancer therapy. *Cancers*, 9(7):1–10, 2017. ISSN 20726694. doi: 10.3390/cancers9070092.

[5] Daniel S. Chen and Ira Mellman. Oncology meets immunology: The cancer-immunity cycle. *Immunity*, 39(1):1–10, 2013. ISSN 10747613. doi: 10.1016/j.immuni.2013.07.012.

[6] Lauren Sompayrac. *How the Immune System Works*. Wiley-Blackwell, 4th edition, 2014. ISBN 978-0-4706-5729-4.

[7] Alberts, Johnson, Lewis, Morgan, Raff, Roberts, and Walter. *Molecular Biology of the Cell*. Garland Science, 6th edition, 2014. ISBN 9780815344322.

[8] Morten M. Nielsen, Deborah A. Witherden, and Wendy L. Havran. $\gamma\delta$ T cells in homeostasis and host defence of epithelial barrier tissues. *Nature Reviews Immunology*, 17(12):733–745, 2017. ISSN 14741741. doi: 10.1038/nri.2017.101.

[9] David Redmond, Asaf Poran, and Olivier Elemento. Single-cell TCRseq: paired recovery of entire T-cell alpha and beta chain transcripts in T-cell receptors from single-cell RNAseq. *Genome Medicine*, 8(1):80, Jul 2016. ISSN 1756-994X. doi: 10.1186/s13073-016-0335-7.

[10] Jonathan W Yewdell, Eric Reits, and Jacques Neefjes. Making sense of mass destruction: quantitating MHC class I antigen presentation. *Nature reviews. Immunology*, 3 (12):952–961, 2003. ISSN 1474-1733. doi: 10.1038/nri1250.

[11] Alan S. Perelson and Frederik W. Wiegel. Scaling aspects of lymphocyte trafficking. *Journal of Theoretical Biology*, 257(1):9–16, 2009. ISSN 00225193. doi: 10.1016/j.jtbi.2008.11.007.

[12] Philippe Bousso. T-cell activation by dendritic cells in the lymph node: lessons from the movies. *Nature reviews. Immunology*, 8(9):675–684, 2008. ISSN 1474-1733. doi: 10.1038/nri2379.

[13] Michael D Cahalan and Ian Parker. Choreography of cell motility and interaction dynamics imaged by two-photon microscopy in lymphoid organs. *Annual Review of Immunology*, 26:585–626, 2008. ISSN 0732-0582. doi: 10.1146/annurev.immunol.24.021605.090620.

[14] A J Young. The physiology of lymphocyte migration through the single lymph node in vivo. *Seminars in immunology*, 11(2):73–83, 1999. ISSN 1044-5323. doi: 10.1006/smim.1999.0163.

[15] Lois L Cavanagh and Wolfgang Weninger. Dendritic cell behaviour in vivo: lessons learned from intravital two-photon microscopy. *Immunology and Cell Biology*, 86(5):428–438, 2008. ISSN 0818-9641. doi: 10.1038/icb.2008.25.

[16] A Martin-Fontecha, A Martin-Fontecha, A Lanzavecchia, A Lanzavecchia, F Sallusto, and F Sallusto. Dendritic cell migration to peripheral lymph nodes. *Handb Exp Pharmacol*, 188:31–49, 2009. doi: 10.1007/978-3-540-71029-5_2.

[17] Mark J Miller, Arsalan S Hejazi, Sindy H Wei, Michael D Cahalan, and Ian Parker. T cell repertoire scanning is promoted by dynamic dendritic cell behavior and random T cell motility in the lymph node. *Proceedings of the National Academy of Sciences of the United States of America*, 101(4):998–1003, 2004. ISSN 0027-8424. doi: 10.1073/pnas.0306407101.

[18] Mark J Miller, Olga Safrina, Ian Parker, and Michael D Cahalan. Imaging the single cell dynamics of CD4+ T cell activation by dendritic cells in lymph nodes. *The Journal of Experimental Medicine*, 200(7):847–856, 2004. ISSN 0022-1007. doi: 10.1084/jem.20041236. See also other data from same author.

[19] Susanna Celli, Mark Day, Andreas J Mu, Carmen Molina-paris, Grant Lythe, and Philippe Bousso. Brief report: How many dendritic cells are required to initiate a T-cell response? *Blood*, 120(19):3945–3948, 2012. doi: 10.1182/blood-2012-01-408260.

[20] Thomas Riggs, Adrienne Walts, Nicolas Perry, Laura Bickle, Jennifer N Lynch, Joanne Flynn, Jennifer J Linderman, Mark J Miller, and Denise E Kirschner. A comparison of random vs. chemotaxis driven contacts of T cells with dendritic cells during repetoire scanning. *Journal of Theoretical Biology*, 250(4):732–751, 2009. doi: 10.1016/j.jtbi.2007.10.015.A.

[21] Joanna R Groom. Moving to the suburbs: T-cell positioning within lymph nodes during activation and memory. *Immunology and Cell Biology*, 93(4):330–336, 2015. ISSN 0818-9641. doi: 10.1038/icb.2015.29.

[22] Joost B Beltman, Athanasius F M Marée, Jennifer N Lynch, Mark J Miller, and Rob J de Boer. Lymph node topology dictates T cell migration behavior. *The Journal of experimental medicine*, 204(4):771–780, 2007. ISSN 0022-1007. doi: 10.1084/jem.20061278.

[23] Hélène D. Moreau and Philippe Bousso. Visualizing how T cells collect activation signals in vivo. *Current Opinion in Immunology*, 26(1):56–62, 2014. ISSN 09527915. doi: 10.1016/j.coi.2013.10.013.

[24] Michael L. Dustin. T-cell activation through immunological synapses and kinapses. *Immunological Reviews*, 221(1):77–89, 2008. ISSN 01052896. doi: 10.1111/j.1600-065X.2008.00589.x.

[25] Alexandre P. Benechet, Manisha Menon, and Kamal M. Khanna. Visualizing T cell migration in situ. *Frontiers in Immunology*, 5(JUL):1–12, 2014. ISSN 16643224. doi: 10.3389/fimmu.2014.00363.

[26] Thorsten R Mempel, Sarah E Henrickson, and Ulrich H Von Andrian. T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature*, 427(6970):154–159, 2004. ISSN 0028-0836. doi: 10.1038/nature02238.

[27] Paulus Mrass, Janka Petravic, Miles P. Davenport, and Wolfgang Weninger. Cell-autonomous and environmental contributions to the interstitial migration of T cells. *Seminars in Immunopathology*, 32(3):257–274, 2010. ISSN 18632297. doi: 10.1007/s00281-010-0212-1.

[28] Edward J. Banigan, Tajie H. Harris, David A. Christian, Christopher A. Hunter, and Andrea J. Liu. Heterogeneous CD8+ T Cell Migration in the Lymph Node in the Absence of Inflammation Revealed by Quantitative Migration Analysis. *PLOS Computational Biology*, 11(2):e1004058, 2015. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1004058.

[29] Wolfgang Weninger, Maté Biro, and Rohit Jain. Leukocyte migration in the interstitial space of non-lymphoid organs. *Nature reviews. Immunology*, 14(4):232–46, 2014. ISSN 1474-1741. doi: 10.1038/nri3641.

[30] Aleksandra J. Ozga, Federica Moalli, Jun Abe, Jim Swoger, James Sharpe, Dietmar Zehn, Mario Kreutzfeldt, Doron Merkler, Jorge Ripoll, and Jens V. Stein. pMHC affinity controls duration of CD8+ T cell–DC interactions and imprints timing of effector differentiation versus expansion. *Journal of Experimental Medicine*, 213(12): 2811–2829, 2016. ISSN 0022-1007. doi: 10.1084/jem.20160206.

[31] Frederik Graw and Roland R. Regoes. Influence of the fibroblastic reticular network on cell-cell interactions in lymphoid organs. *PLoS Computational Biology*, 8(3), 2012. ISSN 1553734X. doi: 10.1371/journal.pcbi.1002436.

[32] Allan McI Mowat. Anatomical basis of tolerance and immunity to intestinal antigens. *Nature Reviews Immunology*, 3:331–341, 2003. doi: 10.1038/nri1057.

[33] Reina E. Mebius and Georg Kraal. Structure and function of the spleen. *Nature Reviews Immunology*, 5:606–616, 2005. doi: 10.1038/nri1669.

[34] Shiv Pillai and Annaiah Cariappa. The follicular versus marginal zone B lymphocyte cell fate decision. *Nature Reviews Immunology*, 9:767–777, 2009. doi: 10.1038/nri2656.

[35] Ann Ager, H Angharad Watson, Sophie C Wehenkel, and Rebar N Mohammed. Homing to solid cancers: a vascular checkpoint in adoptive cell therapy using CAR T-cells. *Biochemical Society transactions*, 44(2):377–85, 2016. ISSN 1470-8752. doi: 10.1042/BST20150254.

[36] Hongmei Fu, Amu Wang, Claudio Mauro, and Federica Marelli-Berg. T lymphocyte trafficking: molecules and mechanisms. *Frontiers in bioscience (Landmark edition)*, 18(January):422–440, 2013. ISSN 10939946. doi: 10.2741/4111.

[37] M. E. Smith, A. F. Martin, and W. L. Ford. Migration of Lymphoblasts in the Rat. In *Essays of the Anatomy and Physiology of Lymphoid Tissues*, volume 16, pages 203–232. Thür AG Offsetdruck, Pratteln, 1980.

[38] Zamaneh Mikhak, James P Strassner, and Andrew D Luster. Lung dendritic cells imprint T cell lung homing and promote lung immunity through the chemokine receptor CCR4. *The Journal of experimental medicine*, 210(9):1855–69, 2013. ISSN 1540-9538. doi: 10.1084/jem.20130091.

[39] Jamie A. Lopez, Amelia J. Brennan, James C. Whisstock, Ilia Voskoboinik, and Joseph A. Trapani. Protecting a serial killer: pathways for perforin trafficking and self-defence ensure sequential target cell death. *Trends in Immunology*, 33(8):406 – 412, 2012. ISSN 1471-4906. doi: https://doi.org/10.1016/j.it.2012.04.001.

[40] Aurelie Wiedemann, David Depoil, Mustapha Faroudi, and Salvatore Valitutti. Cytotoxic T lymphocytes kill multiple targets simultaneously via spatiotemporal uncoupling of lytic and stimulatory synapses. *Proceedings of the National Academy of Sciences*, 103(29):10985–10990, 2006. ISSN 0027-8424. doi: 10.1073/pnas.0600651103.

[41] Stephan Halle, Kirsten Anja Keyser, Felix Rolf Stahl, Andreas Busche, Anja Marquardt, Xiang Zheng, Melanie Galla, Vigo Heissmeyer, Katrin Heller, Jasmin Boelter, Karen Wagner, Yvonne Bischoff, Rieke Martens, Asolina Braun, Kathrin Werth, Alexey Uvarovskii, Harald Kempf, Michael Meyer-Hermann, Ramon Arens, Melanie Kremer, Gerd Sutter, Martin Messerle, and Reinhold Förster. In Vivo Killing Capacity of Cytotoxic T Cells Is Limited and Involves Dynamic Interactions and T Cell Cooperativity. *Immunity*, pages 233–245, 2016. ISSN 10747613. doi: 10.1016/j.immuni.2016.01.010.

[42] Zilton Vasconcelos, Sabina Müller, Delphine Guipouy, Wong Yu, Claire Christophe, Sébastien Gadat, Salvatore Valitutti, and Loïc Dupré. Individual Human Cytotoxic T Lymphocytes Exhibit Intraclonal Heterogeneity during Sustained Killing. *Cell Reports*, 11(9):1474–1485, 2015. ISSN 22111247. doi: 10.1016/j.celrep.2015.05.002.

[43] Andrew M. Intlekofer and Craig B. Thompson. At the Bench: Preclinical rationale for CTLA-4 and PD-1 blockade as cancer immunotherapy. *Journal of Leukocyte Biology*, 94(1):25–39, 2013. doi: 10.1189/jlb.1212621.

[44] E Valk, CE Rudd, and H Schneider. CTLA-4 trafficking and surface expression. *Trends in immunology*, 29(6):272–279, 2008. doi: 10.1016/j.it.2008.02.011.

[45] Kristen E. Pauken and E. John Wherry. Overcoming T cell exhaustion in infection and cancer. *Trends in Immunology*, 36(4):265–276, 2015. ISSN 14714981. doi: 10.1016/j.it.2015.02.008.

[46] Carl-Henrik Heldin, Kristofer Rubin, Kristian Pietras, and Arne Östman. High interstitial fluid pressure — an obstacle in cancer therapy. *Nature Reviews Cancer*, 4: 806–813, 2004. doi: 10.1038/nrc1456.

[47] Stephane Ferretti, Peter R Allegrini, , Mike M Becquet, and Paul Mj McSheehy. Tumor interstitial fluid pressure as an early-response marker for anticancer therapeutics. *Neoplasia*, 11(9):874–881, 2009.

[48] David Eastwood, Chris Bird, Paula Dilger, Jason Hockley, Lucy Findlay, Stephen Poole, Susan J. Thorpe, Meenu Wadhwa, Robin Thorpe, and Richard Stebbings. Severity of the TGN1412 trial disaster cytokine storm correlated with IL-2 release. *British Journal of Clinical Pharmacology*, 76(2):299–315, 2013. doi: 10.1111/bcp.12165.

[49] Carlos L. Arteaga, Peter C. Adamson, Jeffrey A. Engelman, Margaret Foti, Richard B. Gaynor, Susan G. Hilsenbeck, Paul J. Limburg, Scott W. Lowe, Elaine R. Mardis, Scott Ramsey, Timothy R. Rebbeck, Andrea L. Richardson, Eric H. Rubin, and George J. Weiner. AACR Cancer Progress Report 2014. *Clinical Cancer Research*, 20(19 Supplement):S1–S112, 2014. ISSN 1078-0432. doi: 10.1158/1078-0432.CCR-14-2123.

[50] NCT01118091. Steven Rosenberg *et al.* Prospective Randomized Comparative Study of Cell Transfer Therapy Using Short-Term Cultured Anti-Tumor Autologous Lymphocytes Following a Non-Myeloablative Lymphocyte Depleting Chemotherapy Regimen Compared to High-dose Aldesleukin in Metastatic Melanoma. *ClinicalTrials.gov*, trial NCT01118091, April 2010. Retrieved October 2018.

[51] NCT01082926. City of Hope Medical Center. Phase I Study of Cellular Immunotherapy for Recurrent/Refractory Malignant Glioma Using Intratumoral Infusions of GRm13Z40-2, An Allogeneic CD8+ Cytolitic T-Cell Line Genetically Modified to Express the IL 13-Zetakine and HyTK and to be Resistant to Glucocorticoids, in Combination With Interleukin-2. *ClinicalTrials.gov*, trial NCT01082926, March 2010. Retrieved October 2018.

[52] Steffen Walter et al. Multipeptide immune response to cancer vaccine IMA901 after single-dose cyclophosphamide associates with longer patient survival. *Nature Medicine*, 18(8):1254–1261, 2012. ISSN 1078-8956. doi: 10.1038/nm.2883.

[53] Brian I Rini, Arnulf Stenzl, Romauld Zdrojowy, Mikhail Kogan, Mikhail Shkolnik, Stephane Oudard, Steffen Weikert, Sergio Bracarda, Simon J Crabb, Jens Bedke, Joerg Ludwig, Dominik Maurer, Regina Mendrzyk, Claudia Wagner, Andrea Mahr, Jens Fritsche, Toni Weinschenk, Steffen Walter, Alexandra Kirner, Harpreet Singh-Jasuja, Carsten Reinhardt, and Tim Eisen. IMA901, a multipeptide cancer vaccine, plus sunitinib versus sunitinib alone, as first-line therapy for advanced or

metastatic renal cell carcinoma (IMPRINT): a multicentre, open-label, randomised, controlled, phase 3 trial. *The Lancet Oncology*, 17(11):1599 – 1611, 2016. doi: 10.1016/S1470-2045(16)30408-9.

[54] Jessica Hartmann, Martina Schüßler-Lenz, Attilio Bondanza, and Christian J Buchholz. Clinical development of CAR T cells—challenges and opportunities in translating innovative treatment concepts. *EMBO Molecular Medicine*, 9(9):e201607485, 2017. ISSN 1757-4676. doi: 10.15252/emmm.201607485.

[55] Charlotte Graham, Rebecca Hewitson, Antonio Pagliuca, and Reuben Benjamin. Cancer immunotherapy with CAR-T cells - Behold the future. *Clinical Medicine, Journal of the Royal College of Physicians of London*, 18(4):324–328, 2018. ISSN 14734893. doi: 10.7861/clinmedicine.18-4-324.

[56] Hagop Kantarjian, Anthony Stein, Nicola Gökbuget, Adele K. Fielding, Andre C. Schuh, Josep-Maria Ribera, Andrew Wei, Hervé Dombret, Robin Foà, Renato Bassan, Önder Arslan, Miguel A. Sanz, Julie Bergeron, Fatih Demirkan, Ewa Lech-Maranda, Alessandro Rambaldi, Xavier Thomas, Heinz-August Horst, Monika Brüggemann, Wolfram Klapper, Brent L. Wood, Alex Fleishman, Dirk Nagorsen, Christopher Holland, Zachary Zimmerman, and Max S. Topp. Blinatumomab versus chemotherapy for advanced acute lymphoblastic leukemia. *New England Journal of Medicine*, 376 (9):836–847, 2017. doi: 10.1056/NEJMoa1609783. PMID: 28249141.

[57] L.V. Brown, E.A. Gaffney, J. Wagg, and M.C. Coles. Applications of mechanistic modelling to clinical and experimental immunology: an emerging technology to accelerate immunotherapeutic discovery and development. *Clinical & Experimental Immunology*, 193(3):284–292, 2018. doi: 10.1111/cei.13182.

[58] R A Gatenby and E T Gawlinski. A reaction-diffusion model of cancer invasion. *Cancer research*, 56(31):5745–5753, 1996. ISSN 0008-5472.

[59] Curtis A. Gravenmier, Miriam Siddique, and Robert A. Gatenby. Adaptation to Stochastic Temporal Variations in Intratumoral Blood Flow: The Warburg Effect as a Bet Hedging Strategy. *Bulletin of Mathematical Biology*, pages 1–17, 2017. ISSN 15229602. doi: 10.1007/s11538-017-0261-x.

[60] Robert A Gatenby, Ariosto S Silva, Robert J Gillies, and B Roy Frieden. Adaptive therapy. *Cancer Research*, 69(11):4894–4903, 2009. ISSN 00085472. doi: 10.1158/0008-5472.CAN-08-3658.

[61] Kevin Leder, Ken Pitter, Quincey Laplant, Dolores Hambardzumyan, Brian D Ross, Timothy A Chan, Eric C Holland, and Franziska Michor. Mathematical Modeling of PDGF-Driven Glioblastoma Reveals Optimized Radiation Dosing Schedules. *Cell*, 156(3):603–616, 2014. ISSN 00928674. doi: 10.1016/j.cell.2013.12.029.

[62] Laurent Claret, Pascal Girard, Paulo M. Hoff, Eric Van Cutsem, Klaas P. Zuideveld, Karin Jorga, Jan Fagerberg, and René Bruno. Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics. *Journal of Clinical Oncology*, 27(25):4103–4108, 2009. ISSN 0732183X. doi: 10.1200/JCO.2008.21.0807.

[63] Larry Norton and Richard Simon. Growth curve of an experimental solid tumor following radiotherapy. *Journal of the National Cancer Institute*, 58(6):1735–1741, 1977. ISSN 14602105. doi: 10.1093/jnci/58.6.1735.

[64] Laurent Claret, Jin Y. Jin, Charles Ferte, Helen Winter, Sandhya Girish, Mark Stroh, Pei He, Marcus Ballinger, Alan Sandler, Amita Joshi, Achim Rittmeyer, David Gandara, Jean Charles Soria, and Rene Bruno. A model of overall survival predicts treatment outcomes with atezolizumab versus chemotherapy in non–small cell lung cancer based on early tumor kinetics. *Clinical Cancer Research*, 24(14):3292–3298, 2018. ISSN 15573265. doi: 10.1158/1078-0432.CCR-17-3662.

[65] Veronica Estrella, Tingan Chen, Mark Lloyd, Jonathan Wojtkowiak, Heather H. Cornnell, Arig Ibrahim-Hashim, Kate Bailey, Yoganand Balagurunathan, Jennifer M. Rothberg, Bonnie F. Sloane, Joseph Johnson, Robert A. Gatenby, and Robert J. Gillies. Acidity generated by the tumor microenvironment drives local invasion. *Cancer Research*, 73(5):1524–1535, 2013. ISSN 00085472. doi: 10.1158/0008-5472.CAN-12-2796.

[66] Xavier Elharrar, Dominique Barbolosi, Joseph Ciccolini, Christophe Meille, Christian Faivre, Bruno Lacarelle, Nicolas Andre, and Fabrice Barlesi. A phase Ia/Ib clinical trial of metronomic chemotherapy based on a mathematical model of oral vinorelbine in metastatic non-small cell lung cancer and malignant pleural mesothelioma: rationale and study protocol. *BMC cancer*, 16(1):278, 2016. ISSN 1471-2407 (Electronic). doi: 10.1186/s12885-016-2308-z.

[67] Goldie and Coldman. *Drug Resistance in Cancer: Mechanisms and Models*. Cambridge University Press, 1998. ISBN 978-0521482738.

[68] HC Monro and EA Gaffney. Modelling chemotherapy resistance in palliation and failed cure. *J. Theoretical Biology*, 257:292–302, 2009. doi: 10.1016/j.jtbi.2008.12.006.

[69] Jacob G Scott, David Basanta, Alexander R A Anderson, Philip Gerlee, and Jacob G Scott. A mathematical model of tumour self-seeding reveals secondary metastatic deposits as drivers of primary tumour growth. *Journal of The Royal Society Interface*, 10(20130011), 2013. doi: 10.1098/rsif.2013.0011.

[70] Hélène D Moreau, Gib Bogle, and Philippe Bousso. A virtual lymph node model to dissect the requirements for T cell activation by synapses and kinapses. *Immunology and cell biology*, 94(November 2015):680–688, 2016. ISSN 1440-1711. doi: 10.1038/icb.2016.36.

[71] Gib Bogle and P Rod Dunbar. Agent-based simulation of T-cell activation and proliferation within a lymph node. *Immunology and Cell Biology*, 88(2):172–179, 2010. ISSN 0818-9641. doi: 10.1038/icb.2009.78.

[72] Chang Gong, Joshua T. Mattila, Mark Miller, JoAnne L. Flynn, Jennifer J. Linderman, and D. Kirschner. Predicting lymph node output efficiency using systems biology. *Journal of Theoretical Biology*, 335:169–184, 2013. ISSN 00225193. doi: 10.1016/j.jtbi.2013.06.016.

[73] Anass Bouchnita, Gennady Bocharov, Andreas Meyerhans, and Vitaly Volpert. Hybrid approach to model the spatial regulation of T cell responses. *BMC Immunology*, 18(Suppl 1), 2017. ISSN 14712172. doi: 10.1186/s12865-017-0205-0.

[74] Liam V. Brown, Eamonn A. Gaffney, Jonathan Wagg, and Mark C. Coles. An in silico model of cytotoxic T-lymphocyte activation in the lymph node following short peptide vaccination. *Journal of The Royal Society Interface*, 15(140), 2018. ISSN 1742-5689. doi: 10.1098/rsif.2018.0041.

[75] Teerachat Saeheng, Kesara Na-Bangchang, and Juntra Karbwang. Utility of physiologically based pharmacokinetic (PBPK) modeling in oncology drug development and its accuracy: a systematic review. *European Journal of Clinical Pharmacology*, July 2018. ISSN 0031-6970. doi: 10.1007/s00228-018-2513-6.

[76] T Teorell. Kinetics of distribution of substances administered to the body. I. The extravascular modes of administration. *Arch. Int. Pharmacodyn. Ther*, 57:205–225, 1937.

[77] H. M. Jones, Y. Chen, C. Gibson, T. Heimbach, N. Parrott, S. A. Peters, J. Snoeys, V. V. Upreti, M. Zheng, and S. D. Hall. Physiologically based pharmacokinetic modeling in drug discovery and development: a pharmaceutical industry perspective. *Clinical pharmacology and therapeutics*, 97(3):247–262, 2015. ISSN 15326535. doi: 10.1002/cpt.37.

[78] Rakesh Jain. Delivery of molecular and cellular medicine to solid tumors. *Advanced Drug Delivery Reviews*, 29(6):997–1003, 2012. ISSN 15378276. doi: 10.1016/j.biotechadv.2011.08.021.Secreted.

[79] K Yoshida, N Budha, and J. Jin. Impact of Physiologically Based Pharmacokinetic Models on Regulatory Reviews and Product Labels: Frequent Utilization in the Field of Oncology. *Clinical Pharmacology and Therapeutics*, 101(5):597–602, 2017. doi: 10.1002/cpt.622.

[80] Amit Garg and Joseph P. Balthasar. Physiologically-based pharmacokinetic (PBPK) model to predict IgG tissue kinetics in wild-type and FcRn-knockout mice. *Journal of Pharmacokinetics and Pharmacodynamics*, 34(5):687–709, 2007. ISSN 1567567X. doi: 10.1007/s10928-007-9065-1.

[81] Hui Zhu, Robert J Melder, Laurence T Baxter, and Rakesh K Jain. Physiologically Based Kinetic Model of Effector Cell Biodistribution in Mammals: Implications for Adoptive Immunotherapy. (617):3771–3781, 1996.

[82] Robert J Melder, Lance L Munn, Brian R Stoll, Edgardo M Marecos, Laurence T Baxter, Ralph Weissleder, and Rakesh K Jain. Systemic distribution and tumor localization of adoptively transferred lymphocytes in mice: comparison with physiologically based pharmacokinetic model. *Neoplasia*, 4(1):3–8, 2002. ISSN 1522-8002. doi: 10.1038/sj/neo/7900209.

[83] Stuart W. Friedrich, Stephany C. Lin, Brian R. Stoll, Laurence T. Baxter, Lance L. Munn, and Rakesh K. Jain. Antibody-directed effector cell therapy of tumors: Analysis and optimization using a physiologically based pharmacokinetic model. *Neoplasia*, 4(5):449–463, 2002. ISSN 15228002. doi: 10.1038/sj.neo.7900260.

[84] Jan T. Poleszczuk, Kimberly A. Luddy, Sotiris Prokopiou, Mark Robertson-Tessi, Eduardo G. Moros, Mayer Fishman, Julie Y. Djeu, Steven E. Finkelstein, and Heiko Enderling. Abscopal benefits of localized radiotherapy depend on activated T-cell trafficking and distribution between metastatic lesions. *Cancer Research*, 76(5):1009–1018, 2016. ISSN 15387445. doi: 10.1158/0008-5472.CAN-15-1423.

[85] Rachel Walker, Jan Poleszczuk, Shari Pilon-Thomas, Sungjune Kim, Alexander A.R.A. Anderson, Brian J. Czerniecki, Louis B. Harrison, Eduardo G. Moros, and Heiko Enderling. Immune interconnectivity of anatomically distant tumors as a potential mediator of systemic responses to local therapy. *Scientific Reports*, 8(1): 1–11, 2018. ISSN 20452322. doi: 10.1038/s41598-018-27718-1.

[86] Haolu Wang, Xiaowen Liang, Zhi Ping Xu, Darrell H G Crawford, Xin Liu, and Michael S Roberts. A physiologically based kinetic model for elucidating the in vivo distribution of administered mesenchymal stem cells. *Nature Publishing Group*, (February):1–12, 2016. ISSN 20452322. doi: 10.1038/srep22293.

[87] Eamonn A Gaffney. The application of mathematical modelling to aspects of adjuvant chemotherapy scheduling. *Journal of Mathematical Biology*, 48:375–422, 2004.

[88] Wojciech Krzyzanski. Pharmacodynamic models of age-structured cell populations. *Journal of Pharmacokinetics and Pharmacodynamics*, 42(5):573–589, 2015. ISSN 15738744. doi: 10.1007/s10928-015-9446-9.

[89] G. Chiorino, J. A.J. Metz, D. Tomasoni, and P. Ubezio. Desynchronization rate in cell populations: Mathematical modeling and experimental data. *Journal of Theoretical Biology*, 208(2):185–199, 2001. ISSN 00225193. doi: 10.1006/jtbi.2000.2213.

[90] Bambang S. Adiwijaya, Eva Herrmann, Brian Hare, Tara Kieffer, Chao Lin, Ann D. Kwong, Varun Garg, John C.R. Randle, Christoph Sarrazin, Stefan Zeuzem, and Paul R. Caron. A multi-variant, viral dynamic model of genotype 1 HCV to assess the in vivo evolution of protease-inhibitor resistant variants. *PLoS Computational Biology*, 6(4), 2010. ISSN 1553734X. doi: 10.1371/journal.pcbi.1000745.

[91] Bambang S. Adiwijaya, Tara L. Kieffer, Joshua Henshaw, Karen Eisenhauer, Holly Kimko, John J. Alam, Robert S. Kauffman, and Varun Garg. A viral dynamic model for treatment regimens with direct-acting antivirals for chronic hepatitis c infection. *PLoS Computational Biology*, 8(1), 2012. ISSN 1553734X. doi: 10.1371/journal.pcbi.1002339.

[92] Massimo Cella, Catherijne Knibbe, Saskia N. de Wildt, Joop Van Gerven, Meindert Danhof, and Oscar Della Pasqua. Scaling of pharmacokinetics across paediatric populations: The lack of interpolative power of allometric models. *British Journal of Clinical Pharmacology*, 74(3):525–535, 2012. ISSN 03065251. doi: 10.1111/j.1365-2125.2012.04206.x.

[93] Liam V. Brown, Eamonn A. Gaffney, Jonathan Wagg, and Mark C. Coles. An in silico model of cytotoxic T-lymphocyte activation in the lymph node following short peptide vaccination. *Journal of The Royal Society Interface*, 15(140), 2018. ISSN 1742-5689. doi: 10.1098/rsif.2018.0041.

[94] Amy L Olex, Elizabeth M Hiltbold, Xiaoyan Leng, and Jacquelyn S Fetrow. Dynamics of dendritic cell maturation are identified through a novel filtering strategy applied to biological time-course microarray replicates. *BMC immunology*, 11:41, 2010. ISSN 1471-2172. doi: 10.1186/1471-2172-11-41.

[95] Michele Weber, Robert Hauschild, Jan Schwarz, Christine Moussion, Ingrid de Vries, Daniel F Legler, Sanjiv a Luther, Tobias Bollenbach, and Michael Sixt. Interstitial dendritic cell guidance by haptotactic chemokine gradients. *Science (New York, N.Y.)*, 339(6117):328–32, 2013. ISSN 1095-9203. doi: 10.1126/science.1228456.

[96] Menna R Clatworthy, Caren E Petrie Aronin, Rebeccah J Mathews, Nicole Y Morgan, Kenneth G C Smith, and Ronald N Germain. Immune complexes stimulate CCR7-dependent dendritic cell migration to lymph nodes. *Nature Medicine*, 20(12):1458–1463, 2014. ISSN 1078-8956. doi: 10.1038/nm.3709.

[97] Michio Tomura, Akihiro Hata, Satoshi Matsuoka, Francis H. W. Shand, Yasutaka Nakanishi, Ryoyo Ikebuchi, Satoshi Ueha, Hidekazu Tsutsui, Kayo Inaba, Kouji Matsushima, Atsushi Miyawaki, Kenji Kabashima, Takeshi Watanabe, and Osami Kanagawa. Tracking and quantification of dendritic cell migration and antigen trafficking between the skin and lymph nodes. *Scientific Reports*, 4:6030, 2014. ISSN 2045-2322. doi: 10.1038/srep06030.

[98] Peter S. Kim, Peter P. Lee, and Doron Levy. Modeling regulation mechanisms in the immune system. *Journal of Theoretical Biology*, 246(1):33–69, 2007. ISSN 00225193. doi: 10.1016/j.jtbi.2006.12.012.

[99] Masakazu Akiyama, Kazuma Tateiri, Takamichi Sushida, Ryo Kobayashi, Yuuto Hirayoshi, and Mamoru Okamoto. GLSC3D v3.0.0. URLs `https://github.com/GLSC3DProject/GLSC3D` and `http://www-mmc.es.hokudai.ac.jp/~masakazu/`. Accessed Dec 2017. Japanese language source.

[100] N.G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North Holland, 3rd edition, 5 2007. ISBN 978-0444529657.

[101] Kamonwad Ngamchuea, Richard O. D. Clark, Stanislav V. Sokolov, Neil P. Young, Christopher Batchelor-McAuley, and Richard G. Compton. Single Oxidative Collision Events of Silver Nanoparticles: Understanding the Rate-Determining Chemistry. *Chemistry – A European Journal*, 23(63):16085–16096, Oct 2017. ISSN 1521-3765. doi: 10.1002/chem.201703591.

[102] Shaltiel Eloul, Enno Kätelhön, Christopher Batchelor-McAuley, Kristina Tschulik, and Richard G. Compton. Diffusional impacts of nanoparticles on microdisc and microwire electrodes: The limit of detection and first passage statistics. *J. Electroanal. Chem.*, 755:136–142, Oct 2015. ISSN 1572-6657. doi: 10.1016/j.jelechem.2015.07.042.

[103] Attila Szabo, Klaus Schulten, and Zan Schulten. First passage time approach to diffusion controlled reactions. *J. Chem. Phys.*, 72(8):4350–4357, Apr 1980. ISSN 0021-9606. doi: 10.1063/1.439715.

[104] Volker Teichgräber, Vaios Karanikas, Christine Trumpheller, Mailen Kootsey, Eamonn A. Gaffney, Liam V. Brown, Alex Phipps, Jean Eric Charoin, Gonzalo Acuna, Pavel Pisa, and Jonathan Wagg. Therapeutic Cancer Vaccines: Optimizing the Delivery of Short peptide Vaccines to Achieve Successful Clinical Outcome. Submitted.

[105] Eric Reits, Alexander Griekspoor, Joost Neijssen, Tom Groothuis, Kees Jalink, Peter Van Veelen, Hans Janssen, Jero Calafat, Jan Wouter Drijfhout, and Jacques Neefjes. Peptide Diffusion, Protection, and Degradation in Nuclear and Cytoplasmic Compartments before Antigen Presentation by MHC Class I. *Immunity*, 18(1):97–108, 2003. ISSN 10747613. doi: 10.1016/S1074-7613(02)00511-3.

[106] Ivo R. Ruttekolk, J. Joris Witsenburg, Heike Glauner, Petra H M Bovee-Geurts, Emer S. Ferro, Wouter P R Verdurmen, and Roland Brock. The intracellular pharmacokinetics of terminally capped peptides. *Molecular Pharmaceutics*, 9(5):1077–1086, 2012. ISSN 15438384. doi: 10.1021/mp200331g.

[107] Mario Castro. *A Physicist's Approach to Immunology*, pages 339–350. Springer New York, New York, NY, 2011. ISBN 978-1-4419-7725-0. doi: 10.1007/978-1-4419-7725-0_16.

[108] Christopher J.H. Porter and Susan A. Charman. Lymphatic transport of proteins after subcutaneous administration. *Journal of Pharmaceutical Sciences*, 89(3):297–310, 2000. ISSN 00223549. doi: 10.1002/(SICI)1520-6017(200003)89:3⟨297::AID-JPS2⟩3.0.CO;2-P. URL `http://dx.doi.org/10.1002/(SICI)1520-6017(200003)89:3{\%}3C297::AID-JPS2{\%}3E3.0.CO;2-P`.

[109] Janosch Deeg, Markus Axmann, Jovana Matic, Anastasia Liapis, David Depoil, Jehan Afrose, Silvia Curado, Michael L. Dustin, and Joachim P. Spatz. T cell activation is determined by the number of presented antigens. *Nano Letters*, 13(11):5619–5626, 2013. ISSN 15306984. doi: 10.1021/nl403266t.

[110] David R Fooksman, Santosh Vardhana, Gaia Vasiliver-Shamis, Jan Liese, David a Blair, Janelle Waite, Catarina Sacristán, Gabriel D Victora, Alexandra Zanin-Zhorov, L Michael, and Michael L Dustin. Functional anatomy of T cell activation and synapse formation. *Annual Review of Immunology*, 28:79–105, 2010. ISSN 0732-0582. doi: 10.1146/annurev-immunol-030409-101308.

[111] Michael D. Cahalan and Ian Parker. Close encounters of the first and second kind: T-DC and T-B interactions in the lymph node. *Seminars in Immunology*, 17(6): 442–451, 2005. ISSN 10445323. doi: 10.1016/j.smim.2005.09.001.

[112] Gib Bogle and P. Rod Dunbar. T cell responses in lymph nodes. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 2(1):107–116, 2010. ISSN 19395094. doi: 10.1002/wsbm.47.

[113] Hélène Beuneu, Zacarias Garcia, and Philippe Bousso. Cutting edge: cognate CD4 help promotes recruitment of antigen-specific CD8 T cells around dendritic cells. *Journal of Immunology (Baltimore, Md. : 1950)*, 177(3):1406–10, 2006. ISSN 0022-1767. doi: 10.4049/jimmunol.177.3.1406.

[114] J R Delaney, Y Sykulev, H N Eisen, and S Tonegawa. Differences in the level of expression of class I major histocompatibility complex proteins on thymic epithelial and dendritic cells influence the decision of immature thymocytes between positive and negative selection. *Proceedings of the National Academy of Sciences of the United States of America*, 95(9):5235–5240, 1998. ISSN 0027-8424. doi: 10.1073/pnas.95.9.5235.

[115] Donald F Hunt, Jeffrey Shabanowitz, Kazuyasu Sakaguchi, Hanspeter Michel, Noelle Sevilir, Andrea L Cox, Etrore Appella, Robert a Henderson, and Victor H Engelhard. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science*, 255(5049):1261–1263, 1992.

[116] S Stevanović and H Schild. Quantitative aspects of T cell activation–peptide generation and editing by MHC class I molecules. *Seminars in immunology*, 11(6):375–384, 1999. ISSN 1044-5323. doi: 10.1006/smim.1999.0195.

[117] Theodore J Tsomides, B D Walker, Herman N Eisen, and Bruce D Walkert. An optimal viral peptide recognized by CD8+ T cells binds very tightly to the restricting class I major histocompatibility complex protein on intact cells but not to the purified class I protein. *Proceedings of the National Academy of Sciences of the United States of America*, 88(December):11276–11280, 1991. ISSN 0027-8424. doi: 10.1073/pnas.88.24.11276.

[118] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. URL `http://www.scipy.org/`. [Online; accessed Dec 2017].

[119] Teerawan Rattanapak, James C. Birchall, Katherine Young, Atsuko Kubo, Sayumi Fujimori, Masaru Ishii, and Sarah Hook. Dynamic visualization of dendritic cell-antigen interactions in the skin following transcutaneous immunization. *PLoS ONE*, 9(2):1–12, 2014. ISSN 19326203. doi: 10.1371/journal.pone.0089503.

[120] Erik H J G Aarntzen, Mangala Srinivas, Fernando Bonetto, Luis J. Cruz, Pauline Verdijk, Gerty Schreibelt, Mandy Van De Rakt, W. Joost Lesterhuis, Maichel Van Riel, Cornelius J a Punt, Gosse J. Adema, Arend Heerschap, Carl G. Figdor, Wim J. Oyen, and I. Jolanda M De Vries. Targeting of 111in-labeled dendritic cell human vaccines improved by reducing number of cells. *Clinical Cancer Research*, 19(6):1525–1533, 2013. ISSN 10780432. doi: 10.1158/1078-0432.CCR-12-1879.

[121] Gerald Linette. Passenger mutations as target antigens for immuno-oncology. In *ESMO Symposium on Immuno-Oncology*. European Society for Medical Oncology, November 2016.

[122] Li Li, Shiro Mori, Maya Sakamoto, Shoki Takahashi, and Tetsuya Kodama. Mouse Model of Lymph Node Metastasis via Afferent Lymphatic Vessels for Development of Imaging Modalities. *PLoS ONE*, 8(2):e55797, 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0055797.

[123] Tetsuya Kodama, Yuriko Hatakeyama, Shigeki Kato, and Shiro Mori. Visualization of fluid drainage pathways in lymphatic vessels and lymph nodes using a mouse model to test a lymphatic drug delivery system. *Biomedical Optics Express*, 6(1):124, 2015. ISSN 2156-7085. doi: 10.1364/BOE.6.000124.

[124] Vasiliki Economopoulos, Jennifer C. Noad, Shruti Krishnamoorthy, Brian K. Rutt, and Paula J. Foster. Comparing the MRI appearance of the lymph nodes and spleen in wild-type and immuno-deficient mouse strains. *PLoS ONE*, 6(11):2–9, 2011. ISSN 19326203. doi: 10.1371/journal.pone.0027508.

[125] Marc Bajénoff, Béatrice Breart, Alex Y C Huang, Hai Qi, Julie Cazareth, Veronique M Braud, Ronald N Germain, and Nicolas Glaichenhaus. Natural killer cell behavior in lymph nodes revealed by static and real-time imaging. *The Journal of Experimental Medicine*, 203(3):619–631, 2006. ISSN 0022-1007. doi: 10.1084/jem.20051474.

[126] Cynthia L Willard-Mack. Normal structure, function, and histology of lymph nodes. *Toxicologic pathology*, 34(5):409–424, 2006. ISSN 0192-6233. doi: 10.1080/01926230600867727.

[127] A Grakoui, S K Bromley, C Sumen, M M Davis, A S Shaw, P M Allen, and M L Dustin. The immunological synapse: a molecular machine controlling T cell activation. *Science (New York, N.Y.)*, 285(5425):221–227, 1999. ISSN 00368075. doi: 10.1126/science.285.5425.221.

[128] E. Rosemary Christinck, Mark A. Luscher, Brian H. Barber, and David B. Williams. Peptide binding to class I MHC on living cells and quantitation of complexes required for CTL lysis. *Nature*, 352:67–70, 1991.

[129] Darrell J Irvine, Marco a Purbhoo, Michelle Krogsgaard, and Mark M Davis. Direct observation of ligand recognition by T cells. *Nature*, 419(6909):845–849, 2002. ISSN 0028-0836. doi: 10.1038/nature01076.

[130] Antonella Viola and Antonio Lanzavecchia. T cell activation determined by T cell receptor number and tunable thresholds. *Science (New York, N.Y.)*, 273(5271):104–106, 1996. ISSN 0036-8075. doi: 10.1126/science.273.5271.104.

[131] Zacarias Garcia, Emmanuelle Pradelli, Susanna Celli, Hélène Beuneu, Aurélie Simon, and Philippe Bousso. Competition for antigen determines the stability of T cell-dendritic cell interactions during clonal expansion. *Proceedings of the National Academy of Sciences of the United States of America*, 104(11):4553–4558, 2007. ISSN 0027-8424. doi: 10.1073/pnas.0610019104.

[132] Joseph N Blattman. Estimating the Precursor Frequency of Naive Antigen-specific CD8 T Cells. *Journal of Experimental Medicine*, 195(5):657–664, 2002. ISSN 00221007. doi: 10.1084/jem.20001021.

[133] Fabrice Lemaitre, Helen K W Law, Milena Hasan, and Matthew L Albert. Enumeration of human antigen – specific naive CD8+ T cells reveals conserved precursor frequencies. *Blood*, 115(18):3718–25, 2010. ISSN 1528-0020. doi: 10.1182/blood-2009-10-251124.The.

[134] L Breiman. Random Forests. *Machine Learning*, 45(1):5–32, 2001. ISSN 0885-6125. doi: 10.1023/A:1010933404324.

[135] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[136] David Makowski, Cédric Naud, Marie-Hélène Jeuffroy, Aude Barbottin, and Hervé Monod. Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model prediction. *Reliability Engineering and System Safety*, 91(10-11):1142–1147, 2006. ISSN 09518320. doi: 10.1016/j.ress.2005.11.015.

[137] A Saltelli, S Tarantola, and K. P.-S. Chan. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics*, 41(1):39–56, 1999. ISSN 0040-1706. doi: 10.1080/00401706.1999.10485594.

[138] A Saltelli, S Tarantola, and F Campolongo. Sensitivity Analysis as an Ingredient of Modeling. *Statistical Science*, 15(4):377–395, 2000.

[139] Elena Milner, Eilon Barnea, Ilan Beer, and Arie Admon. The turnover kinetics of major histocompatibility complex peptides of human cancer cells. *Molecular & cellular proteomics : MCP*, 5(2):357–365, 2006. ISSN 1535-9476. doi: 10.1074/mcp.M500241-MCP200.

[140] H. N. Eisen, X. H. Hou, C. Shen, K. Wang, V. K. Tanguturi, C. Smith, K. Kozyrytska, L. Nambiar, C. a. McKinley, J. Chen, and R. J. Cohen. Promiscuous binding of extracellular peptides to cell surface class I MHC protein. *Proceedings of the National Academy of Sciences*, 109(12):4580–4585, 2012. ISSN 0027-8424. doi: 10.1073/pnas.1201586109.

[141] Alexandra Mant, Fay Chinnery, Tim Elliott, and Anthony P. Williams. The pathway of cross-presentation is influenced by the particle size of phagocytosed antigen. *Immunology*, 136(2):163–175, 2012. ISSN 00192805. doi: 10.1111/j.1365-2567.2012.03558.x.

[142] V Cerundolo, T Elliott, J Elvin, J Bastin, H G Rammensee, and A Townsend. The binding affinity and dissociation rates of peptides for class I major histocompatibility complex molecules. *Eur J Immunol*, 21(9):2069–2075, 1991. doi: 10.1002/eji.1830210915.

[143] Malgorzata A. Garstka, Alexander Fish, Patrick H. N. Celie, Robbie P. Joosten, George M. C. Janssen, Ilana Berlin, Rieuwert Hoppes, Magda Stadnik, Lennert Janssen, Huib Ovaa, Peter a. van Veelen, Anastassis Perrakis, and Jacques Neefjes. The first step of peptide selection in antigen presentation by MHC class I molecules. *Proceedings of the National Academy of Sciences*, 112(5):1505–1510, 2015. ISSN 0027-8424. doi: 10.1073/pnas.1416543112.

[144] Melissa Lever, Hong-Sheng Lim, Philipp Kruger, John Nguyen, Nicola Trendel, Enas Abu-Shah, Philip K Maini, P. Anton van der Merwe, and Omer Dushek. A minimal signalling architecture explains the T cell response to a 1,000,000-fold variation in antigen affinity and dose. *Proceedings of the National Academy of Sciences*, 2016. ISSN 1091-6490. doi: 10.1101/071878.

[145] Tajie H. Harris, Edward J. Banigan, David A. Christian, Christoph Konradt, Elia D.Tait Wojno, Kazumi Norose, Emma H. Wilson, Beena John, Wolfgang Weninger, Andrew D. Luster, Andrea J. Liu, and Christopher A. Hunter. Generalized Lévy walks and the role of chemokines in migration of effector CD8+T cells. *Nature*, 486(7404):545–548, 2012. ISSN 00280836. doi: 10.1038/nature11098.

[146] Gib Bogle and P. Rod Dunbar. On-Lattice Simulation of T Cell Motility, Chemotaxis, and Trafficking in the Lymph Node Paracortex. *PLoS ONE*, 7(9), 2012. ISSN 19326203. doi: 10.1371/journal.pone.0045258.

[147] A Carlier, A Vasilevich, M Marechal, J de Boer, and L Geris. In silico clinical trials for pediatric orphan diseases. *Scientific Reports*, 8(2465), 2018. doi: 10.1038/s41598-018-20737-y.

[148] C Hermann, A van Hateren, N Trautwein, A Neerincx, PJ Duriez, S Stevanović, J Trowsdale, JE Deane, T Elliott, and LH Boyle. TAPBPR alters MHC class I peptide presentation by functioning as a peptide exchange catalyst. *Elife*, 6(4), 2015. doi: 10.7554/eLife.09617.

[149] D N Garboczi, D T Hung, and D C Wiley. HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in Escherichia coli and complexed with single antigenic peptides. *Proceedings of the National Academy of Sciences*, 89(8):3429–3433, 1992. ISSN 0027-8424. doi: 10.1073/pnas.89.8.3429.

[150] A. Steinle and D. J Schendel. HLA class I alleles of LCL 721 and 174 x CEM.T2 (T2). *Tissue Antigens*, 44(4):268–270, 1994.

[151] P. Parham and F. M Brodsky. Partial purification and some properties of BB7.2. A cytotoxic monoclonal antibody with specificity for HLA-A2 and a variant of HLA-A28. *Hum Immunol*, 3(4):277–299, 1981.

[152] K C Parker, M a Bednarek, and J E Coligan. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *Journal of immunology (Baltimore, Md. : 1950)*, 152(1):163–175, 1994.

[153] Morten Nielsen, Claus Lundegaard, Peder Worning, Sanne Lise Lauemøller, Kasper Lamberth, Søren Buus, Søren Brunak, and Ole Lund. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Science*, 12(5):1007–1017, 2003. ISSN 09618368. doi: 10.1110/ps.0239403.

[154] Gabriel Balaban, Henrik Finsberg, Simon Funke, Trine F. Håland, Einar Hopp, Joakim Sundnes, Samuel Wall, and Marie E. Rognes. In vivo estimation of elastic heterogeneity in an infarcted human heart. *Biomechanics and Modeling in Mechanobiology*, 17(5):1317–1329, Oct 2018. ISSN 1617-7940. doi: 10.1007/s10237-018-1028-5. URL `https://doi.org/10.1007/s10237-018-1028-5`.

[155] Neil Dalchau, Andrew Phillips, Leonard D. Goldstein, Mark Howarth, Luca Cardelli, Stephen Emmott, Tim Elliott, and Joern M. Werner. A peptide filtering relation quantifies MHC class I peptide optimization. *PLoS Computational Biology*, 7(10), 2011. ISSN 1553734X. doi: 10.1371/journal.pcbi.1002144.

[156] Alistair Bailey, Neil Dalchau, Rachel Carter, Stephen Emmott, Andrew Phillips, Jörn M. Werner, and Tim Elliott. Selector function of MHC I molecules is determined by protein plasticity. *Scientific Reports*, 5:1–15, 2015. ISSN 20452322. doi: 10.1038/srep14928.

[157] Yanhong Gu, Wei Zhao, Fanyu Meng, Bingqian Qu, Xu Zhu, Yang Sun, Yongqian Shu, and Qiang Xu. Sunitinib impairs the proliferation and function of human peripheral T cell and prevents T-cell-mediated immune response in mice. *Clinical Immunology*, 135(1):55–62, 2010. ISSN 15216616. doi: 10.1016/j.clim.2009.11.013.

[158] Ritika Jaini, Patricia Rayman, Peter A. Cohen, James H. Finke, and Vincent K. Tuohy. Combination of sunitinib with anti-tumor vaccination inhibits T cell priming and requires careful scheduling to achieve productive immunotherapy. *International Journal of Cancer*, 134(7):1695–1705, 2014. ISSN 00207136. doi: 10.1002/ijc.28488.

[159] Madeleine M Hipp, Norbert Hilf, Steffen Walter, Daniela Werth, Katharina M Brauer, Markus P Radsak, Toni Weinschenk, Harpreet Singh-jasuja, and Peter Brossart. Sorafenib, but not Sunitinib, affects function of dendritic cells and induction of primary immune responses. 111(12):5610–5620, 2008. doi: 10.1182/blood-2007-02-075945. clinical.

[160] L.V. Brown, E.A. Gaffney, J. Wagg, and M.C. Coles. Applications of mechanistic modelling to clinical and experimental immunology: an emerging technology to accelerate immunotherapeutic discovery and development. *Clinical & Experimental Immunology*, 193(3):284–292, 2018. doi: 10.1111/cei.13182.

[161] Kieran Alden, Mark Read, Jon Timmis, Paul S. Andrews, Henrique Veiga-Fernandes, and Mark Coles. Spartan: A comprehensive tool for understanding uncertainty in simulations of biological systems. *PLOS Computational Biology*, 9(2):1–9, 02 2013. doi: 10.1371/journal.pcbi.1002916.

[162] Jake P. Taylor-King, E. Emiel van Loon, Gabriel Rosser, and S. Jon Chapman. From birds to bacteria: Generalised velocity jump processes with resting states. *Bulletin of Mathematical Biology*, 77(7):1213–1236, Jul 2015. ISSN 1522-9602. doi: 10.1007/s11538-015-0083-7.

[163] David Pettitt, Zeeshaan Arshad, James Smith, Tijana Stanic, Georg Holländer, and David Brindley. CAR-T Cells: A Systematic Review and Mixed Methods Analysis of the Clinical Trial Landscape. *Molecular Therapy*, 26(2):342–353, 2018. doi: 10.1016/j.ymthe.2017.10.019.

[164] He Li, Yao Huang, Du-Qing Jiang, Lian-Zhen Cui, Zhou He, Chao Wang, Zhi-Wei Zhang, Hai-Li Zhu, Yong-Mei Ding, Lin-Fang Li, Qiang Li, Hua-Jun Jin, and Qi-Jun Qian. Antitumor activity of EGFR-specific CAR T cells against non-small-cell lung cancer cells in vitro and in mice. *Cell Death & Disease*, 9(2):177, 2018. ISSN 2041-4889. doi: 10.1038/s41419-017-0238-6.

[165] Nobuhiro Nishio, Iulia Diaconu, Hao Liu, Vincenzo Cerullo, Ignazio Caruana, Valentina Hoyos, Lisa Bouchier-Hayes, Barbara Savoldo, and Gianpietro Dotti. Armed oncolytic virus enhances immune functions of chimeric antigen receptor-modified T cells in solid tumors. *Cancer Research*, 74(18):5195–5205, 2014. ISSN 15387445. doi: 10.1158/0008-5472.CAN-14-0697.

[166] Mauro P. Avanzi, Oladapo Yeku, Xinghuo Li, Dinali P. Wijewarnasuriya, Dayenne G. van Leeuwen, Kenneth Cheung, Hyebin Park, Terence J. Purdon, Anthony F.

Daniyan, Matthew H. Spitzer, and Renier J. Brentjens. Engineered Tumor-Targeted T Cells Mediate Enhanced Anti-Tumor Efficacy Both Directly and through Activation of the Endogenous Immune System. *Cell Reports*, 23(7):2130–2141, 2018. ISSN 22111247. doi: 10.1016/j.celrep.2018.04.051.

[167] M E Smith and W L Ford. The recirculating lymphocyte pool of the rat: a systematic description of the migratory behaviour of recirculating lymphocytes. *Immunology*, 49 (1):83–94, 1983. ISSN 0019-2805.

[168] Michael H. Kershaw, Jennifer A. Westwood, Linda L. Parker, Gang Wang, Zelig Eshhar, Sharon A. Mavroukakis, Donald E. White, John R. Wunderlich, Silvana Canevari, Linda Rogers-Freezer, Clara C. Chen, James C. Yang, Steven A. Rosenberg, and Patrick Hwu. A Phase I Study on Adoptive Immunotherapy Using Gene-Modified T Cells for Ovarian Cancer. *Clinical Cancer Research*, 12(20):6106–6115, 2006. ISSN 1078-0432. doi: 10.1158/1078-0432.CCR-06-1183.

[169] Ana Caterina Parente-Pereira, Jerome Burnet, David Ellison, Julie Foster, David Marc Davies, Sjoukje Van der Stegen, Sophie Burbridge, Laura Chiapero-Stanke, Scott Wilkie, Stephen Mather, and John Maher. Trafficking of CAR-Engineered Human T Cells Following Regional or Systemic Adoptive Transfer in SCID Beige Mice. *Journal of Clinical Immunology*, 31(4):710–718, 2011. doi: 10.1007/s10875-011-9532-8.

[170] Luca Gattinoni, Christopher A. Klebanoff, Douglas C. Palmer, Claudia Wrzesinski, Keith Kerstann, Zhiya Yu, Steven E. Finkelstein, Marc R. Theoret, Steven A. Rosenberg, and Nicholas P. Restifo. Acquisition of full effector function in vitro paradoxically impairs the in vivo antitumor efficacy of adoptively transferred CD8+ T cells. *Journal of Clinical Investigation*, 115(6):1616–1626, 2005. ISSN 00219738. doi: 10.1172/JCI24480.

[171] Chengcheng Zhang, Zhe Wang, Zhi Yang, Meiling Wang, Shiqi Li, Yunyan Li, Rui Zhang, Zhouxing Xiong, Zhihao Wei, Junjie Shen, Yongli Luo, Qianzhen Zhang, Limei Liu, Hong Qin, Wei Liu, Feng Wu, Wei Chen, Feng Pan, Xianquan Zhang, Ping Bie, Houjie Liang, Gabriele Pecher, and Cheng Qian. Phase I Escalating-Dose Trial of CAR-T Therapy Targeting CEA+Metastatic Colorectal Cancers. *Molecular Therapy*, 25(5):1248–1258, 2017. ISSN 15250024. doi: 10.1016/j.ymthe.2017.03.010.

[172] Dhaval K. Shah and Alison M. Betts. Towards a platform PBPK model to characterize the plasma and tissue disposition of monoclonal antibodies in preclinical species

and human. *Journal of Pharmacokinetics and Pharmacodynamics*, 39(1):67–86, 2012. ISSN 1567567X. doi: 10.1007/s10928-011-9232-2.

[173] ICRP. Basic Anatomical and Physiological Data for Use in Radiological Protection Reference Values. *ICRP Publication 89, Ann. ICRP*, 32(3-4), 2002.

[174] R P Brown, M D Delp, S L Lindstedt, L R Rhomberg, and R P Beliles. Physiological parameter values for physiologically based pharmacokinetic models. *Toxicology and Industrial Health*, 13(4):407–84, 1997. ISSN 0748-2337. doi: 10.1177/074823379701300401.

[175] John F. Graf, Bernhard J. Scholz, and Maria I. Zavodszky. BioDMET: A physiologically based pharmacokinetic simulation tool for assessing proposed solutions to complex biological problems. *Journal of Pharmacokinetics and Pharmacodynamics*, 39(1):37–54, 2012. ISSN 1567567X. doi: 10.1007/s10928-011-9229-x.

[176] L T Baxter, H Zhu, D G Mackensen, and R K Jain. Physiologically based pharmacokinetic model for specific and nonspecific monoclonal antibodies and fragments in normal tissues and human tumor xenografts in nude mice. *Cancer Res*, 54(6): 1517–1528, 1994. ISSN 0008-5472.

[177] Sheila Annie Peters. *Physiologically-Based Pharmacokinetic (PBPK) Modeling and Simulations: Principles, Methods, and Applications in the Pharmaceutical Industry.* Wiley, 2012. ISBN 978-0-470-48406-7.

[178] Matteo Bellone and Arianna Calcinotto. Ways to enhance lymphocyte trafficking into tumors and fitness of tumor infiltrating lymphocytes. *Frontiers in oncology*, 3 (September):231, 2013. ISSN 2234-943X. doi: 10.3389/fonc.2013.00231.

[179] F. Kallinowski, K. H. Schlenger, S. Runkel, M. Kloes, M. Stohrer, P. Okunieff, and P. Vaupel. Blood Flow, Metabolism, Cellular Microenvironment, and Growth Rate of Human Tumor Xenografts. *Cancer Res.*, 49(14):3759–3764, 1989. ISSN 0008-5472.

[180] H Lyng, A Skretting, and E K Rofstad. Blood flow in six human melanoma xenograft lines with different growth characteristics. *Cancer research*, 52(3):584–592, 1992. ISSN 00085472.

[181] Michail A. Allias, K. Wårdell, M. Stücker, C. Anderson, and E. G. Salerud. Assessment of pigmented skin lesions in terms of blood perfusion estimates. *Skin Research and Technology*, 10(1):43–49, 2004. ISSN 0909752X. doi: 10.1111/j.1600-0846.2004.00052.x.

[182] Gemma Lancaster, Aneta Stefanovska, Margherita Pesce, Gian Marco Vezzoni, Barbara Loggini, Raffaele Pingitore, Fabrizio Ghiara, Paolo Barachini, Gregorio Cervadoro, Marco Romanelli, and Marco Rossi. Dynamic markers based on blood perfusion fluctuations for selecting skin melanocytic lesions for biopsy. *Scientific Reports*, 5:1–13, 2015. ISSN 20452322. doi: 10.1038/srep12825.

[183] Quan-Sing Ng, Vicky Goh, Jessica Milner, Anwar R. Padhani, Michele I. Saunders, and Peter J. Hoskin. Acute tumor vascular effects following fractionated radiotherapy in human lung cancer: In vivo whole tumor assessment using volumetric perfusion computed tomography. *International Journal of Radiation Oncology*Biology*Physics*, 67(2):417–424, 2007. ISSN 03603016. doi: 10.1016/j.ijrobp.2006.10.005.

[184] Quan Sing Ng, Vicky Goh, Ernst Klotz, Heinz Fichte, Michele I. Saunders, Peter J. Hoskin, and Anwar R. Padhani. Quantitative assessment of lung cancer perfusion using MDCT: Does measurement reproducibility improve with greater tumor volume coverage? *American Journal of Roentgenology*, 187(4):1079–1084, 2006. ISSN 0361803X. doi: 10.2214/AJR.05.0889.

[185] F Kiessling, J Boese, C Corvinus, J R Ederle, I Zuna, S O Schoenberg, G Brix, a Schmähl, S Tuengerthal, F Herth, H U Kauczor, and M Essig. Perfusion CT in patients with advanced bronchial carcinomas: a novel chance for characterization and treatment monitoring? *European radiology*, 14(7):1226–33, 2004. ISSN 0938-7994. doi: 10.1007/s00330-004-2288-2.

[186] Jianwei Wang, Ning Wu, Matthew D. Cham, and Ying Song. Tumor response in patients with advanced non-small cell lung cancer: Perfusion CT evaluation of chemotherapy and radiation therapy. *American Journal of Roentgenology*, 193(4): 1090–1096, 2009. ISSN 0361803X. doi: 10.2214/AJR.08.1367.

[187] Joline S. W. Lind, Martijn R. Meijerink, Anne-Marie C. Dingemans, Cornelis van Kuijk, Michel C. Öllers, Dirk de Ruysscher, Pieter E. Postmus, and Egbert F. Smit. Dynamic contrast-enhanced CT in patients treated with sorafenib and erlotinib for non-small cell lung cancer: a new method of monitoring treatment? *European Radiology*, 20(12):2890–2898, 2010. ISSN 0938-7994. doi: 10.1007/s00330-010-1869-5.

[188] Minming Zhang and Michio Kono. Solitary Pulmonary Nodules: Evaluation of Blood Flow Patterns with Dynamic CT. *Radiology*, 205:471–478, 1997. ISSN 00338419.

[189] Robert Hermans, Philippe Lambin, Ann Van Der Goten, Walter Van Den Bogaert, Berit Verbist, Caroline Weltens, and Pierre R. Delaere. Tumoural perfusion as measured by dynamic computed tomography in head and neck carcinoma. *Radiotherapy and Oncology*, 53(2):105–111, 1999. ISSN 01678140. doi: 10.1016/S0167-8140(99)00132-2.

[190] Robert Hermans, Martijn Meijerink, Walter Van Den Bogaert, Alexis Rijnders, Caroline Weltens, and Philippe Lambin. Tumor perfusion rate determined noninvasively by dynamic computed tomography predicts outcome in head-and-neck cancer after radiotherapy. *International Journal of Radiation Oncology Biology Physics*, 57(5): 1351–1356, 2003. ISSN 03603016. doi: 10.1016/S0360-3016(03)00764-8.

[191] Dheeraj Gandhi, D. B. Chepeha, T. Miller, R. C. Carlos, C. R. Bradford, R. Karamchandani, F. Worden, A. Eisbruch, T. N. Teknos, G. T. Wolf, and S. K. Mukherji. Correlation between initial and early follow-up CT perfusion parameters with endoscopic tumor response in patients with advanced squamous cell carcinomas of the oropharynx treated with organ-preservation therapy. *American Journal of Neuroradiology*, 27(1):101–106, 2006. ISSN 01956108. doi: ⟨p⟩⟨/p⟩.

[192] H I Bicher. The physiological effects of hyperthermia. *Radiology*, 137(2):511–513, 1980. ISSN 0033-8419 (Print). doi: 10.1148/radiology.137.2.7433684.

[193] Ali Syed Arbab, Shigeki Aoki, Keiji Toyama, Nobuhiko Miyazawa, Hiroshi Kumagai, Takako Umeda, Takao Arai, Tsutomu Araki, Hiroyuki Kabasawa, and Yoshiyuki Takahashi. Quantitative measurement of regional cerebral blood flow with flow-sensitive alternating inversion recovery imaging: Comparison with [iodine 123]-iodoamphetamin single photon emission CT. *American Journal of Neuroradiology*, 23(3):381–388, 2002. ISSN 01956108.

[194] Vicky Goh, Steve Halligan, and Clive I. Bartram. Quantitative Tumor Perfusion Assessment with Multidetector CT: Are Measurements from Two Commercial Software Packages Interchangeable? *Radiology*, 242(3):777–782, 2007. ISSN 0033-8419. doi: 10.1148/radiol.2423060279.

[195] Vicky Goh, Steve Halligan, Jo Ann Hugill, and Clive I. Bartram. Quantitative assessment of tissue perfusion using MDCT: Comparison of colorectal cancer and skeletal muscle measurement reproducibility. *American Journal of Roentgenology*, 187(1): 164–169, 2006. ISSN 0361803X. doi: 10.2214/AJR.05.0050.

[196] Vicky Goh, Steve Halligan, Jo-ann Hugill, Louise Gartner, and Clive I Bartram. Quantitative Colorectal Cancer Perfusion Measurement Using Dynamic Contrast-Enhanced Multidetector-Row Computed Tomography: Effect of Acquisition Time and Implications for Protocols. *Journal of Computer Assisted Tomography*, 29(1):59–63, 2005. ISSN 0363-8715. doi: 10.1097/01.rct.0000152847.00257.d7.

[197] Vicky Goh, Steve Halligan, Jo Ann Hugill, Paul Bassett, and Clive I. Bartram. Quantitative assessment of colorectal cancer perfusion using MDCT: Inter- and intraobserver agreement. *American Journal of Roentgenology*, 185(1):225–231, 2005. ISSN 0361803X. doi: 10.2214/ajr.185.1.01850225.

[198] Dushyant V. Sahani, Sanjeeva P. Kalva, Leena M. Hamberg, Peter F. Hahn, Christopher G. Willett, Sanjay Saini, Peter R. Mueller, and Ting-Yim Lee. Assessing Tumor Perfusion and Treatment Response in Rectal Cancer with Multisection CT: Initial Observations. *Radiology*, 234(3):785–792, 2005. ISSN 0033-8419. doi: 10.1148/radiol.2343040286.

[199] R J Touloukian, R R Rickert, R C Lange, and R P Spencer. The micro-vascular circulation of lymphangiomas: a study of Xe 133 clearance and pathology. *Pediatrics*, 48(1):36–40, 1971. ISSN 00314005.

[200] U Plengvanit, R Suwanik, O Chearanai, S Intrasupt, S Sutayavanich, C Kalayasiri, and V Viranuvatti. Regional hepatic blood flow studied by intrahepatic injection of 133 xenon in normals and in patients with primary carcinoma of the liver, with particular reference to the effect of hepatic artery ligation. *Australian and New Zealand journal of medicine*, 2(1):44–48, 1972. ISSN 0004-8291. doi: 10.1111/j.1445-5994.1972.tb03907.x.

[201] R H Wheeler, H A Ziessman, and B R Medvec. Tumor blood flow and systemic shunting in patients receiving intraarterial chemotherapy for head and neck cancer. *Cancer Research*, 46(8):4200–4204, 1986. ISSN 0008-5472.

[202] F M Waterman, R E Nerlinger, D J 3Rd Moylan, and Dennis B Leeper. Response of human tumor blood flow to local hyperthermia. *Int J Radiat Oncol Biol Phys*, 13(1): 75–82, 1987. ISSN 0360-3016.

[203] Clinton Nyström, Lars Forssman, and Bengt Roos. Myometrial Blood Flow Studies in Carcinoma of the Corpus Uteri: A preliminary report on the clearance method using xenon 133. *Acta Oncologica*, 8(3):193–198, 1969. ISSN 0284-186X. doi: 10.3109/02841866909134450.

[204] Matti J. Mäntylä, Jarmo T. Toivanen, Maunu A. Pitkänen, and Ahti H. Rekonen. Radiation-induced changes in regional blood flow in human tumors. *International Journal of Radiation Oncology*Biology*Physics*, 8(10):1711–1717, 1982. ISSN 03603016. doi: 10.1016/0360-3016(82)90291-7.

[205] M J Mantyla. Regional blood flow in human tumors. *Cancer research*, 39(6 Pt 2): 2304–2306, 1979.

[206] I. Taylor, R. Bennett, and S. Sherriff. The blood supply of colorectal liver metastases. *British Journal of Cancer*, 38(6):749–756, 1978. ISSN 15321827. doi: 10.1038/bjc.1978.283.

[207] Thaddeus V. Samulski, Peter Fessenden, Riccardo Valdagni, and Daniel S. Kapp. Correlations of thermal washout rate, steady state temperatures, and tissue type in deep seated recurrent or metastatic tumors. *International Journal of Radiation Oncology, Biology, Physics*, 13(6):907–916, 1987. ISSN 03603016. doi: 10.1016/0360-3016(87)90106-4.

[208] Tapani Lahtinen, Paavo Karjalainen, and Esko M. Alhava. Measurement of bone blood flow with a 133Xe washout method. *European Journal of Nuclear Medicine*, 4: 437, 1979.

[209] Richard Johnson. A thermodynamic method for investigation of radiation induced changes in the microcirculation of human tumors. *International Journal of Radiation Oncology, Biology, Physics*, 1(7-8):659–670, 1976. ISSN 03603016. doi: 10.1016/0360-3016(76)90148-6.

[210] G. Ribeiro, D. A. Jones, and M. Jones. Carcinoma of the breast associated with pregnancy. *British Journal of Surgery*, 73(8):607–609, 1986. ISSN 13652168. doi: 10.1002/bjs.1800730805.

[211] C G Rhodes, R J S Wise, J M Gibbs, R S J Frackowiak, J Hatazawa, A J Palmer, D G Thomas, and T Jones. In vivo disturbance of the oxidative metabolism of glucose in human cerebral gliomas. *Annals of neurology*, 14:614–626, 1983. ISSN 0364-5134. doi: 10.1002/ana.410140604.

[212] J Olesen and O B Paulson. The effect of intra-arterial papaverine on the regional cerebral blood flow in patients with stroke or intracranial tumor. *Stroke; a journal of cerebral circulation*, 2(2):148–159, 1971. ISSN 0039-2499. doi: 10.1161/01.STR.2.2.148.

[213] A A Lammertsma. Positron emission tomography and in vivo measurements of tumor perfusion and oxygen utilization. *Canc Meta Rev*, 6:521–539, 1987.

[214] M Ito, A A Lammertsma, R J Wise, Silvia Bernardi, R S Frackowiak, J D Heather, C G McKenzie, D G Thomas, and T Jones. Measurement of regional cerebral blood flow and oxygen utilisation in patients with cerebral tumours using 15O and positron emission tomography: analytical techniques and preliminary results. *Neuroradiology*, 23:63–74, 1982. ISSN 00283940. doi: 10.1007/BF00367239.

[215] D J Brooks, R P Beaney, A A Lammertsma, D R Turton, J Marshall, D G Thomas, and T Jones. Studies on regional cerebral haematocrit and blood flow in patients with cerebral tumours using positron emission tomography. *Microvasc Res*, 31(3):267–276, 1986.

[216] R P Beaney, D K Brooks, and K L Leenders et al. Blood flow and oxygen utilization in the contralateral cerebral cortex of patients with untreated intracranial tumors as studied by positron emission tomography with observations on the effect of decompressive surgery. *J Neurol, Neurosurg, Psychiat*, 48:310–319, 1985.

[217] R P Beaney. Positron emission tomography in the study of human tumors. *Seminars in nuclear medicine*, 14(4):324–341, 1984. ISSN 0001-2998 (Print).

[218] Massimo Bellomi, Giuseppe Petralia, Angelica Sonzogni, Maria Giulia Zampino, and Andrea Rocca. CT perfusion for the monitoring of neoadjuvant chemotherapy and radiation therapy in rectal carcinoma: initial experience. *Radiology*, 244(2):486–493, 2007. ISSN 0033-8419. doi: 10.1148/radiol.2442061189.

[219] Mike Notohamiprodjo, Steven Sourbron, Michael Staehler, Henrik J. Michaely, Ulrike I. Attenberger, Gerwin P. Schmidt, Holger Boehm, Annie Horng, Christian Glaser, Christian Stief, Maximilian F. Reiser, and Karin A. Herrmann. Measuring perfusion and permeability in renal cell carcinoma with dynamic contrast-enhanced MRI: A pilot study. *Journal of Magnetic Resonance Imaging*, 31(2):490–501, 2010. ISSN 10531807. doi: 10.1002/jmri.22028.

[220] Chaan S. Ng, Xiaohong Wang, Silvana C. Faria, E. Lin, Chusilp Charnsangavej, and Nizar M. Tannir. Perfusion CT in patients with metastatic renal cell carcinoma treated with interferon. *American Journal of Roentgenology*, 194(1):166–171, 2010. ISSN 0361803X. doi: 10.2214/AJR.09.3105.

[221] Yan Chen, Jin Zhang, Jingrui Dai, Xiaoli Feng, Haizhen Lu, and Chunwu Zhou. Angiogenesis of renal cell carcinoma: Perfusion CT findings. *Abdominal Imaging*, 35 (5):622–628, 2010. ISSN 09428925. doi: 10.1007/s00261-009-9565-0.

[222] Cedric De Bazelaire, Neil M. Rofsky, Guillaume Duhamel, M. Dror Michaelson, Daniel George, and David C. Alsop. Arterial spin labeling blood flow magnetic resonance imaging for the characterization of metastatic renal cell carcinoma. *Academic Radiology*, 12(3):347–357, 2005. ISSN 10766332. doi: 10.1016/j.acra.2004.12.012.

[223] I Wang, S Andersson-Engels, G E Nilsson, K Wårdell, and K Svanberg. Superficial blood flow following photodynamic therapy of malignant non-melanoma skin tumours measured by laser Doppler perfusion imaging. *The British Journal of Dermatology*, 136(2):184–9, 1997. ISSN 0007-0963. doi: 10.1046/j.1365-2133.1997.d01-1166.x.

[224] H.S. Reinhold. Physiological effects of hyperthermia. Recent Results. *Cancer Research*, 107:32–43, 1988.

[225] Petter Brodin, Vladimir Jojic, Tianxiang Gao, Sanchita Bhattacharya, Cesar J Lopez Angel, David Furman, Shai Shen-Orr, Cornelia L. Dekker, Gary E. Swan, Atul J. Butte, Holden T. Maecker, and Mark M. Davis. Variation in the human immune system is largely driven by non-heritable influences. *Cell*, 160(1-2):37–47, 2015. ISSN 10974172. doi: 10.1016/j.cell.2014.12.020.

[226] Petter Brodin and Mark M Davis. Human immune system variation. *Nature Reviews Immunology*, 17:21–29, 2016. ISSN 1474-1733. doi: 10.1038/nri.2016.125.

[227] Julien Vibert and Véronique Thomas-Vaslin. Modelling T cell proliferation: Dynamics heterogeneity depending on cell differentiation, age, and genetic background. *PLoS Computational Biology*, 13(3):1–25, 2017. ISSN 15537358. doi: 10.1371/journal.pcbi.1005417.

[228] Yanping Xiao and Gordon J. Freeman. The microsatellite instable subset of colorectal cancer is a particularly good candidate for checkpoint blockade immunotherapy. *Cancer Discovery*, 5(1):16–18, 2015. ISSN 21598290. doi: 10.1158/2159-8290.CD-14-1397.

[229] Expert Discusses Microsatellite Instability, Immunotherapy in CRC. `http://www.onclive.com/web-exclusives/expert-discusses-microsatellite-instability-immunotherapy-in-crc`. [Online; accessed Dec 2017].

[230] Steven A Rosenberg. Interleukin 2 for patients with renal cancer. *Nature clinical practice. Oncology*, 4(9):497, 2007. ISSN 1743-4262. doi: 10.1038/ncponc0926.

[231] Martin Reck, Delvys Rodríguez-Abreu, Andrew G. Robinson, Rina Hui, Tibor Csoszi, Andrea Fülöp, Maya Gottfried, Nir Peled, Ali Tafreshi, Sinead Cuffe, Mary O'Brien, Suman Rao, Katsuyuki Hotta, Melanie A. Leiby, Gregory M. Lubiniecki, Yue Shentu, Reshma Rangwala, and Julie R. Brahmer. Pembrolizumab versus Chemotherapy for PD-L1–Positive Non–Small-Cell Lung Cancer. *New England Journal of Medicine*, 375 (19):1823–1833, 2016. ISSN 0028-4793. doi: 10.1056/NEJMoa1606774.

[232] Chris Fellner. Ipilimumab (Yervoy) Prolongs Survival In Advanced Melanoma. *P&T*, 37(9):503–512, 2012. ISSN 1052-1372.

[233] A Amin and RL White. Interleukin-2 in Renal Cell Carcinoma: A Has-Been or a Still-Viable Option? *Journal of Kidney Cancer and VHL*, 1(7):74–83, 2014. doi: 10.15586/jkcvhl.2014.18.

[234] Christoph Scheiermann, Julie Gibbs, Louise Ince, and Andrew Loudon. Clocking in to immunity. *Nature Reviews Immunology*, 18:423–437, 2018. doi: 10.1038/s41577-018-0008-4.

[235] Markus Chmielewski, Olga Hahn, Gunter Rappl, Michael Nowak, Ingo H. Schmidt-Wolf, Andreas A. Hombach, and Hinrich Abken. T cells that target carcinoembryonic antigen eradicate orthotopic pancreatic carcinomas without inducing autoimmune colitis in mice. *Gastroenterology*, 143(4):1095–1107.e2, 2012. ISSN 00165085. doi: 10.1053/j.gastro.2012.06.037.

[236] Shan-Shan Jiang, Yan Tang, Yao-Jun Zhang, D-Sheng Weng, Zhong-Guo Zhou, Ke Pan, Qiu-Zhong Pan, Qi-Jing Wang, Qing Liu, Jia He, Jing-Jing Zhao, Jiang Li, Min-Shan Chen, Alfred E Chang, Qiao Li, and Jian-Chuan Xia. A phase I clinical trial utilizing autologous tumor-infiltrating lymphocytes in patients with primary hepatocellular carcinoma. *Oncotarget*, 6(38):41339–49, 2015. ISSN 1949-2553. doi: 10.18632/oncotarget.5463.

[237] M. H. Geukes Foppen, M. Donia, I. M. Svane, and J. B.A.G. Haanen. Tumor-infiltrating lymphocytes for the treatment of metastatic cancer. *Molecular Oncology*, 9(10):1918–1935, 2015. ISSN 18780261. doi: 10.1016/j.molonc.2015.10.018.

[238] Richard A Morgan, Mark E Dudley, John R Wunderlich, Marybeth S Hughes, James C Yang, Richard M Sherry, Richard E Royal, Suzanne L Topalian, Udai S Kammula,

Nicholas P Restifo, Zhili Zheng, Azam Nahvi, Christiaan R De Vries, Linda J Rogers-freezer, Sharon A Mavroukakis, and Steven A Rosenberg. Cancer Regression in Patients. *In Vitro*, 126(2006):126–130, 2007. doi: 10.1126/science.1129003.

[239] Eric Tran, Paul F. Robbins, Yong-Chen Lu, Todd D. Prickett, Jared J. Gartner, Li Jia, Anna Pasetto, Zhili Zheng, Satyajit Ray, Eric M. Groh, Isaac R. Kriley, and Steven A. Rosenberg. T-Cell Transfer Therapy Targeting Mutant KRAS in Cancer. *New England Journal of Medicine*, 375(23):2255–2262, 2016. ISSN 0028-4793. doi: 10.1056/NEJMoa1609279.

[240] Cassandra Willyard. The mice with human tumours: Growing pains for a popular cancer model. *Nature*, 560:156–157, 2018. doi: 10.1038/d41586-018-05890-8.

[241] Sten Friberg and Stefan Mattson. On the growth rates of human malignant tumors: Implications for medical decision making. *Journal of Surgical Oncology*, 65 (4):284–297, 1997. ISSN 00224790. doi: 10.1002/(SICI)1096-9098(199708)65:4⟨284::AID-JSO11⟩3.0.CO;2-2.

[242] Rebar N. Mohammed, H. Angharad Watson, Miriam Vigar, Julia Ohme, Amanda Thomson, Ian R. Humphreys, and Ann Ager. L-selectin Is Essential for Delivery of Activated CD8+ T Cells to Virus-Infected Organs for Protective Immunity. *Cell Reports*, 14(4):760–771, 2016. ISSN 22111247. doi: 10.1016/j.celrep.2015.12.090.

[243] Cor H.J. Lamers, Yarne Klaver, Jan W. Gratama, Stefan Sleijfer, and Reno Debets. Treatment of metastatic renal cell carcinoma (mRCC) with CAIX CAR-engineered T-cells–a completed study overview. *Biochemical Society Transactions*, 44(3):951–959, 2016. ISSN 0300-5127. doi: 10.1042/BST20160037.

[244] Priti S. Hegde, Vaios Karanikas, and Stefan Evers. The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition. *Clinical Cancer Research*, 22(8):1865–1874, 2016. ISSN 1078-0432. doi: 10.1158/1078-0432.CCR-15-1507.

[245] Vitaly V. Ganusov and Michio Tomura. Experimental and mathematical approaches to quantify recirculation kinetics of lymphocytes. *bioRxiv*, page 268326, Aug 2018. doi: 10.1101/268326.

[246] Randall J. LeVeque. Advection Equations and Hyperbolic Systems. In *Finite Difference Methods for Ordinary and Partial Differential Equations*, chapter 10. Society for Industrial and Applied Mathematics, 2007. ISBN 978-0-898716-29-0.

[247] J.R. Dormand and P.J. Prince. A family of embedded runge-kutta formulae. *Journal of Computational and Applied Mathematics*, 6(1):19 – 26, 1980. ISSN 0377-0427. doi: https://doi.org/10.1016/0771-050X(80)90013-3.

[248] E. Hairer, S.P. Norsett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer-Verlag, 2 edition, 1993.

[249] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, volume 1. Cambridge University Press, 2 edition, 1992.

[250] B.A. Waaler. Physiology of the Pulmonary Circulation. *Angiologica*, 8(3–5):266–284, 1971.