Maynooth University
National University of Ireland Maynooth

Mathematical Modelling and Statistical Inference from Immune Response Data

by

Alexander S. Miles

Research Masters Thesis

Hamilton Institute

National University of Ireland Maynooth

Maynooth

Co. Kildare

August 2017

Research Supervisor: Prof. Ken Duffy

Contents

Abstract:

A hallmark of the adaptive immune response is the proliferation of pathogen-specific lymphocytes that leave in their wake a long lived population of cells that provide lasting immunity. A subject of debate is at which time point post infection those memory cells are produced during an adaptive immune response. In two ground-breaking studies, [Buchholz et al., 2013] and [Gerlach et al., 2013] introduced a new experimental method that allowed them to determine the number offspring from individual lymphocytes *in vivo* at a single harvesting time point. Through the development, application and fitting of a mathematical model, the authors of [Buchholz et al., 2013] concluded that memory cell precursors are produced before the effector cells that clear the original pathogen, contrary to prior understanding. Cohort level cell data in the paper [Kinjyo et al., 2015], however, challenges that deduction. In this thesis we sought to quantitatively reconcile these two reports by adopting the mathematical methodology of [Buchholz et al., 2013] to make it suitable for drawing inferences from the data in [Badovinac et al., 2007], [Schlub et al., 2010] and [Kinjyo et al., 2015]. When fitting to spleen and blood data reported in these papers, under the assumptions of the model, our conclusion is consistent with [Buchholz et al., 2013]: memory precursor cells appear before effector cells. However, an alternative possibility supported by the data in [Kinjyo et al., 2015] is that memory is created after the expansion phase, a deduction not possible from the data or mathematical methods in [Buchholz et al., 2013].

Acknowledgements:

I would like to thank my supervisor, professor Ken Duffy, for the great deal of time, effort and expertise he invested in me and this work. His commitment to rigour, enthusiasm for the subject and contribution of ideas made this thesis what it is. I would like to thank my past supervisor Neil O'Connell for recommending me for this position, without whom, none of this would have been possible. I would also like to thank Muriel Médard and the Network Coding and Reliable Communications Group for hosting me at MIT in the USA for a year. I would like to thank my parents, Richard and Elizabeth Miles and my girlfriend Fátima Lopez for all their encouragement throughout. Thank you to my colleges and office mates Giulio Prevedello, Gianfelice Meli and Harry Tideswell for all the feedback and for all the emotional support throughout this project. Thank you to Rosemary Hunt, Kate Moriarty and Michael Lewy for all the essential behind the scenes administration work that made everything run smoothly. I am grateful for the stipend that funded this work, which was provided by the Science Foundation Ireland Grant 12IP1263.

1 Summary

This thesis focuses on a cell known as the T lymphocyte, which is part of the immune system. We seek to provide insight into an important question of immunology: when these T cells rapidly multiply upon infection, what order do they differentiate into their cell subtypes, known as effector T cells and memory T cells? We do this by fitting mathematical models of possible T cell behaviour to experimental data reported in published papers. Our scope is to look at a subgroup of T cells known as $CD8^+$ T cells, the definition of which is covered in the next chapter. While there are many similarities between mouse and human immune systems, there are also many differences ([Mestas and Hughes, 2004]) and unless otherwise stated, all information in this thesis describes murine results.

In Chapter 2 we review some of the basics of the immune system and where $CD8^+$ T cells fit in to it. We describe the phase rapid $CD8^+$ T cell proliferation, how it is triggered and how it ends. We also describe the experimental techniques used to acquire the data that we eventually use to fit the mathematical model to.

In Chapter 3 we review some of the controversies ([Ahmed et al., 2009]) around $CD8^+$ T cell differentiation and discuss at two recent experiments that come to seemingly different conclusions on the question: [Buchholz et al., 2013] and [Kinjyo et al., 2015]. We will later fit mathematical models to the data from these papers. The authors of [Buchholz et al., 2013] used an innovative approach to the problem, observing the $CD8^+$ T cell expansion phase while tracking cells at a clonal level (looking at individual families descended from one progenitor cell) to obtain clonal statistics on the observed cell subtypes. The authors of [Buchholz et al., 2013] fit mathematical models of the expansion phase to these data and from this they determined the precursors of memory $CD8^+$ T cells appear before effector $CD8^+$ T cells. We compare this paper to another, [Kinjyo et al., 2015], because their experimental setups are similar. The [Kinjyo et al., 2015] paper does not track statistics of $CD8^+$ T cells at a clonal level in this experiment, only at a cohort level (across the whole population). However, [Kinjyo et al., 2015] provides some additional data that challenges the assumptions of [Buchholz et al., 2013], such as distribution of cell lifetimes. From this cohort data [Kinjyo et al., 2015] suggests a qualitative effector first model. We use data from two further experimental studies [Schlub et al., 2010] and [Badovinac et al., 2007], with similar setups. Evidence from these papers suggest one major cause of the differences between the [Buchholz et al., 2013] and [Kinjyo et al., 2015] results is the number of cells adoptively transferred in their experiments, which have large effect on the properties of the expansion phase. This in combination with [Kinjyo et al., 2015] data also sheds light on some further assumptions of the model in [Buchholz et al., 2013], particularly the assumption that memory precursor cells first appear in the expansion phase and not later.

In Chapter 4 we review the mathematics behind the methods that the authors of [Buchholz et al., 2013] use to fit statistical models to the observed data. These models are known as Bellman-Harris Processes ([Harris, 1964], Ch. 6) and we review some of their properties. We do this to adapt the

fitting process so as to be suitable for data in other papers.

Chapter 5 contains the novel elements of the thesis's contribution. We apply the model fitting method described in [Buchholz et al., 2013] to its reported cohort data rather than clonal data, and obtain the same deduction of memory precursors appearing before effectors in the expansion phase. This gives us confidence to apply the method to data in other papers where cohort data is more readily available. We make further changes and simplifications to the fitting method to make it appropriate for data reported in [Badovinac et al., 2007], [Schlub et al., 2010], and [Kinjyo et al., 2015]. We also use the data in [Schlub et al., 2010] and [Badovinac et al., 2007] to adjust for the difference between the papers in the number of cells adoptively transferred. After these adjustments, the memory first model gives the best fit to all data sets from the spleen and blood taken from all four papers, albeit with questionable parametrisations in some cases, and only the memory first models give a biologically plausible fit.

In Chapter 6 we consider some of the core assumptions of the model, for instance the assumption that memory appears in the expansion phase and not after. We also note the possibility that some of the observed cells characterized as early memory could instead be in the process of losing their naive characteristics. We discuss other limitations of the thesis, such as having few observed data points for models to fit to in some instances. We suggest further work, such as alternative models that could evaluate the data outside the assumptions taken. Finally we review the conclusions, that under the assumptions of the model, memory appears first, but with the caveat of heavy assumptions and fitting to limited data.

2 Review of the CD8$^+$ T cell's background

The immune system protects the body from diseases. The immune system must neutralize threats such as **pathogens** (for example viruses and bacteria), cancer cells ([Schumacher and Schreiber, 2015]) and non-living substances, such as toxins, while at the same time exercise **self tolerance**, i.e. not damage friendly tissue ([Murphy et al., 2008, Ch. 1.1]). There are many components to the immune response, and this thesis looks at a type of **white blood cell** known as a **CD8$^+$ T cell** (CD standing for **cluster of differentiation** which are molecules on the cell's surface, the positive indicating a high number of such molecules). T cells are a type of **lymphocyte**, (named so because of the high number of them found in the **lymphatic system**) which seek out and destroy infected cells. T cells become activated during infection and enter a period of fast reproduction, known as the **expansion phase**, to produce a clonal army that can combat the threat ([Kaech et al., 2002b]). It is the behaviour of CD8$^+$ T cells during this expansion phase that this thesis focuses on.

Within this chapter we review the background of CD8$^+$ T cells. The first section subdivides the mature immune system into different subsystems to establish the place of CD8$^+$ T cells within them. The second and third section describes how T cells behave during an immune response, showing how the expansion phase relates part of the overall immune response. The fourth section looks at some of physical characteristics of T cells. Finally we review some of the techniques and terms used by experiments in papers, the data from which we use in our analysis. If the reader is familiar with immunological theory and experimental concepts this chapter can be skipped.

2.1 Segmentation of the immune system: Where T cells fit in

The immune system is often segmented into **innate** and **adaptive** subsystems (Figure 1).

The cellular part of the **innate immune system** recognizes threats by detecting molecular patterns common to pathogens ([Murphy et al., 2008, Ch. 2.1]). The innate system responds quickly, defeating most threats within a few hours or days, whereas the adaptive system usually takes seven-eight days to remove a threat, and because of this, the innate immune system deals with the majority of infections before they cause a response from the adaptive immune system. However, the innate system does not improve its response to a secondary infection ([Murphy et al., 2008, Ch. 2.1]).

The innate system's cellular component includes **phagocyte** cells that engulf targeted cells. Phagocytes in turn have subtypes, two of which are **dendritic** cells and **macrophages**, which have an additional role in helping activate the adaptive immune response ([Murphy et al., 2008, Ch. 1.3]). There are many other components of the innate system, examples of which are: **mast cells** which release histamine, to cause inflation, dilation of the blood vessels and attract more cells to combat the threat and **anatomical barriers** which block the pathogens path and include skin and internal membranes.

The **adaptive immune system** is present in all jawed vertebrates (it has a different form in

jawless vertebrates, [Alder et al., 2005]). Factors that distinguish it from the innate system include the ability to improve the response upon a threat reappearing and any two sampled cells from the adaptive system are likely to be specific to different threats. The adaptive immune system is split into its two most important components **B cells** and **T cells**, together called **lymphocytes** (some other lymphocytes, such as the **natural killer** cell belong to the innate immune system, these kill tumours and virus infested cells, [Alberts et al., 2002, Ch. 22]).

An **antigen** (antibody generator) is any substance that can trigger an adaptive immune response ([Alberts et al., 2002, Ch. 24]). These antigens are found on the surface of harmful cells, floating freely in the blood plasma or on certain kinds of host cells that present antigens and help induce a immune response, which are known as **antigen presenting cells** ([Murphy et al., 2008, Ch. 2.10]). A dendritic cell or macrophage are examples of an antigen presenting cell. These present antigens sampled from its environment or taken from parts of cells engulfed via phagocytosis. Lymphocytes of the adaptive immune response do not all respond to the same antigens, but each cell responds only to the small repertoire that will bind with its **T cell receptors** (**TCR**) or **B cell receptors** (**BCR**), which are different between any two cells, in a lock and key style mechanism, and thus, cells of the adaptive immune system are referred to as **antigen specific**. **Affinity** refers to the strength of the fit between the antigen and cell receptor and a stronger fit will be more likely to induce a strong immune response from the cell. Although there are only a small number of cells in a mouse that will respond to any specific antigen (est. 100-1000, [Schlub et al., 2010]) they rapidly divide to fight infection. This mechanism, where an initially small number of cells specific to an antigen, upon infection activate and rapidly divide is known as **Clonal Selection** ([Burnet, 1957]) and allows for the immune system to respond to a wide repertoire of possible antigens and threats.

B cells provide **humoral immunity** by producing **antibodies**, Y shaped proteins that lock onto their cognate antigen. These antibodies mark a cell for death, block the cell from functioning or cluster cells together making them less harmful ([Alberts et al., 2002, Ch. 24])).

T cells provide **cell-mediated immunity**. They are categorized into two important types **CD4$^+$** and **CD8$^+$** ([Murphy et al., 2008, Ch. 1.19]). CD4$^+$ cells, sometimes known as **helper cells** co-ordinate other parts of the immune response by releasing **cytokines**, signal proteins recognized by other cells ([Alberts et al., 2002, Ch. 24]). CD8$^+$ T cells target harmful cells expressing their cognate antigen and secrete **granzyme-B**, which triggers cell self-destruction in the target through a chemical cascade, known as **apoptosis**. Cells that perform this action are also called **cytotoxic** cells. This thesis concerns itself with CD8$^+$ T cells ([Alberts et al., 2002, Ch. 24]).

CD8$^+$ T cells are further categorized into **naive** cells, which are mature T cells that have not yet come into contact with their cognate antigen ([Murphy et al., 2008, Ch. 8.11]) and **activated** T cells, those that have. Naive cells are slow to divide and do not cause apoptosis in other cells. While the distinction between CD8$^+$ and CD4$^+$ T cell and naive and activated T cells is broadly agreed upon, there is some controversy and blurred lines around further subcategories of CD8$^+$ activated T cells.

Heterogeneity is observed in activated T cell populations, both at a cellular level in their behaviour (such as the rate of proliferation and cytokine production) and physical traits (size and cell markers) and a clonal level (i.e. family trees of T cells behave differently upon activation), while at the same time at a population level, the immune response remains robust ([Ahmed et al., 2009], [Kinjyo et al., 2015], [Buchholz et al., 2013], [Hodgkin et al., 2014]). It is generally agreed there are two subgroups of activated $CD8^+$ T cells: T cells that have the ability to induce apoptosis are called **effector** T cells (**TEF**), these are often short lived ([Alberts et al., 2002, Ch. 24], [Ahmed et al., 2009]). T cells that remain after the infection and have the ability to quickly mount a secondary immune response in the case of reinfection are called **memory T cells**. These memory cells are of a longer lineage than their effector counterparts. Memory cells are sometimes subdivided into **effector memory** T cells (**TEM**) and **central memory** T cells (**TCM**, [Wherry et al., 2003]). As the name suggests, effector memory cells sit on the overlap between the memory and effector groups, retaining some effector functionality ([Ahmed et al., 2009]), but less so than effector cells ([Schlub et al., 2010]) and these are also longer lived than effector cells. Whereas central memory cells express **CD62L** and **CCR7**, both of which cause the cell to home to the lymph tissue, the effector memory cells express less ([Ahmed et al., 2009]). Thus, effector memory and central memory are distinguished by their homing tendency, with effector memory residing in inflamed tissue, central memory residing in the **peripheral lymphoid tissue**. Effector memory cells are less likely to proliferate upon reinfection than their central memory counterparts ([Ahmed et al., 2009]) but may fulfil an effector role in the secondary response. Cells that reside in the **peripheral tissues**, where reinfection is most likely, are called **resident memory** cells and are sometimes considered separate subgroup ([Mueller et al., 2013], [Masopust and Schenkel, 2013]), however this thesis does not discuss this subgroup of memory further.

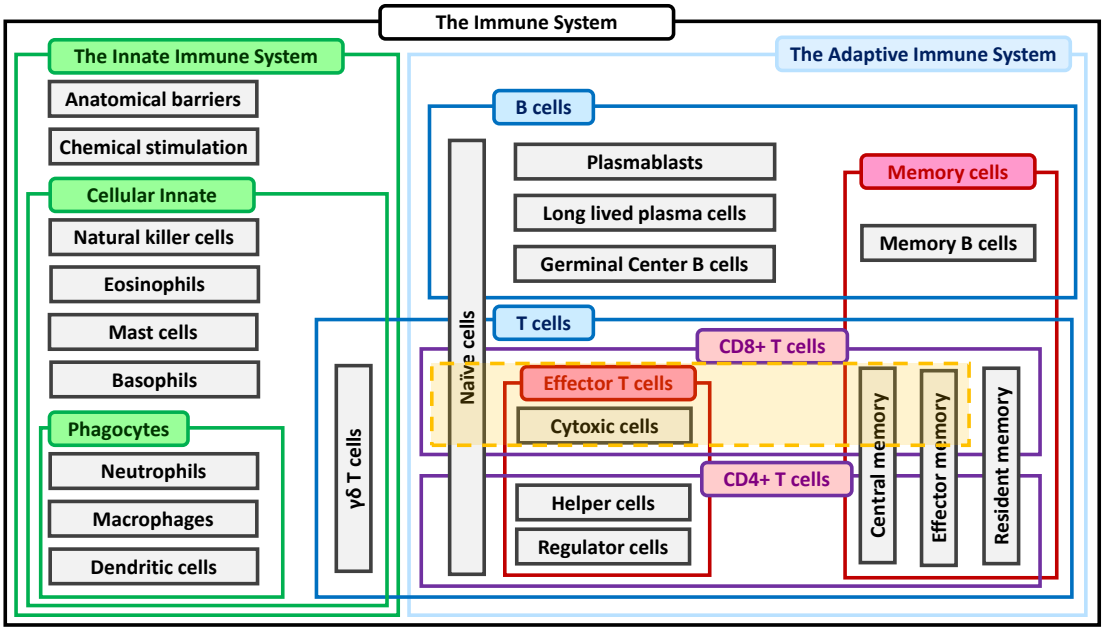


Figure 1: Basic components of the mature immune system. Orange shaded areas are those this thesis focuses on. A further categorization is that of lymphocytes: T cells, B cells and Natural Killer cells. The cells in the Adaptive Immune System and the cellular part of the Innate Immune system are known white blood cells (although mast cells and dendritic cells primarily do not reside in the blood). The diagram is not exhaustive ([Alberts et al., 2002],[Murphy et al., 2008]).

2.2 Stages of an adaptive immune response: What T cells do

The behaviour of $CD8^+$ T cells during an adaptive immune response is divided into four phases: **activation**, **expansion**, **contraction** and **memory**. Figure 2 provides a representative caricature of this response.

In order to activate and trigger a full immune response, a naive T cell requires at least two signals ([Murphy et al., 2008, Ch. 8.12], [Marchingo et al., 2014]):

- Signal 1. Cognate antigen. The T cell's TCR must detect its cognate antigen presented on the **Major histocompatibility complex** (**MHC**, a type of glycoprotein) on the surface of an antigen presenting cell ([Murphy et al., 2008, Ch. 1.20]). This interaction often occurs in the spleen or lymph nodes (Table 1). The antigen signal is all that is necessary for activation, but without the other signals, it triggers a weak immune response.

- Signal 2. Co-stimulation. This comes from the antigen presenting cell via expression of molecules such as **CD80**, **CD86** or **CD70** which bind to **CD28** or **CD27** on the T cell surface ([Murphy et al., 2008, Ch. 8.12, 8.14]).

- Signal 3. Detection of cytokines. For example the cytokine **interleukin-12** (**IL-12**), although

other combinations could also produce a response ([Murphy et al., 2008, Ch. 8.12], [Marchingo et al., 2014]).

To launch a substantial response signal one is necessary and either signal two or three are required in addition. If the T cell receives deficient stimulation it may not activate or activate but trigger only a weak immune response. During the expansion phase, the strongly activated T cells secrete further cytokines, for example **interleukin-2** (**IL-2**), which induces further proliferation and differentiation of other T cells.

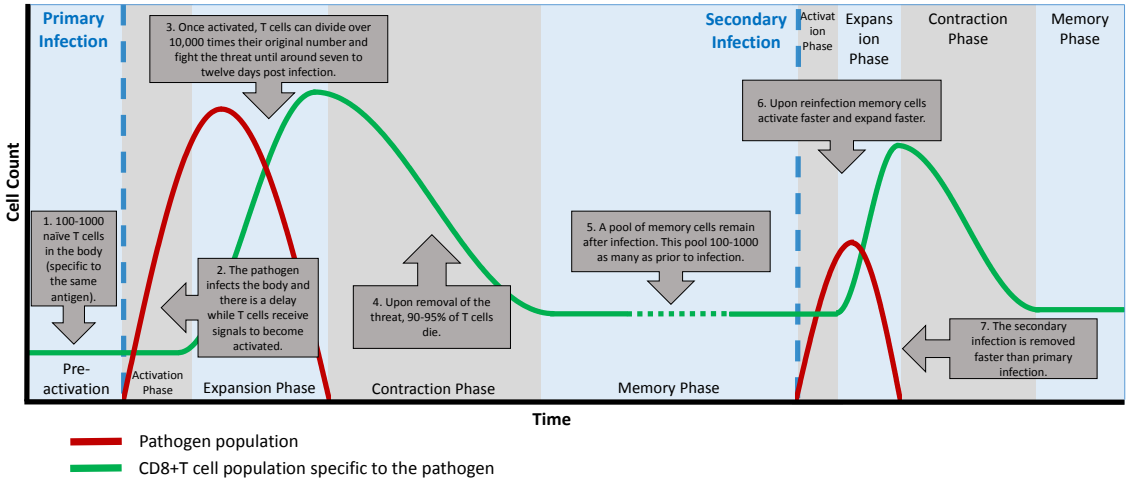


Figure 2: Non-data schematic of CD8$^+$ T cells at different stages of the immune response. Adapted from [Kaech et al., 2002b].

.

In order to fight the infection, the small number of antigen specific T cells rapidly divide, and can in some cases reach upwards of 10,000 fold its initial number ([Schlub et al., 2010]) depending on the strength of activation. This is known as the expansion phase. Cells change their surface markers and behaviour during the expansion phase when compared to their naive counterparts. For example, cells increase in size, increase their expression of **CD44** and decrease their expression of **CD62L**, ([Murphy et al., 2008, Ch. 10.16], [Kinjyo et al., 2015], [Buchholz et al., 2013], [Schlub et al., 2010]). The mechanisms that drive T cell expansion are not fully understood. One possible candidate for regulating this phase is **Myc**, a **transcription factor** (a protein that converts DNA to RNA), that has its levels set after activation and which may decide the number of divisions a cell goes through ([Heinzel et al., 2017], [Man and Kallies, 2015]).

The contraction or death phase occurs after the expansion has ended, when the threat has been removed and about 90-95% of the T cells created during the expansion phase die off ([Schlub et al., 2010]). However, a pool of memory cells remain after the expansion phase has ended, and these

respond with greater effectiveness in case of re-infection. These memory cells are quicker to respond upon antigen detection (division, differentiation and production of cytokines) and also start from a more numerous baseline (100-1000 fold above the initial frequency) than the naive cells that fought the primary infection ([Murphy et al., 2008, Ch. 10.16]).

Body part	**Relevance to T cells**	**Background information**
Bone Marrow	Produces immature native double positive ($CD8^+$ $CD4^+$) T cells from haematopoietic stem cells.	**Haematopoiesis**, blood cell creation ([Alberts et al., 2002, Ch. 24]). Part of the **primary lymphoid tissues**.
Thymus	Allowing double positive T cells to reach maturity as either $CD8^+$ or $CD4^+$ naive T cells. Removes harmful and useless T cells.	Haematopoiesis of T cells ([Alberts et al., 2002, Ch. 24]). Part of the primary lymphoid tissues.
Spleen	Holds a large number of T lymphocytes. T lymphocytes get activated here by antigen presenting cells.	Immunity, creates antibodies and removes antibody covered pathogens, filters and recycles the bloods. ([Murphy et al., 2008, Ch. 1.15]). Part of the **secondary lymphoid tissue**.
Lymph Nodes (Secondary lymphoid tissues)	T lymphocytes get activated here by antigen presenting cells. Acts as transportation for T cells.	Immunity, filters the blood ([Murphy et al., 2008, Ch. glos.]). **Mesenteric lymph nodes** (**MLN**) are the largest lymph nodes and found in the intestine. Part of the secondary lymphoid tissue.
Blood	Transport for T lymphocytes.	Transport of resources around the body ([Alberts et al., 2002, Ch. 22]).
Peripheral tissues and/or site of infection	Mature effector T cells reside here to fight infection. Later residential memory T cells remain here to fight reinfection.	For example, lungs in the case of a flu virus ([Pan et al., 2017]).

Table 1: Relevance of organs and locations in the body and with respect to T cells.

2.3 Markers and phenotype: $CD8^+$ T cells traits

A **phenotype** is a segregation of a group of cells by their observable characteristics ([Alberts et al., 2002, Ch. 24]). For example, cell size, expression of molecules on the cell surface or fluorescent reporters. We often wish to segregate cells by their functions or potential functions and we refer to these as different **cell types** or **cell subsets**. A common issue in immunology is accurately partitioning cells into cell types from information on their phenotype. When a phenotype correlates with a cell-type,

we say the phenotype **purifies** or **enriches** for that cell-type. **Gating** is the discrete categorization of cells based on their surface markers, the most simple categorization is be decided by being over or under a threshold (**gate**), which can be non linear if looking at multiple phenotypic traits. Table 2 shows some phenotypic traits of T cells against their proposed cell-type and these are used in the experiments we discuss.

While the cell-type is defined by a cell's functionality ([Wherry et al., 2003], [Sarkar et al., 2007]), in some literature it is defined by a cells markers ([Kinjyo et al., 2015], [Buchholz et al., 2013]). This sometimes leads to an issue in the field immunology of unaligned definitions. This thesis assumes CD62L markers to be a good indicator of memory functionality, but we note that fitting mathematical models the differentiation of markers is not the same as fitting directly to the differentiation of functionality.

2.4 Experimental procedures: Observing T cells

Experiments are said to be **in vivo** (within the living) if they observe biological phenomenon within the whole living organism. **In vitro** (within the glass) refers to experiments observing a phenomenon in a controlled environment outside the body ([Alberts et al., 2002, Ch. glos.]).

Flow cytometry is a method of observing properties of cells, such as the expression of markers on the cells surface or measuring cell size ([Murphy et al., 2008, Ch. A-22]). This is done by suspending cells in fluid and quickly passing them through beams of light. The cell's surface markers are indicated by the refraction of light passing through dye molecules bound to them. Flow cytometry can be used to purify cells with certain characteristics to be used in experiments, this is known as **Fluorescence-activated cell sorting** (**FACS**) or flow cytometry can evaluate cells harvested at the end of an experiment.

Mice are often bred or genetically engineered for special properties related to the experiment. **Transgenic** means that an organism has artificially introduced or removed DNA in its genetic code ([Alberts et al., 2002, Ch. 8]). Sometimes mice are **cross-bred** with other mice so that their offspring acquire desired properties. In these experiments **Black-6 mice** (**C57BL/6**) are often used, a specific strain of laboratory mouse. **OT-1 mice** are mice genetically engineered so that all their T cells are specific to the **SIINFEKL** antigen (an antigen from **chicken ovalbumin**, **OVA**), allowing researchers to introduce pathogens expressing OVA and induce a controlled immune response ([Kinjyo et al., 2015],[Buchholz et al., 2013]). When introducing a pathogen to a host, it is not always possible to have an exact count of the pathogen introduced, so the concentration of pathogens is often measured by their visible effect: **colony-forming units** (**cfu**) are the counted number of colonies a substance makes when applied to agar plates ([Sieuwerts et al., 2008]). **Plaque forming units** (**pfu**) is a measure of the visible zones of infection (plaques) a substance makes in a layer of host cells ([Cammack et al., 2006]).

A common tool used in *in vivo* immunology experiments is **adoptive transfer**, where cells are taken from a donor and transferred to a host. This is combined with breeding techniques so that progeny of host cells have distinguishing cell markers from the donors, and thus can be distinguished by flow cytometry after extraction.

Time lapse imaging/ microscopy is a technique to observe cells over time where images of the experiment are taken at intervals. An example of this in immunology is images being recorded of cells proliferating in wells.

Carboxyfluorescein succinimidyl (**CFSE**) is a fluorescent dye used to stain cells so that that they are fluorescent under specific light frequencies ([Lyons and Parish, 1994]). As a cell divides, the dye is divided equally between its progeny, so that the intensity of fluorescence from a cell can be used to estimate the number of divisions between the observed cell and the dyes introduction. **Cell-trace violet** (**CTV**) is a similar dye using a different light spectrum which has the advantage it can be used in conjunction with ubiquitous **GFP** (**green fluorescent protein**) reporters.

Mice can been genetically engineered so that their cells express **Fluorescent ubiquitination-based cell cycle indicator**, (**FUCCI**, [Sakaue-Sawano et al., 2008]). Cells from such mice become reversibility fluorescent in either **mKusabira-Orange 2** (**mKO2**, orange) or **mAzami-Green** (**mAG**, green) under the appropriate frequency of light, and this fluorescence is determined by the stage of the cell life-cycle. Cell that are $mKO2^{+}mKO2^{-}$ are more likely to be **quiescent** (not dividing). Other combinations of fluorescence ($mKO2^{-}mKO2^{-}$, $mKO2^{-}mKO2^{+}$ $mKO2^{+}mKO2^{+}$) indicate the cell is at some stage of division. This information can be used to gage how quickly a population is proliferating.

Sometimes an experimental setup allows for the cells being observed to be identified as being descended from the same initial naive cell, i.e of the same family. When this is the case we use the convention throughout of saying the data are available at a **clonal** or **family** level. If these data are not available we say the data are at a **cohort** or **population** level.

Cell Marker/Trait	Role	Purifies for...
CD4	Assists with activation.	Very low in cytotoxic cells, very high in helper cells.
CD8	Assists with activation.	Very high in cytotoxic cells, very low in helper cells.
CD62L	Causes cells to home to secondary lymphoid tissues.	High in central memory and naive cells. Low on effector and effector memory cells ([Kinjyo et al., 2015], [Buchholz et al., 2013]).
CD27	Co-stimulation. Binds to CD70.	High on central memory and effector memory, low on effector cells. Expressed on naive cells but it is unclear if this can be used to identify them. (Hendriks et al. [2000], Baars et al. [2005], [Kinjyo et al., 2015], [Buchholz et al., 2013]).
CD127	IL-7 receptor	High on naive, central memory and effector memory cells. Low on effector cells. ([Schlub et al., 2010]).
CD44	Activation and homing.	Low values indicate naivety ([Kinjyo et al., 2015], [Buchholz et al., 2013]).
CD45	Unknown.	In experiments cells are engineered to give differing expressions to allow T cell lineages to be distinguished([Murphy et al., 2008, Ch. App II]).
CD90 (Also known as Thy-1)	Augments signalling.	In experiments cells are engineered to give differing expressions to allow T cell lineages to be distinguished ([Murphy et al., 2008, Ch. App II]).
CCR7	Causes cells to home to secondary lymphoid tissues.	High in central memory and low in effector memory ([Sallusto et al., 1999]).
Cell size	N/A	Effector cells are larger, memory and naive cells are smaller ([Kinjyo et al., 2015]).
Speed of division	N/A	Higher speeds of division in effector cells ([Kinjyo et al., 2015], [Buchholz et al., 2013]).
Interleukin-2 (IL-2) production	Helps T cells activate, proliferate and differentiate. Encourages cells to differentiate to effectors.	Higher production indicates memory ([Alberts et al., 2002, Ch. 24], [Buchholz et al., 2013]).
Interferon-γ (IFN–γ) production	Encourages macrophage activation and expression of MHC molecules.	Indicates multifunctional cells ([Buchholz et al., 2013],[Murphy et al., 2008, Ch. AppIII])
T box transcription factor (T-bet)	Regulates T cell development.	Associated with effector cells ([Buchholz et al., 2013],[Tantisira et al., 2004])
T-box transcription factor eomesodermin (Eomes)	Required for memory generation.	Associated with memory cells ([Buchholz et al., 2013], [Knudson et al., 2017])
Killer-cell-lectin-like-receptor-G1 (KLRG1)	Helps activation of T cells.	Effector cell marker. ([Montufar-Solis et al., 2017],[Buchholz et al., 2013])

Table 2: Phenotypic indicators for different T cell types.

3 Description of experiments evaluating the order of CD8$^+$ T cell differentiation

During an adaptive immune response, at some time after activation, the expanding population of CD8$^+$ T cells becomes heterogeneous, containing a sub-population of effector cells whose offspring die off during the contraction phase, and a sub-population of memory cells of longer lineage (see Section 2.2 for a details on the expansion phase). Many core aspects of the mechanisms that govern this differentiation and expansion remain subject to controversy ([Ahmed et al., 2009]). Contemporary questions discussed in the literature include:

- What CD8$^+$ T cell categorisation of cell subtypes are most representative of the underlying mechanics that drive the immune response and what phenotype best enriches for these cell types?
- Upon activation, assuming there is an inherent ordering to T cell subtypes, which subtypes do naive cells first differentiate into and is this behaviour consistent between different naive cells?
- Which T cell subtypes retain the potential to become other T cell subtypes?
- What drives the heterogeneity of CD8$^+$ T cell behaviour? Candidates could include: type and strength of initial activation, cytokines, the location within the body, an internal stochastic mechanism within the cells or a combination.
- Can the answers to the above be formalised into a model, quantitative or qualitative, so as to give a comprehensive view of the expansion phase of the immune response?

This thesis aims to answer one of the above questions. It examines the order which naive T cells differentiate into effector and memory subsets. The questions are difficult to tackle independently. For example, problems with cell subtype and phenotype classification have been put forward as causes of confusion and seemingly contradictory evidence between theories ([Ahmed et al., 2009]). Likewise if activation potency is what drives variation in CD8$^+$ T cell behaviour, different experimental setups may yield different results.

The following subsection gives an overview of different standpoints from past papers on the question of T cell differentiation.

The subsections 3.2 to 3.5 review four articles reporting results from similar experimental setups: adoptive transfer of OT-1 cells. Firstly, we examine two recent articles that come to different conclusions as to the order of T cell differentiation: [Buchholz et al., 2013] and [Kinjyo et al., 2015]. [Buchholz et al., 2013] is of interest to us because of its novel methodology of tracking cells at a clonal level and its method of fitting a mathematical model to statistics observed from experiment. We seek to adapt

this method to data available in other papers (which we do in Chapter 5). The experimental and analytical method described in [Buchholz et al., 2013] suggests that memory precursor cells appear before effector cells (for an overview of the mathematical methods and models used in [Buchholz et al., 2013] see Section 4.3). [Kinjyo et al., 2015] is of interest because of its use of generational and proliferation speed data to give insight into the mechanics of the early days of the adaptive immune response. The model suggested in [Kinjyo et al., 2015] is that effector cells appear before memory precursor cells.

One key difference in the experimental setup between [Buchholz et al., 2013] and [Kinjyo et al., 2015] is the number of naive cells adoptively transferred. The papers [Badovinac et al., 2007] and [Schlub et al., 2010] describe the effects of different adoptive transfer numbers on OT-1 experimental results. While the focus of the papers [Badovinac et al., 2007] and [Schlub et al., 2010] is not the order of T cell subtype differentiation, they also employ an experimental setup similar to that of [Buchholz et al., 2013] and [Kinjyo et al., 2015] and we use these data to supplement the mathematical model fitting later in Chapter 5.

In the last subsection we compare the different experimental setups, results and conclusions between these four papers and suggest ways to reconcile them.

3.1 History of viewpoints on CD8$^+$ T cell differentiation

There are many ideas and theories behind CD8$^+$ differentiation during the expansion phase. To give an overview of opinion we simplify the problem down to three basic versions of championed theories (Figure 3, [Ahmed et al., 2009], [Kinjyo et al., 2015]):

- **Linear Effector First Model:** Naive cells first differentiate into effector cells, which proliferate. Later in the immune response, these cells either differentiate into memory cells or continue to proliferate as effectors before eventually dying.
- **Linear Memory First Model:** Naive cells first differentiate into memory cell precursors, which proliferate and differentiate into effector cells. Once the expansion phase is completed the effector cells die.
- **Bifurcation Model:** A naive cell has the potential to divide into either effector or memory subsets which then continue to proliferate inheriting their cell-type. The effector cells die at the end of the expansion phase while the memory cells remain.

These three theories are merely basic representations of more diverse and subtle theories. For example the above assume three basic CD8$^+$ cell-types (naive, effector and memory cell), however further cell-types have been included in models ([Wherry et al., 2003]). These three models are also not exhaustive, for example they all exclude the possibility of **bidirectional** differentiation (i.e. cells being able to return to states they previously differentiated from) between cell-types.

The Linear Effector First Model could be considered the traditional view and there are a number of earlier papers suggesting that T cells go through an effector phase before differentiating into memory cells ([Opferman et al., 1999],[Ahmed et al., 2009]). The issue was complicated by the observation that some T cells had an overlap between memory and effector functionality and phenotypic markers. This subcategory is often referred to as effector memory cells (TEMp, the p standing for precursor) to distinguish them from effectors (TEF) and central memory cells (TCMp), but the exact role and classification of TEMp cells remains a controversy in itself (see Section 2.3 on phenotypic identification). There was also some early evidence that central memory appears before memory [Sallusto et al., 1999]. In 2007 the article [Chang et al., 2007] proposed asymmetric division theory to explain heterogeneity with $CD8^+$ T cells. This became a popular theory for explaining the variable population dynamics of T cells families as being dependent on how naive cells are activated ([King et al., 2012]). In 2013 two papers were published jointly in Science [Buchholz et al., 2013] and [Gerlach et al., 2013] that used distinct systems to further address these issues. [Buchholz et al., 2013] developed a methodology of fitting mathematical models to observed summary statistics of $CD8^+$ T cell phenotypes and found the Linear Memory First Model to be the best fit to observed data. We do not review the sister paper [Gerlach et al., 2013] so closely, as the measurements they use to identify families, **polymerase chain reaction** (**PCR**) amplification of retrovirally installed **DNA tags**, does not scale linearly with cell counts, meaning the proportional data gathered is unsuitable for the fitting method we use.

Table 3 provides a non-exhaustive list of papers from the last 20 years supporting or giving insight into the three basic theories. The papers listed as supporting theories are only a simplification. For example [Kinjyo et al., 2015] is a article that supports the Linear Effector First Model, but accepts that perhaps a naive cell may still have the ability to directly differentiate into a memory cell, however this it is not a driving force behind the population dynamics.

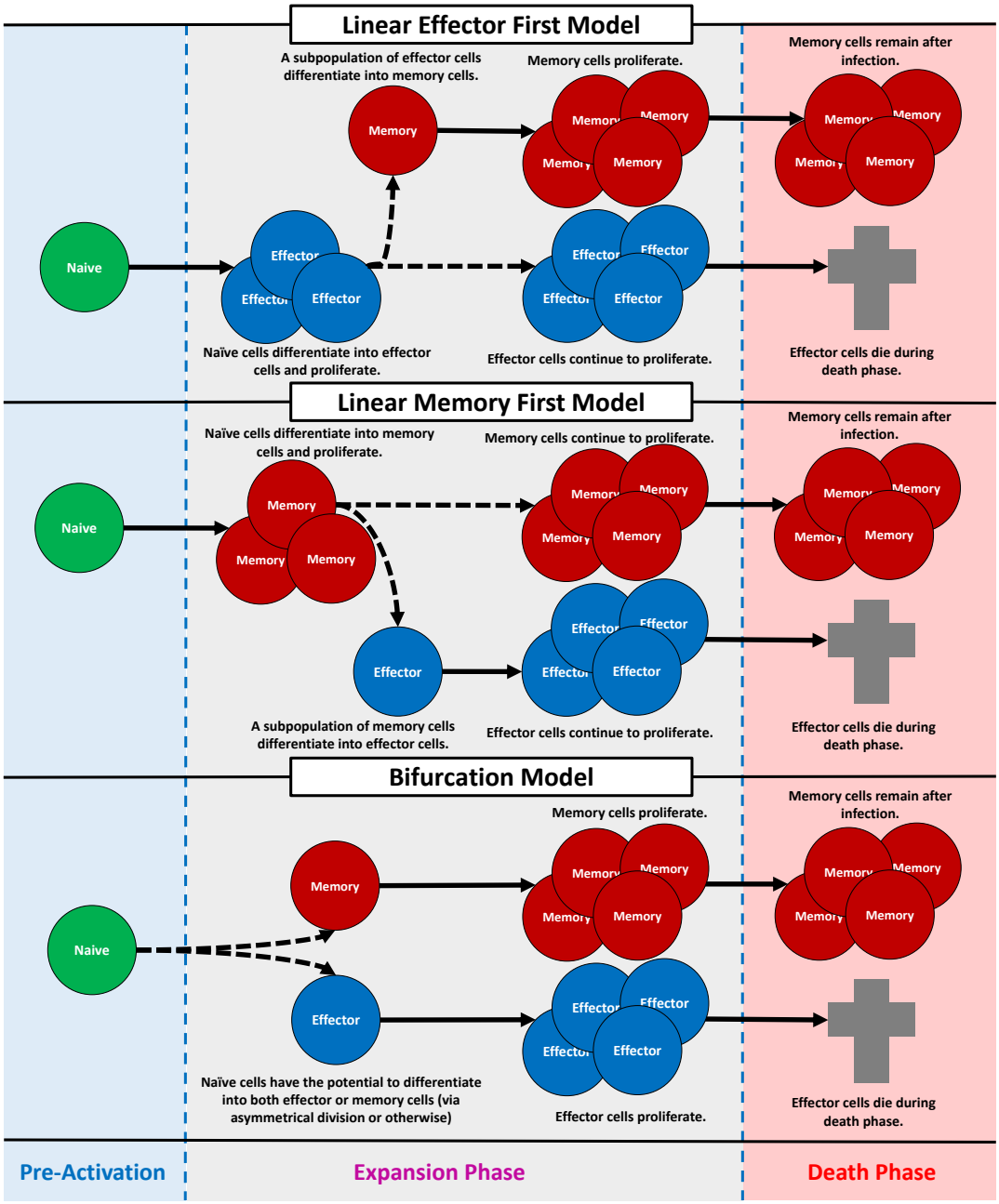


Figure 3: Three simplified models showing possibilities of ways CD8$^+$ T cells differentiate during the expansion phase. In the Linear Effector First Model, naive cells differentiate first into effector cells and proliferate, only then do they differentiate into memory cell precursors. In the Linear Memory First Model the reverse is true. In the bifurcation model naive cells have the potential to become either phenotype. Models listed are not exhaustive.

Reference	Model	Notes
[Opferman et al., 1999]	EF	Memory cells were found to be the progeny of cytotoxic effectors.
[Jacob and Baltimore, 1999]	EF	The effectors that develop into memory cells can be identified in the primary response.
[Sallusto et al., 1999]	MF	Separation of memory into effector memory and central memory. TCMp appears before then TEMp (human).
[Kaech et al., 2002a]	-	Memory cell precursors were present at the peak of the immune response but their functional traits as memory T cells: self-renewal and rapid recall to antigen, did not develop until later.
[Wherry et al., 2003]	EF	TEMp's are the intermediate phenotype in the linear pathway. Naive $\rightarrow$ TEF $\rightarrow$ TEMp $\rightarrow$ TCMp.
[Bouneaud et al., 2005]	-	Bi-directional differentiation between TEMp and TCMs. Memory cells are present throughout the expansion phase stage.
[Bachmann et al., 2005]	-	In the expansions phase the T cell population initially expresses higher levels of $CD62L^+$ which are down-regulated and later increase again in the death phase.
[D'Souza and Hedrick, 2006]	-	Timing of recruitment of individual T cell clones determines the population dynamics of the subsequent immune response. Activation theory: duration and strength of activation signal regulates differentiation.
[Masopust et al., 2006]	-	Memory cell phenotype differed depending on the number re-challenges and that secondary and tertiary responses resulted in the memory CD8 T cells with effector-like properties.
[Sarkar et al., 2007]	EF	Some TEMp do not progress to TCMp. The majority of TCMp's are derived from TEMp's.
[Chang et al., 2007]	B	Activation causes the cell to split asymmetrically with the two progeny typically having different phenotypic traits that are inherited.
[Joshi et al., 2007]	EF	Signals from the innate immune system regulate which effector cells become memory cells.
[Badovinac et al., 2007]	-	Number of cells adoptively transferred can shorten the expansion phase.
[Harrington et al., 2008]	EF	There are parallels to the Linear Effector First Model for CD4 T cells.
[Sarkar et al., 2008]	EF	IL-2 and duration of antigen stimulation define which effectors become memory cells.
[Teixeiro et al., 2009]	B	Effector and memory states are separable fates, determined by differential TCR signalling. Activation theory: duration and strength of activation signal regulates differentiation.
[Schlub et al., 2009]	EF	Support that differentiation is division-linked.
[Schlub et al., 2010]	EF	The number of divisions a cell goes through is linked to level of CD62L. Further evidence that the number of cells adoptively transferred can shorten the expansion phase.
[King et al., 2012]	B	Activation strength above a threshold causes asymmetric division and below the threshold leads to an abortive immune response.
[Gerlach et al., 2013]	-	Single cell response is not predictable and a predictable immune summary statistics are achieved through averaging of populations.
[Buchholz et al., 2013]	MF	Clonal summary statistics match to the Linear Memory First Model.
[Marchingo et al., 2014] / [Marchingo et al., 2016]	-	Strength of affinity and co-stimulation can be summed to predict strength of immune response.
[Kinjyo et al., 2015]	EF	An initial homogeneous population effectors after infection is later followed by a sub-population of memory cells later.
[Heinzel et al., 2017]	-	Linking levels of Myc to cell expansion.

Table 3: List of papers offering evidence for the order of differentiation during the expansion phase. Model column is a suggestion of which model the article supports (EF = Linear Effector First Model, B = Bifurcation Model, MF = Linear Memory First Model, - = Not directly supporting one of the described models) Earlier the history of immunology, what might be called the Linear Effector First Model was most broadly supported. In recent years asymmetric theory became popular. More recently [Buchholz et al., 2013] suggested a Linear Memory First Model to fit best to their data.

3.2 Experiment: [Buchholz et al., 2013]

The paper [Buchholz et al., 2013] reported on a collection of experiments designed to determine clonal properties of the $CD8^+$ T cell expansion phase and then fitted models to statistics gathered from these data. This paper was novel in its use of congenic markers allowing for cells of the same family to be identified. The paper also describes and developed the mathematical model that we adapt and apply to data taken from other papers that report on experiments with similar systems, but without familial information (Chapter 5).

3.2.1 [Buchholz et al., 2013] experimental setup

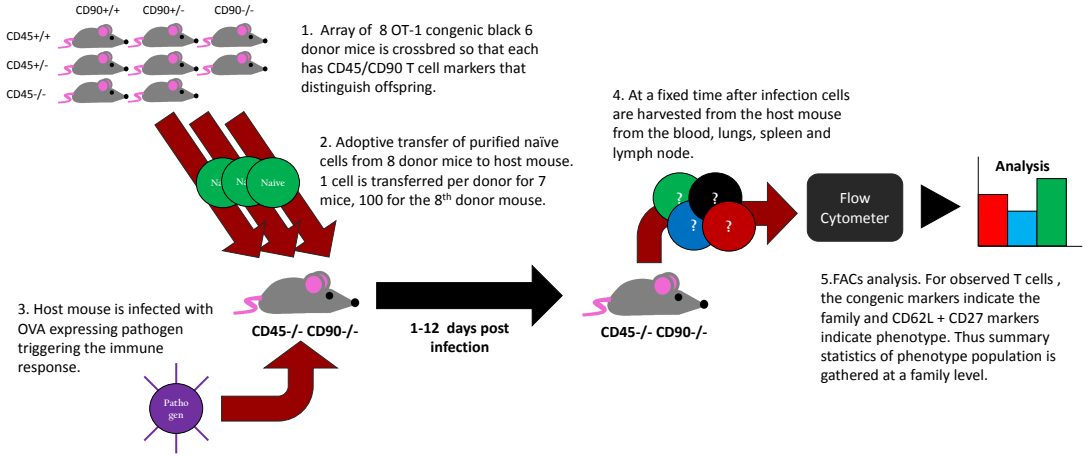


Figure 4: Summary diagram of the experimental adoptive transfer setup in [Buchholz et al., 2013]. Congenic OT-1 mice with differentiated labels are used for an adoptive transfer experiment where the cell markers are used to identify cells of the same family. Harvesting was done on day eight and 12 on 47 host mice. Similar experiments harvested cells on days one, two, three, four, six and eight post infection but without the array of congenic markers returning data at a population level for these time periods and also only harvesting from the spleen.

The authors of [Buchholz et al., 2013] performed an OT-1 adoptive transfer experiment. The experimental set up enabled authors to distinguish the transferred donor progeny of multiple naive cells within a single mouse based on unique cell surface markers, thus allowing for clonal family data to be observed. Donor mice were engineered to they express both **CD45.1** and **CD90.1** markers homozygously, heterozygously or not at all, thus allowing for nine distinguishable combinations of markers in total (Figure 4). There were 47 black-6 host mice chosen from double heterogeneous marker pool. For each of these host mice there were eight OT-1 donor mice, each with a different marker. Purifying for $CD44^-CD8^+$ allowed naive cells to be extracted from the donors. For seven of the donor mice one naive cell was transferred, and for the remaining donor mouse 100 cells where transferred in order to observe the effects of scaling when transferring multiple cells. The host was then

infected with 5×10^3 **Listeria monocytogenes**, a pathogenic bacteria, expressing OVA (**LM-OVA**). This caused the activation of the OT-1 naive $CD8^+$ T cells in the mice. At day eight post infection, the cells were harvested from host mice. Cells were taken from the spleen, lymph nodes (axillary, inguinal, brachial, cervical, lumbar, mesenteric) and lungs. The cells observed through flow cytometry were gated into three phenotypes: the central memory precursor (TCMp; $CD27^+CD62L^+$), effector memory precursor (TEMp; $CD27^+CD62L^-$) and effector cells (TEF; $CD27^-CD62L^-$).

The data collected was used to generate summary statistics for phenotype populations: the average cell count per family, coefficients of variation and correlation coefficients between phenotypes. These summary statistics were used to choose a best fitting model of immune response from a family of 304 models with different orders of differentiation between phenotypes, one way or bidirectionally. For details of the models, the fitting and the mathematics behind it, see Chapter 4.

Other experiments described in [Buchholz et al., 2013] that harvested cell at different times post infection and had slightly different experimental set ups. In those experiments that harvested cells one, two, three, four, six, days post infection, the CD45 and CD90 array of surface markers was not used, rather only one marker was used to distinguish the host from the donor cells which meant only proportional data were available. Rather than one cell transferred per donor there was an adoptive transfer of 10,000 OT1 cells to get proportional data in the case of cells harvested one, two, three and four days post infection and 100 OT1 cells were transferred for the equivalent experiments that harvested six or eight days post infection. In that experiment cells were taken from the spleen only. This population data, along with CFSE data, was used to compare expectations against the proposed model.

In a similar experiment, but harvesting on day 12, other indicators of phenotype were measured. These included IL-2, IFN-γ producers, T-box transcription factor expressed in T cells (T-bet), eomesodermin (Eomes) and Killer-cell-lectin-like-receptor-G1 (KLRG1). While these were not included in model fitting they were used to show phenotypic trends within the data.

3.2.2 [Buchholz et al., 2013] findings

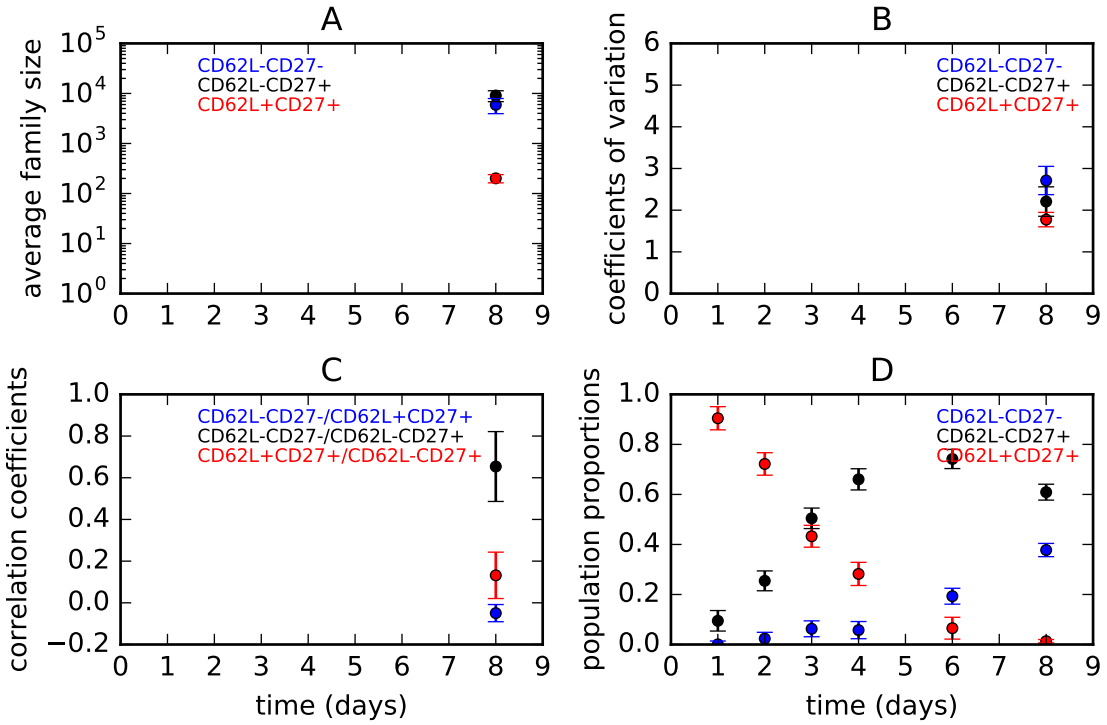


Figure 5: Summary statistics reproduced from the data reported in [Buchholz et al., 2013]. The top left graph shows the average count of cell phenotypes on day eight post infection, top right is the coefficients of variation of each phenotype, bottom left is the correlations between phenotype populations and bottom right is proportions of populations. Error bars represent the sample variation achieved by bootstrapping (see Section 4.1.2 for a description of bootstrapping). All data but that in the bottom right graph comes from the experiment that used the OT-1 array of congenic markers to mark data at a clonal level. The bottom right graph data came from a population level adoptive transfer experiment that harvested cells at different times. The proportional data shows that the $CD62L^+$ cells initially have a large proportion of population at days one to two post infection but over the days they lose share of the population to the faster growing $CD62L^-$ cells. The model defines non naive phenotypes as central memory precursor (TCMp; $CD27^+CD62L^+$), effector memory precursor (TEMp; $CD27^+CD62L^-$) and effector cells (TEF; $CD27^-CD62L^-$). These data were manually read off the data points in graph from their supplementary material S26. Proportional data were taken from Figure 4E.

The summary statistics showing phenotype post infection from the experiments are presented in Figure 5. From 304 possible differentiation networks two models fit the observed data adequately:

- Naive cells differentiate into TCMp cells ($CD62L^+CD27^+$) a few days after activation. TCMp cells have a chance to differentiate into TEMp cells ($CD62L^-CD27^+$) or proliferate. Finally

TEMp cells have a chance to differentiate into TEF cells ($CD62L^{-}CD27^{-}$) or proliferate. The proliferation rate increases for each successive cell-type.

- This model is the same as above, but with the additional small chance of naive cells skipping the differentiation into TCMp cells ($CD62L^{+}CD27^{+}$) and directly differentiating into a TEMp cell ($CD62L^{-}CD27^{+}$).

See Figure 6 for a qualitative description of the besting fitting network of differentiation model. For a description of the recreation of this model, including analysis of checks of observations against predictions, see Section 5.1. Other models were found to be poor fits to the observed data. In addition, the models that best fitted to day eight summary statistics were found to be a good fit when compared against the proportional data at a population level taken from days one to eight.

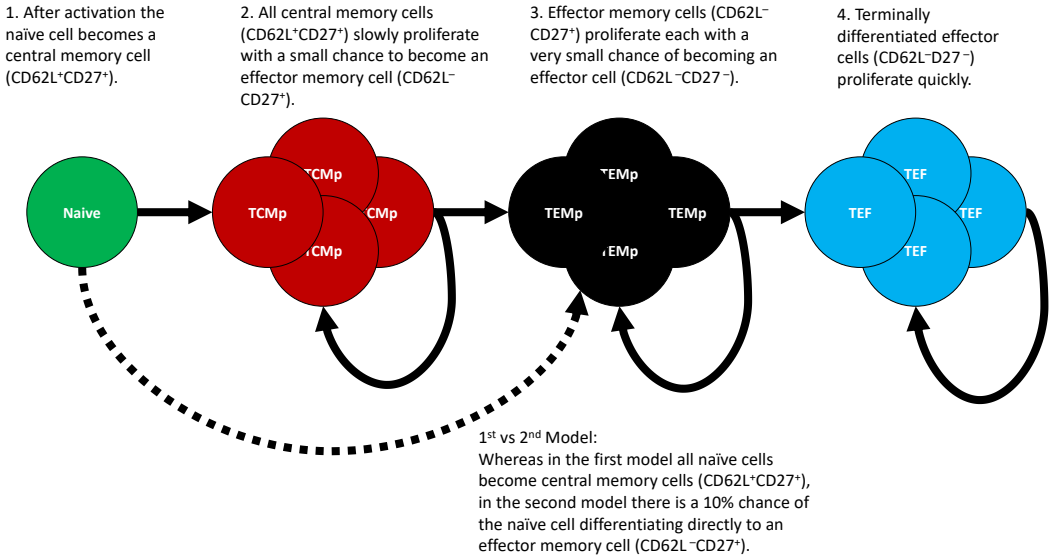


Figure 6: Two best fitting models of $CD8^{+}$ T cell differentiation proposed in [Buchholz et al., 2013]. The first model is a linear differentiation model, where naive cells first go through a memory stage before differentiating into an effector memory cells and finally differentiating into quickly proliferating effectors. The second model is the same as the first but with a 10% chance for naive cells to differentiate straight into effector memory cells after activation.

From the data harvested from the day 12 experiment, additional insights were drawn:

- At day 12 post infection most families had well below the mean of 4000 cells (**dwarf families**) and a few families generating up to 70,000 descendants (**giant families**)
- The dwarf families positively correlated with increased expression of $CD27^{+}CD62L^{+}$ and thus correlated to a memory phenotype, and the giant families expressed less of those markers, in-

dicating a correlation to the effector phenotype. Other indicators of phenotype (IL-2, IFN-γ, Enomes expression and T-bet expression) observed on day 12 post infection confirmed the correlation of dwarf families tending to the memory phenotype and giant families to the effector phenotype.

- At day 12, out of the 329 single naive donor cells (47 mice, 7 donors each), progeny from 93 cells (28%) were recovered.
- The variance in family profile showed that $CD8^+$ T cell immune response was only predictable at a population level. Under ten families cause the population to vary, but 50+ families give a predictable response.
- The article identified initial events in a family history driving its later profile. The majority of the population was accounted for by a few families.

3.2.3 [Buchholz et al., 2013] auxiliary experiments and results

A number of additional experiments and results are reported to give additional nuance to conclusions.

Differences of locations: As a check, the authors of [Buchholz et al., 2013] measured total cell counts in the spleen and observed that it correlated with amalgamated cell counts in other areas of the body.

Effects of different pathogen: Changing the virus to vaccinia virus–expressing OVA (Vaccinia-OVA)s, with 5×10^6 pfu injected, yielded similar results in terms of family size and variance of response.

3.3 Experiment: [Kinjyo et al., 2015]

The authors of [Kinjyo et al., 2015] performed a similar set of OT-1 adoptive transfer experiments to that of [Buchholz et al., 2013], however [Kinjyo et al., 2015] measured $CD8^+$ T cell behaviour at a population level rather than a clonal one. This article was novel in its observation of temporal data on cell proliferation speed. Unlike the [Buchholz et al., 2013] paper, no mathematical model was employed. Rather a qualitative one is proposed. Later we adapt the mathematical model [Buchholz et al., 2013] to make it suitable to analyse data acquired from experiments described in the [Kinjyo et al., 2015] paper (presented in Chapter 5).

3.3.1 [Kinjyo et al., 2015] experimental setup

Experimental Set Up One: Single adoptive transfer. Crossed FUCCI mice, OT-1 mice and mice with CD45.2 marker, had their cells labelled with CTV and adoptively transferred in Black-6 mice

(Figure 7). 10^6 of these donor cells were adoptively transferred into the host mouse (CD45.1). The host mouse was then infected with 100 pfu of influenza A virus **PR8** expressing OVA (**PR8-OVA**).

At day two, three, four and seven post infection cells were harvested from the MLN, lungs and spleen and separately analysed. FUCCI expression and CTV fluorescence allowed for information to be gathered on proliferation rate. To identify correlation of phenotypes, FACS analysis was used to measure the following on every cell: CD62L, CD44, CD27, CD71, Ly6C, CXCR3, KLRG-1, IL-7Rα and cell size was also observed.

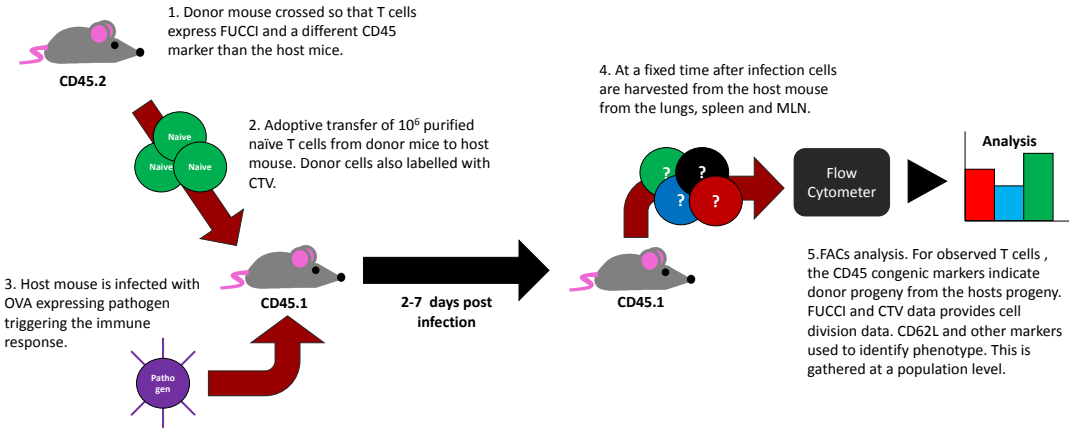


Figure 7: Summary diagram of the experimental setup in [Kinjyo et al., 2015]. Adoptive transfer experiment to observe $CD8^+$ T cell phenotype and division data during the expansion phase.

Experimental Set Up Two: Double adoptive transfer. The authors of [Kinjyo et al., 2015] wished to see if the $CD62L^{int}$ could have both $CD62L^+$ and $CD62L^-$ progeny. The experiment was repeated as described for Experiment One up to the point T cells were harvested. In this new experiment harvesting was only performed on day four post infection. The harvested cells were sorted for **$CD62L^{int}$** (intermediate expression) and for cell trace dye dilution, selecting for cells that were between the fourth and eighth generation. These sorted cells were transferred into secondary recipients (CD45.1), who had been infected at the same time as the primary recipients. These were in turn harvested on day ten after the second transfer and FACS analysis was done on CD62L and FUCCI expression (Figure 8).

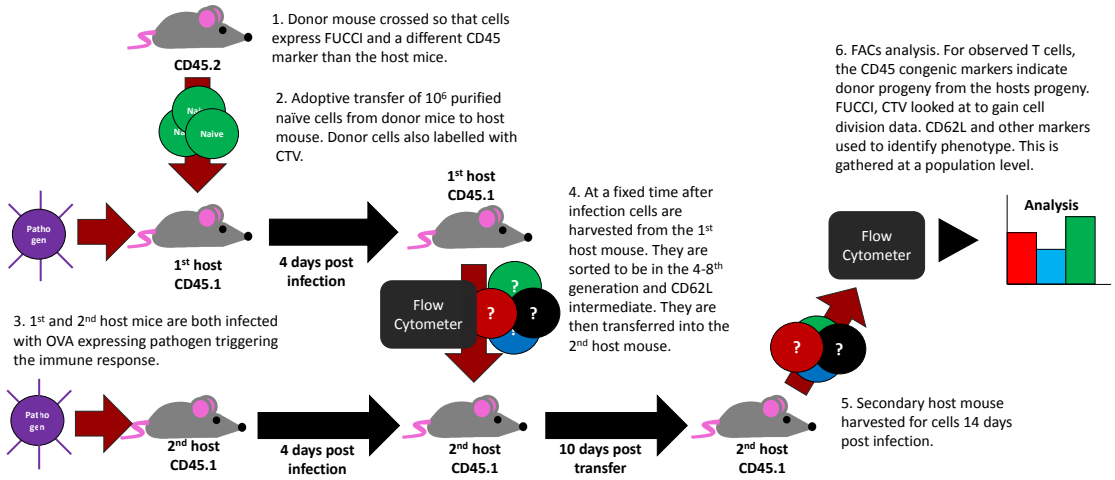


Figure 8: Summary diagram of the experimental setup in [Kinjyo et al., 2015]. A double adoptive transfer experiment to observe the phenotype profile of progeny of activated fourth to eighth generation $CD62L^{int}$ $CD8^+$ T cells.

Experimental Set Three: In vitro cell time lapse microscopy. Previous work had shown $CD8^+$ T cell families had heterogeneous differentiation profiles ([Schlub et al., 2009], [Gerlach et al., 2013], [Buchholz et al., 2013]). To confirm their data were consistent with this, [Kinjyo et al., 2015] sought to show that there was later heterogeneity in cell cycle time, with strong correlations of lifetime within families. In an experiment, they activated $CD8^+$ T cells *in vitro* and allowed them to divide. Cells from the first, third and eight generations were sorted based on their CTV fluorescence. The cells' speed of proliferation was determined using both time lapse imaging and FUCCI expression (Figure 9).

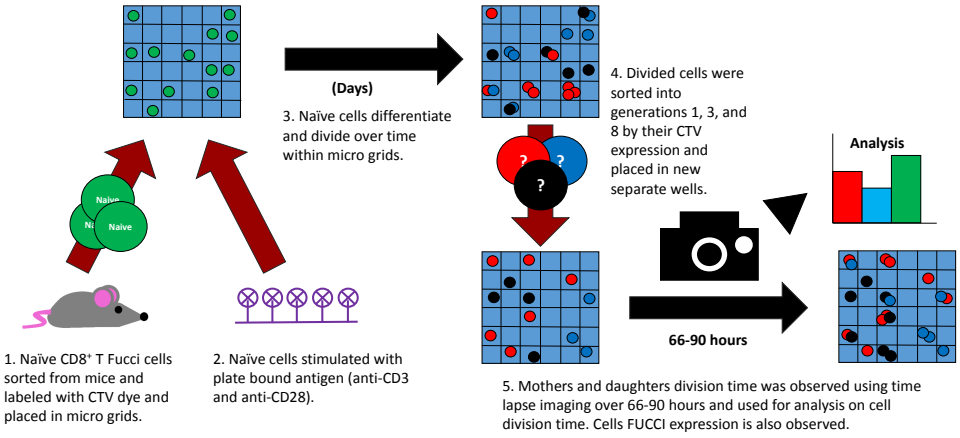


Figure 9: Summary diagram of *in vitro* time lapse experimental setup in [Kinjyo et al., 2015]. An experiment to measure distribution and correlation of $CD8^+$ T cell lifetimes within families.

3.3.2 [Kinjyo et al., 2015] findings

Experiment Set Up One: Single adoptive transfer The article provides data on the percentage of CD62L$^+$ CD8$^+$ T cells at different times post infection (Figure 10). These data show us an initial drop in the proportion of CD62L$^+$ expressing cells during the early expansion phase followed by increase to around 40% of CD62L$^+$ cells by day seven.

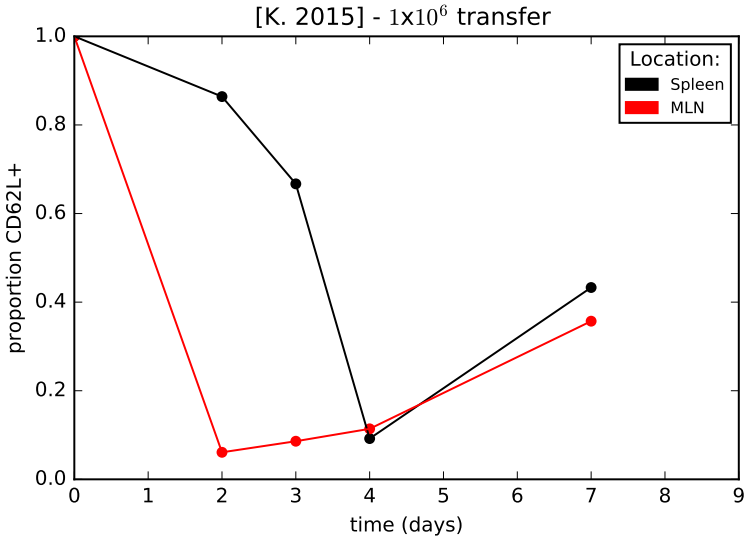


Figure 10: Percentage population of CD8$^+$ cells during the expansion phase that are CD62L$^+$ as described by [Kinjyo et al., 2015]. Values here were taken from data labels on graphs [Kinjyo et al., 2015] from their Figure 3a.

The analysis in [Kinjyo et al., 2015] of these data led them to deduce that

> "By day four post infection, most of the virus-activated cells showed intermediate level of CD62L, as a relatively homogeneous population, compared with the distinct CD62L$^+$ and CD62L$^-$ subsets present on day seven post infection. [Kinjyo et al., 2015, p.5]"

The distinct nature of the latter two populations is also supported by other differences in their phenotypic traits such as size and other markers but CD27, KLRG-1 and IL-7Rα.

[Kinjyo et al., 2015] used FUCCI to give information into cell proliferation speed (see Section 2.4 for information on FUCCI). This combined with CTV information gives understanding into the cells proliferation rates post infection. The article notes the correlation between speed of proliferation and effector traits. Gating on CD62L$^+$ showing slower proliferation (mK02$^+$mAG$^-$), whereas CD62L$^-$ correlated to fast proliferation (mK02$^-$mAG$^+$).

Thus [Kinjyo et al., 2015] concludes upon activation the cell becomes a quickly proliferating homogeneous population on day four at some point giving rise to a memory like sub-population before day

seven.

Experimental Set Up Two: Double adoptive transfer. In Experiment One the complete dilution of cell tracking dye by day seven post infection indicated a large subset of slowly proliferating cells from the start was unlikely. Memory cells must have either been quickly proliferating at some point in the expansion phase or differentiated out of the quickly proliferating $CD62L^{int}$ cells (intermediate expression). The second experiment showed $CD62L^{int}$ cells extracted on day four and adoptively transferred again, gave rise to $CD62L^{+}$ and $CD62L^{-}$ subsets.

Experiment Set Up Three: In vitro cell time lapse The $CD8^{+}$ cells from the first and third generation post infection showed a fast cycle time (averaging 13.4 hours and 14.3 hours). The eighth generation had an average of 20.5 hours division time, and while most cells were still quickly proliferating in this generation there existed a slowly proliferating subset. These results were confirmed by looking at FUCCI expression.

[Kinjyo et al., 2015] investigated correlations in lifetimes of cells from the same family. They found strong correlations of lifetime between siblings (Spearmans rank coefficient, $\rho = 0.91$ to 0.98). They found weaker correlations between mother and daughters (Gen one: $\rho = 0.32$, gen three: $\rho = 0.64$, gen eight: $\rho = 0.66$), which ties in with previous reports on B cell lifetimes ([Markham et al., 2010],[Hawkins et al., 2009]). However, if generation eight was partitioned into a quickly proliferating and slowly proliferating cells, it was found that their progeny's lifetimes were highly likely to be in the same partition as their parents. This was not true for generations one and three where slowly proliferating cells had only a slightly higher correlation in having slow offspring and fast proliferating cells also only had a slightly increased correlation in having quickly proliferating offspring. Thus the data supported the hypothesis of a homogeneous population for generation one and three, which eventually split into two sub-populations, one quickly proliferating and one slowly proliferating, where proliferation speed is inherited.

From the data described as a result of these experiments, [Kinjyo et al., 2015] determines a qualitative model depicted in Figure 11. Upon activation, naive cells differentiate into quickly proliferating homogeneous population of $CD62L^{int}$ cells around two-four days post infection. Before day seven the population became heterogeneous, with cells terminally differentiated into two sub-populations of a central memory phenotype or effector phenotype.

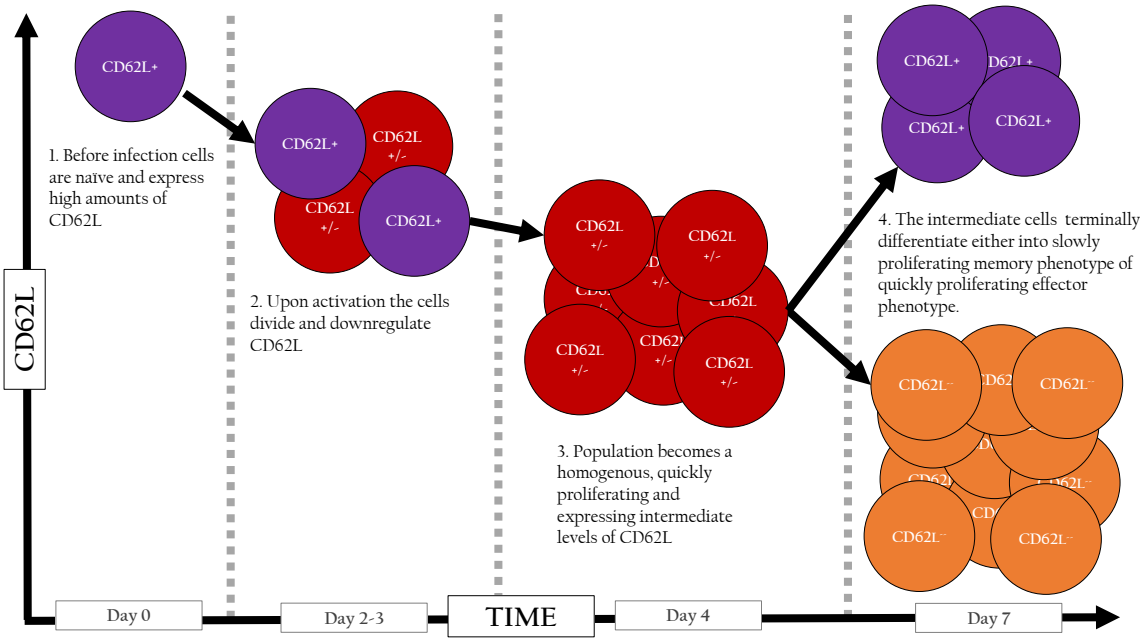


Figure 11: Qualitative model proposed in [Kinjyo et al., 2015]. In this model, after activation, cells become a homogeneous population of quickly proliferating CD62L^{int} effector like cells by day four. This splits into two sub-populations similar to memory and effector phenotypes by day seven.

3.3.3 [Kinjyo et al., 2015] auxiliary experiments and results

Experiments showing that FUCCI expression is a good indicator of cell proliferation speed: [Kinjyo et al., 2015] does a number of experiments to confirm properties of FUCCI. For example, on day seven post infection mice were injected with **BrdU** (Bromodeoxyuridine, a substance cells incorporate into their DNA during DNA synthesis) and spleens where harvested three, five or seven hours later. While some mK02^{+}, mAG^{-} cells incorporated BrdU at every time point, this occurred at a significantly lower rate than its mK02^{-},

Differences in location: The split of the data between location (spleen, lung and MLN) showed many differences T cell population dynamics at different parts of the body. By day four post infection, FUCCI expression indicated cells in the MLN reproduced the quickest, followed by those the lungs and finally the slowest proliferating cells were in the spleen. This was supported by the generational data from CTV. However one must be mindful of the effects of an unknown amount of cell migration might have on these results. Also, similar to other reports ([Kedzierska et al., 2007]), a population of cells with a central memory phenotype (CD44^{+}CD62L^{+}) appeared in the spleen and MLN, but not in the lungs.

Cell life time correlation and environmental factors: In a similar setup to that of Experiment Three, [Kinjyo et al., 2015] found no correlation between unrelated cells that shared a well, supporting

the hypothesis that inheritance of lifetime is driven by internal mechanisms rather than environmental ones.

Cell size: Cell size was observed in many of these experiments and it was found that larger cells both divided faster and had lower expressions of CD62L than their smaller counterparts. This lead to a conclusion that size could be correlated with cell-type.

3.4 Experiment: [Badovinac et al., 2007]

One of the differences in setup between the experiments described in [Buchholz et al., 2013] and [Kinjyo et al., 2015] was the number of T cells adoptively transferred ($1 \times 10^2 - 1 \times 10^4$ and 1×10^6 respectively). The paper [Badovinac et al., 2007] reports data on the effects of transferring different numbers of cells in OT-1 adoptive transfer experiments and allows us to estimate the effect of this difference on T cell behaviour. This paper also contains a similar experimental setup to [Kinjyo et al., 2015] and [Buchholz et al., 2013] (an adoptive transfer experiment where cells are harvested different times post infection at a cohort level) and reports similar population level data. Specifically of interest to us is the proportion of $CD62L^+$ cells at different times post infection.

3.4.1 [Badovinac et al., 2007] experimental setup

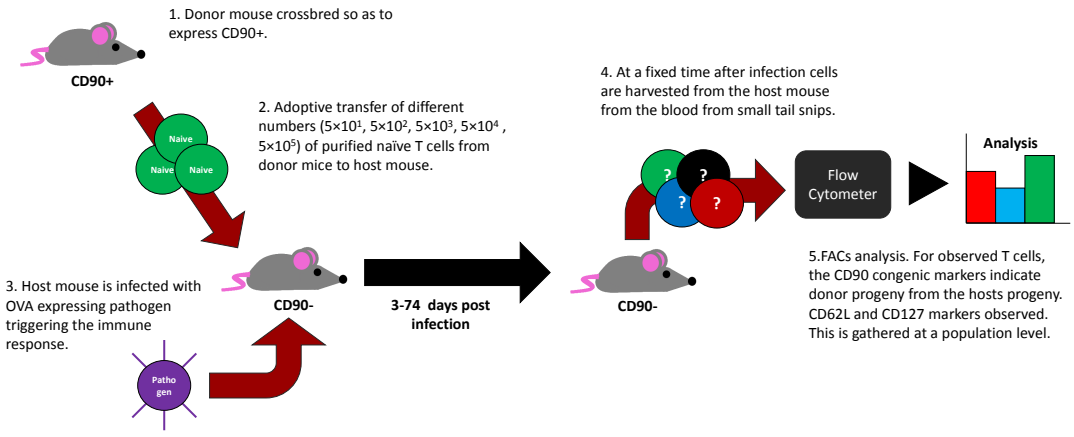


Figure 12: Summary diagram of the experimental setup in [Badovinac et al., 2007]. Adoptive transfer experiment to observe $CD8^+$ T cell phenotype and division data during the expansion phase when varying the number of cells adoptively transferred.

The authors of [Badovinac et al., 2007] describe an adoptive transfer experiment using different numbers of transferred cells: 5×10^1, 5×10^2, 5×10^3, 5×10^4 and 5×10^5. The experiment was repeated with four donor mice per case. Naive cells from $CD90^+$ OT-1 mice were transferred into host $CD90^-$ black-6 mice, and these were infected with 7×10^6 cfu LM-Ova one day after T cell transfer.

Data were collected from blood taken from tail snips done on days three, four, five, six, seven, 10, 13, 26, 40 and 74 post infection. FACS analysis was used to determine CD62L, CD127 and CD90 (Thy-1) expression of harvested cells. CD90 was used to determine which lymphocytes were the progeny of the OT-1 adoptively transferred cells.

3.4.2 [Badovinac et al., 2007] findings

Experiments with a high number of adoptively transferred cells had an earlier peak and shorter expansion phase. For the experiment transferring 5×10^5 cells the peak immune response occurred at day five and while experiments transferring low number of cells (less than 500) had a peak at day seven. When transferring 5×10^5 cells, family size increased only 40 - 400 fold on average up to the peak, compared to the 400,000 fold increase on average for the experiment transferring 5×10^1 cells. We use these results later in Section 5.3.3 to estimate the expansion phase length and average family size at peak immune response for the experiments described in [Buchholz et al., 2013] and [Kinjyo et al., 2015]. Results from [Badovinac et al., 2007] showed proportions of cells expressing high levels of CD127 or CD62L was larger during the memory phase of the immune response for experiments transferring more cells. In all cases, the proportion of $CD62L^+$ cells initially decrease after activation and this proportion eventually increases later in the immune response, but higher transfer number cause this turnaround earlier, as seen in Figure 13. [Badovinac et al., 2007] estimated that excess of 350 cells adoptively transferred will alter the expansion, timescales and phenotype of an immune response.

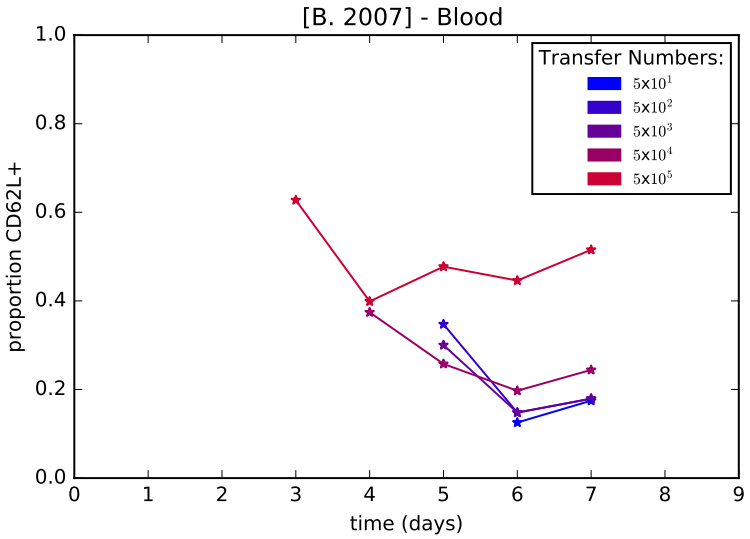


Figure 13: Percentage population of $CD8^+$ T cells during the first eight days that are $CD62L^+$ for different numbers of transferred cells. Values here were taken from graphs in [Badovinac et al., 2007] Figure 4.b

3.4.3 [Badovinac et al., 2007] auxiliary experiments and results

[Badovinac et al., 2007] also investigated how high adoptive transfer number effected results when using a **P14** experimental model as opposed to an OT-1 model. P14 mice have T cells specific to the antigen **LCMV GP33-41**. The P14 TCRs bind to LCMV GP33-41 with less affinity than the OT-1 TCRs binds to SIINFEKL. The results from the experiment described in [Badovinac et al., 2007] showed the P14 experiment had a higher threshold for the number of adoptive transferred cells that would significantly alter the endogenous immune response (estimates of a 350 cell threshold for OT-1 compared to a 6000 threshold for P14).

The authors of [Badovinac et al., 2007] showed that changing the pathogens carrying the antigen, such as virulent LM-ova or vaccine virus expressing **Ova257** did not have a significant effect on the results previously listed for the OT-1 model. This was also the case when using lower doses of the virus. This further supported that the differences in the pathogens used between experiments described in [Buchholz et al., 2013] and [Kinjyo et al., 2015] should have minimal impact.

[Badovinac et al., 2007] reported an experiment that counted the number of OT-1 cells and endogenous T cells in the spleen post infection with different sizes of OT-1 adoptive transfer. This showed that while higher numbers of transferred cells did increase the total count of OT-1 T cells in the spleen at day seven post infection, the increase was not proportional, showing that either cells had died in significant numbers or had not divided as many times. When measuring endogenous T cells at day seven, total counts were lower for very high transfers compared total counts for low transfers, showing high transfers reduce the endogenous immune response size.

3.5 Experiment: [Schlub et al., 2010]

The [Schlub et al., 2010] paper reports data on proportions of $CD8^+$ T cells in the blood that are $CD62L^+$ at different days post infection from an OT-1 adoptive transfer experiment. They repeated this for different numbers of adoptively transferred cells. The purpose of their experiments was to determine the rate at which $CD62L^+$ cells become $CD62L^-$ during the expansion phase. As papers such as [Badovinac et al., 2007] had shown, adoptively transferring very large quantities of cells with pathogen specificity can greatly reduce the number of divisions T cells undergo during the expansion phase. By varying the number of cells transferred, and thus the average number of cell divisions, the authors of [Schlub et al., 2010] looked to find a link between division number and $CD62L^+$ expression. We considered this paper in detail because:

1. It contains a similar experimental setup to one of the experiments in [Kinjyo et al., 2015] and [Buchholz et al., 2013] (an adoptive transfer experiment at cohort level) and reports the proportion of $CD62L^+$ at different times post infection. We later fit an adapted expansion phase models to these data in Section 5.3.5.

2. The additional data on the effects of high numbers of adoptively transferred cells allows us to estimate the impact of this on the [Kinjyo et al., 2015] reported results. We estimate the impact of high transfer numbers in Section 5.3.3.

3. [Schlub et al., 2010] suggests a methodology for estimating the average number of relative divisions cells go through when varying adoptive transfer sizes. In Section 5.3.3 we borrow and extend this method to estimate average family size and we use this for the adapted mathematical model we later employ.

3.5.1 [Schlub et al., 2010] experimental setup

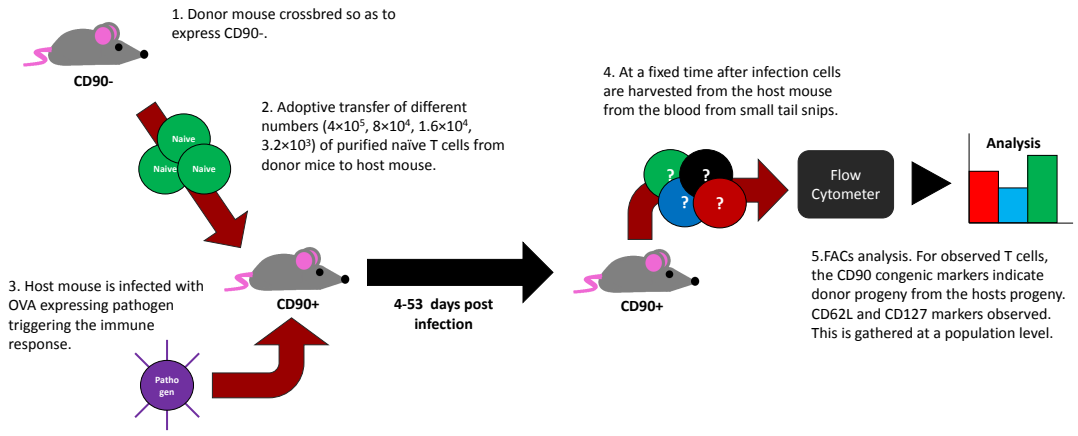


Figure 14: Summary diagram of the experimental setup in [Schlub et al., 2010]. Adoptive transfer experiment to observe $CD8^+$ T cell phenotype and division data during the expansion phase when varying the number of cells adoptively transferred.

The paper [Schlub et al., 2010] describes an adoptive transfer experiment where naive cells from a donor $CD90^-$ OT-1 mouse were transferred into $CD90^+$ black-6 mice. The host mice were infected with 8×10^6 LM-Ova cfu one day after transfer. Mice were bled from tail snips at days four, five, six, seven, eight, 10, 13, 20, 35 and 53 post infection. FACS was used to determine CD62L, CD27 and CD90 (Thy-1) expression of harvested cells (the last of these to determine which cells are the progeny of the OT-1 adoptively transferred cells). This experiment was repeated using different numbers of transferred cells: 3.2×10^3, 1.6×10^4, 8×10^4 and 4×10^5 and repeated with four donor mice per adoptive transfer experiment.

A second similar experiment was performed where cells were harvested from spleen, but only with 3.2×10^3 or 4×10^5 cells adoptively transferred and observations were taken only on fewer days and CD62L expression was not reported. However counts of lymphocytes in the spleen (rather than proportions), split by OT-1 and non OT-1 were reported.

3.5.2 [Schlub et al., 2010] findings

Confirming [Badovinac et al., 2007], these data showed that experiments with high numbers of cells adoptively transferred showed an increase in the proportion of CD62L$^+$ cells before day eight whereas low number of adoptively transferred cells are seen to continually decrease (Figure 15).

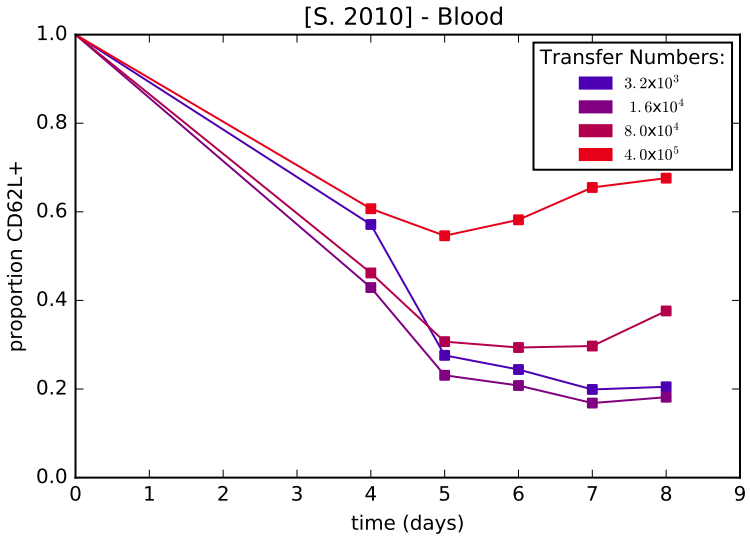


Figure 15: % population of CD8$^+$ cells during the expansion phase that are CD62L$^+$ as described by [Schlub et al., 2010]. Values taken from [Schlub et al., 2010] Figure 1.

While the key objectives of the [Schlub et al., 2010] paper, using a mathematical model to determine the rate CD62L$^+$ cells become CD62L$^-$ is relevant to us, a number of core differences in their assumptions will mean we only touch on their mathematical method briefly. The most significant assumptions [Schlub et al., 2010] make, which we do not, is that CD62L$^+$ and CD62L$^-$ cells divide at approximately the same rate, that naive cells initially divide into CD62L$^+$ cells then and eventually differentiate into CD62L$^-$ cells.

Using this experimental data, [Schlub et al., 2010] employs three methodologies to calculate the relative number of divisions a family goes through and correlates this to the rate at which the proportion of CD62L$^+$ cells decreases during the expansion phase. We describe this method more thoroughly in Appendix A. The counts of T cells in the spleen indicated that the total numbers of lymphocytes (endogenous plus transferred) were similar at peak for both high and low transfers. Fitting an exponential line between the calculated difference in number of divisions and the expression of CD62L give a good fit (r=-0.91, p <0.0001, Spearman's tied-rank correlation) and it is interpreted that each CD62L$^+$ cell had a 21.2% chance to differentiate to a CD62L$^-$ cell upon division.

The data from their experiments gave further evidence of the effects of adoptive transfer of large numbers of OT-1 cells on the immune response. Under the assumptions of the model, in an experiment with 4×10^5 cells adoptively transferred a family will go through approximately 5-6 less divisions than

an experiment transferring 3.2×10^3 cells. The peak of adoptive transfer, was seen to be earlier for higher adoptive transfer numbers, thus while 3.2×10^3 had peak immune response on day six, the adoptive transfer of 4×10^5 cells peaked at day four or earlier. For the adoptive transfer of 3.2×10^3 cells, at peak immune response, approximately 25% of OT-1 cells were $CD62L^+$, whereas for the experiment that transferred 4×10^5 cells, this was around 60% (Figure 15).

3.5.3 [Schlub et al., 2010] auxiliary experiments and results

[Schlub et al., 2010] sought to determine if there was a similar link with division rate and CD127 expression as they did with CD62L, where CD127 is used to discriminate between effector memory and effector cells. However the exponential fit described in method was not as good as in the CD62L case and there are examples of CD127 actually increasing during periods of the expansion phase.
[Schlub et al., 2010] also noted that less than 0.1% of naive cells were unrecruited at the peak of expansion phase, thus ruling out the possibility of unrecruited cells greatly affecting the $CD62L^+$ proportion.

3.6 Comparing experimental systems and results

3.6.1 Comparing all experimental systems: [Badovinac et al., 2007], [Schlub et al., 2010], [Buchholz et al., 2013] and [Kinjyo et al., 2015]

Experiments described in these papers had many similarities that allow for their comparison. All papers describe adoptive transfer of naive $CD8^+$ T cells from donor OT-1 mice to host black-6 mice, and at varying times post infection FACS analysis was done on harvested cells. All experiments used transgenic markers to distinguish between host and donor cells (CD45 or CD90) and all used pathogens expressing OVA to trigger the immune response.

There were also significant differences and these are compared in Table 4. Only [Buchholz et al., 2013] described an adoptive transfer experiment that observed the immune response at a clonal level, at one time point, day eight post infection. That clonal data were what the mathematical model described in [Buchholz et al., 2013] was fitted to. All the papers (along with another experiment from [Buchholz et al., 2013]), described experiments reporting data at cohort level, and these data are observed at multiple time points. This means that if we are to compare papers we have to do so at a population level rather than a clonal one. We see in Section 4.3 that the model [Buchholz et al., 2013] employed also used CD27 expression data and sample error data, some or all of which was not reported in the other experiments.

As mentioned, one of the significant differences between [Buchholz et al., 2013] and [Kinjyo et al., 2015] was the number of cells adoptively transferred, shown in Table 4. [Kinjyo et al., 2015] and [Buchholz et al., 2013] both gathered their data from the spleen for their population level experiments.

[Badovinac et al., 2007] and [Schlub et al., 2010] report data taken from blood tail snips. Blood data were taken from the same mice at different times whereas spleen measurements require mice to be sacrificed at each time point. Both [Badovinac et al., 2007] and [Schlub et al., 2010] report some spleen data from other experiment but this is too limited for our modelling purposes. The papers [Buchholz et al., 2013] and [Schlub et al., 2010] did experiments to check that samples of T cells taken from specific parts of the body are representative of the whole, however [Kinjyo et al., 2015] reported considerable differences in T cell division speed and generational data between locations. Therefore difference in location between papers must always be kept in mind when comparing results.

The papers [Buchholz et al., 2013] and [Kinjyo et al., 2015] focused on finding the order of differentiation within the expansion phase. [Kinjyo et al., 2015] relies on FUCCI expression indicating speed of cell differentiation and CD62L as the primary indicator of cell-type and then used correlations of other characteristics such as cell size to further support their classification. [Buchholz et al., 2013] used CD62L and CD27 as the primary indicators of cell-type with additional factors such as T-bet to support their classification.

[Kinjyo et al., 2015] used the pathogen PR8-Ova to induce an immune response, whereas the other experiments used LM-Ova. Different quantities of pathogens were used between experiments as seen in Table 4. However, both [Buchholz et al., 2013] and [Badovinac et al., 2007] gave evidence that the different pathogen on the OT-1 model had little effect on results. The gating strategies used to phenotypically define cell-types were different in each paper (Table 4) and this should be kept in mind when comparing cell surface marker expression data between experiments.

These factors will need to be considered, as differing conclusions could have an experimental explanation, such as the treatment of location or phenotype profiling, rather than analytical ones.

Setup	**[Badovinac et al., 2007]**	**[Schlub et al., 2010]**	**[Buchholz et al., 2013] clonal experiment**	**[Buchholz et al., 2013] cohort experiment**	**[Kinjyo et al., 2015]**
Factors in common	Adoptive transfer experiment where OT-1 cells are transferred to black-6 (C57BL/6) mice. Cells are harvested at different times post infection and the proportion of $CD62L^+$ is recorded.				
Naive cells transferred	5×10^1 5×10^2 5×10^3 5×10^4 5×10^5	3.2×10^3 1.6×10^4 8×10^4 4×10^5	107	1×10^2 (6,8 days p.i.) 1×10^4 (1-4 days p.i.)	1×10^6
Location cells harvested from (as reported)	blood	blood	lungs + Lymph nodes (multiple) + spleen (split not reported)	spleen	spleen, MLN (split reported)
Times post infection cells harvested (days)	3, 4, 5, 6, 7, 10, 13, 26, 40, 74 (earlier days missing for larger transfers)	4, 5, 6, 7, 8, 10, 13, 20, 35, 53	8	1, 2, 3, 4, 6, 8	0, 2, 3, 4, 7
Pathogen application	LM-Ova bacteria 7×10^6	LM-Ova bacteria 8×10^6 cfu	LM-Ova bacteria 5×10^3 cfu*	LM-Ova bacteria 5×10^3 cfu*	PR8-Ova virus 1×10^2 pfu
Host identifiers	CD90.1.1	CD90.1.2*	CD90.2.2CD45.2.2	Data not reported	CD45.1
Donor identifiers	CD90.1.2	CD90.1.1	CD90.1.1CD45.1.1 CD90.1.2CD45.1.1 CD90.2.2CD45.1.1 CD90.1.1CD45.1.2 CD90.1.2CD45.1.2 CD90.2.2CD45.1.2 CD90.1.1CD45.2.2 CD90.1.2CD45.2.2	Data not reported	CD45.2
Number of host mice per experiment	4	4	47	47*	9
Effector memory vs effector identifier	CD127	CD127	CD27	CD27	Data not reported
CD62L gating threshold (approximate, read from diagrams)	50*	20	100-150	100-150	2000
Clonal/cohort level data	Cohort	Cohort	Clonal	Cohort	Cohort
Sample error data (or equivalent)	No	Yes	Yes	Yes	No

Table 4: Summary comparison of experimental setups and data reported. As papers include multiple experiments and this table only refers to that which returns proportion of CD62L cells at different times. The * indicates some inference from the reader, usually from a description of another experiment in the paper that is assumed to be carried over between experiments.

3.6.2 Comparing results: [Badovinac et al., 2007], [Schlub et al., 2010], [Buchholz et al., 2013] and [Kinjyo et al., 2015]

All papers have experiments that provide data on the proportion of $CD62L^+$ cells over time up to day eight post infection. This is compared in Figure 16.

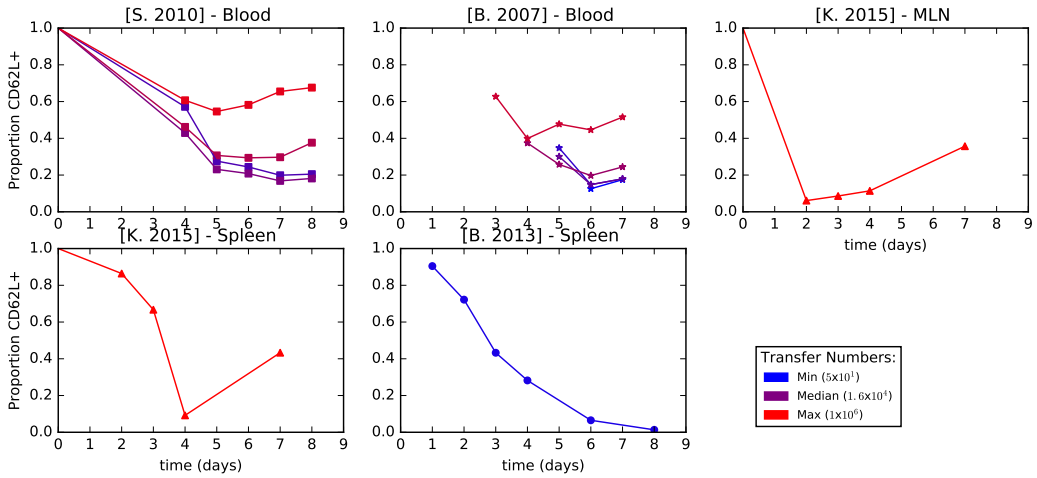


Figure 16: Proportion of CD62L$^+$ OT-1 cells after infection in adoptive transfer experiments reported in papers [Buchholz et al., 2013], [Kinjyo et al., 2015], [Badovinac et al., 2007] and [Schlub et al., 2010]. Colours represent ranking of number of cells adoptively transferred: high number indicated by a red hue and low numbers by a blue hue. All experiments showed an early decrease in the proportion of CD62L$^+$ cells. Experiments with higher numbers of cells adoptively transferred tended to show an increase in the proportion of CD62L$^+$ cells earlier in the immune response. All experiments used different gating thresholds. Data taken from: [Badovinac et al., 2007] graph 4b [Schlub et al., 2010] graph 1d, [Buchholz et al., 2013] graph in the supplementary material S26, [Kinjyo et al., 2015] data labels in Figure 3a. Some data sets did a prior infection observation ([Schlub et al., 2010] and [Kinjyo et al., 2015]) and this point was not fitted to.

All data up to day two showed an initial large CD62L$^+$ population. However, [Kinjyo et al., 2015] associates this with cells still being naive whereas [Buchholz et al., 2013] associates this with an early memory population. All graphs show a decrease in this initial CD62L$^+$ population, which both [Kinjyo et al., 2015] and [Buchholz et al., 2013] put down to the effector pool increasing in size (TEMp and TEFs in [Buchholz et al., 2013]'s model). However, [Kinjyo et al., 2015] reported an increase in the percentage of CD62L$^+$ towards the day seven, similar to that seen in other papers, which also had high numbers of adoptively transferred cells ([Bachmann et al., 2005]), and this was explained by cells differentiating into memory cells, whereas [Buchholz et al., 2013] continues to show a decrease.

As [Badovinac et al., 2007] and [Schlub et al., 2010] have shown, large adoptive transfers can lead to and earlier increase in CD62L$^+$ such as that seen in [Kinjyo et al., 2015]. This leads to the hypothesis that some of data reported in [Kinjyo et al., 2015] may not be in the expansion phase, despite the fact the expansion phase traditionally considered to extend to day seven-eight post infection. Noticeably, while we have seen that high adoptive transfer usually causes high levels of CD62L$^+$, expression reported in [Kinjyo et al., 2015] dips below that of [Buchholz et al., 2013] on day four, though differing gating strategies could perhaps explain this. Also, mesenteric lymph node (MLN) data from [Kinjyo

et al., 2015] exhibited an increase in proportion of CD62L$^+$ cells even earlier, from day two to three (Figure 10), a slightly different pattern from the spleen, which might still indicate some early creation of memory, but the absence of MLN CD62L$^+$ data from the papers we compare makes it is difficult to evaluate this further.

Another area where both [Kinjyo et al., 2015] and [Buchholz et al., 2013] have comparable data is the generation profile of cells at days three and four post infection as seen in Figure 17.

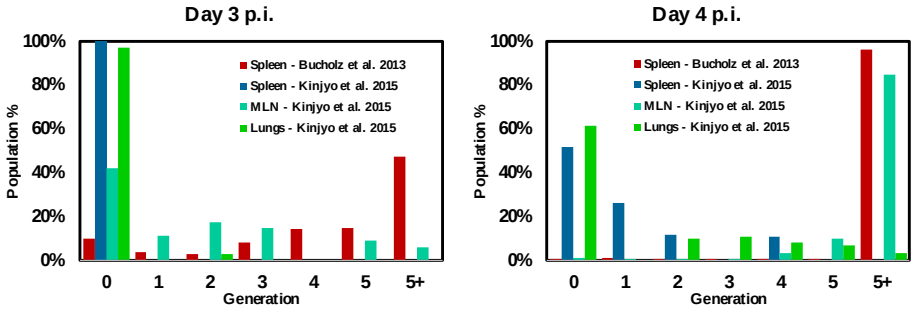


Figure 17: Graphs showing OT-1 cells generation profile at a population level in [Buchholz et al., 2013] and [Kinjyo et al., 2015] experiments. [Buchholz et al., 2013] has almost entirely diluted dye by day four. In addition to the data shown, all dye was diluted in [Kinjyo et al., 2015] at all locations by day seven (no data from day seven was reported [Buchholz et al., 2013]). This could be considered evidence against a slowly proliferating subset in the population. The overall slower proliferation of [Kinjyo et al., 2015] compared to [Buchholz et al., 2013] could be due to the size of adoptive transfer which has been shown to reduce the number of divisions cells go through. Values for [Kinjyo et al., 2015] were manually extracted from CTV flow plots, their Figure 1.b and [Buchholz et al., 2013] data were taken from graph points in their supplementary material S22.

The T cells [Buchholz et al., 2013] reported, taken from the spleen, show a faster rate of proliferation than those reported in [Kinjyo et al., 2015] taken from any location. In addition, [Kinjyo et al., 2015] also observed complete cell dye dilution at all locations on day seven, a result which [Buchholz et al., 2013] saw by day four (using CFSE vs CTV). [Kinjyo et al., 2015] uses this fact to conclude that cells cannot be of a slowly proliferating lineage throughout the expansion phase, a requirement for the [Buchholz et al., 2013] model. One possible explanation of the faster proliferation of [Kinjyo et al., 2015]'s data compared to [Buchholz et al., 2013] may be high adoptive transfer size, as we have seen this leads to smaller family size.

The complete CTV dilution seen in [Kinjyo et al., 2015] on day seven suggests the immune response is still in the expansion phase for a significant time after day four, despite evidence from [Badovinac et al., 2007] and [Schlub et al., 2010] that the expansion phase for such a high adoptive transfer should end around this time. Thus, while day seven CD6L$^+$ profile in [Kinjyo et al., 2015] matched that of a high adoptive transfer, the generational data on day seven was seemingly contradictory to it.

As [Schlub et al., 2010] and [Badovinac et al., 2007] do not reported generational data it is hard to compare, but one possible explanation is that cell death could be influencing the appearance of the generational profile of [Kinjyo et al., 2015] by day seven.

3.6.3 Comparing models: [Buchholz et al., 2013] and [Kinjyo et al., 2015]

Due to their novel experimental set-up [Buchholz et al., 2013] is able to propose how cells expand and differentiate at a clonal level, which can be translated into a population level. The authors of [Kinjyo et al., 2015] propose a model for cell differentiation behaviour only at a population level with some unknowns about the clonal level, most notably, the heterogeneous population at day seven described by [Kinjyo et al., 2015] could be made up of homogeneous or heterogeneous individual families.

As [Buchholz et al., 2013] used mathematical model fitting, their model is also quantitative and fully descriptive, for example giving rates of division, whereas [Kinjyo et al., 2015] is qualitative. However, the [Buchholz et al., 2013] relies on strict assumptions, such as independence of cell behaviours (these model assumptions are reviewed in Chapter 6).

Both models propose very different differentiation dynamics for the expansion phase. Most notably, while both require a down regulation of CD62L during initial stages of the expansion phase, [Buchholz et al., 2013]'s model necessitates this trend continues whereas [Kinjyo et al., 2015] allows for it to increase near day seven. Both these models tie with their respective observed proportions of $CD62L^+$ cells against time.

In terms of population growth dynamics, both models are different, with [Kinjyo et al., 2015] suggesting that all cells and families enter a stage of fast proliferation after activation, with a subgroup slowing down to a memory cell-type by day seven. [Buchholz et al., 2013] would suggest all cells having to go through slower proliferation during the early expansion phase, and some cells would speed up division only later once they have proliferated and also suggests that memory cells will have never entered a period of fast average proliferation.

Using an adapted version of [Buchholz et al., 2013]'s method, we compare the models fitted to data in these experiments, looking specifically at CD62L. We do this in Section 5.2.3. We also review assumptions and alternative models in Chapter 6.

4 Mathematical models used to interrogate the CD8$^+$ T cell expansion phase

This chapter describes pre-existing mathematical and statistical results that are used to help model and explain the CD8$^+$ T cell expansion phase. Firstly, we review some of the tools of mathematical modelling. Secondly, we describe the family of models that the authors of [Buchholz et al., 2013] selected from to find best fit models, known as multi–type Bellman-Harris processes. Finally, we shall describe how these tools are used by [Buchholz et al., 2013] specifically. We go on later in the thesis to adapt these tools to fit models to other data such as that of [Kinjyo et al., 2015].

4.1 Using summary statistics to find the best fitting mathematical models

4.1.1 Motivation

While the mechanics that control the immune system are not fully understood, their results are partly visible in experiments such as those described in Chapter 3. If we want to approximate these dynamics with a mathematical model, we inform the model design by basing the assumptions on deductions from biological observations; for example, the fact that cells always divide into no more than two cells. The way T cell families divide and differentiate during the immune response has been observed to be highly heterogeneous ([Gerlach et al., 2013], [Buchholz et al., 2013]); therefore, a deterministic model at a clonal level would be inappropriate. Some elements of the expansion phase can be modelled as if taking samples from a random variables whose distributions are unknown, but are informed from biological experiments.

Once a list of reasonable assumptions for a mathematical model has been selected, this may not have narrowed down the possibilities to one model or a given set of parameters. In order to rank those models that remain, we seek to compare the observed statistics from experiments with the statistics calculated from the mathematical models. We may also want to penalise more complex models based on Occam's razor.

4.1.2 Ranking models

Suppose we have m replicates of an experiment, and each experiment takes n real valued measurements, then we denote the total mn observations as a matrix $O_{m\times n} \in \mathbb{R}^{m\times n}$ and call this the **dataset**. We also define $o_{i,j} \in \mathbb{R}$ as the j^{th} measurement in replicate i. We denote the n measurements seen in replicate i as a vector $\vec{o_i} = (o_{i,1}, ..., o_{i,n})$.

In the case of [Buchholz et al., 2013], there were 47 transferred T cell families recovered ($m = 47$) from an adoptive transfer experiment on day eight post infection. For each of these experiments, the

cells count of three phenotypes was recorded ($CD62L^+CD27^+$, $CD62L^-CD27^+$ and $CD62L^-CD27^-$, $n = 3$). Each clonal family from within the same mouse is assumed to be an independent replicate of an experiment.

Suppose we have a function $\vec{s} : \mathbb{R}^{m\times n} \to \mathbb{R}^k$ defined as $\vec{s}(A_{m\times n}) = (s_1(A_{m\times n}), ..., s_k(A_{m\times n}))$ where $s_i : \mathbb{R}^{m\times n} \to \mathbb{R}$ for $i = 1, ..., k$. Thus, each of the k functions $s_i(O_{m\times n})$ returns a statistic of the data set $O_{m\times n}$ and $\vec{s}(O_{m\times n})$ is a vector of statistics which we call the **observed statistics**. For brevity we shall denote the vector of observed statistics $\vec{s}(O_{m\times n})$ as $\hat{x} = (\hat{x}_1, ...\hat{x}_k)$.

In the case of [Buchholz et al., 2013] the authors calculated nine summary statistics ($k = 9$) on their data set $O_{47\times 3} \in \mathbb{R}^{47\times 3}$. This was made up of three means

$$\hat{x}_j = s_j(O_{47\times 3}) = \frac{1}{47}\sum_{i=1}^{47} o_{i,j} \text{ for } j = 1, 2, 3,$$

three coefficients of variation

$$\hat{x}_{j+3} = s_{j+3}(O_{47\times 3}) = \frac{\sqrt{\frac{1}{47}\sum_{i=1}^{47}(o_{i,j})^2 - \hat{x}_j^2}}{\hat{x}_j} \text{ for } j = 1, 2, 3$$

and three correlation coefficients

$$\hat{x}_7 = s_7(O_{47\times 3}) = \frac{\frac{1}{47}\sum_{i=1}^{47} o_{i,1}o_{i,2} - \hat{x}_1\hat{x}_2}{\hat{x}_4\hat{x}_5\hat{x}_1\hat{x}_2}$$
$$\hat{x}_8 = s_8(O_{47\times 3}) = \frac{\frac{1}{47}\sum_{i=1}^{47} o_{i,1}o_{i,3} - \hat{x}_1\hat{x}_3}{\hat{x}_4\hat{x}_6\hat{x}_1\hat{x}_3}$$
$$\hat{x}_9 = s_9(O_{47\times 3}) = \frac{\frac{1}{47}\sum_{i=1}^{47} o_{i,2}o_{i,3} - \hat{x}_2\hat{x}_3}{\hat{x}_5\hat{x}_6\hat{x}_2\hat{x}_3}.$$

Let us also suppose we have a probabilistic model that we think approximates the experiment and we calculate the expected value of each of the k summary statistics according this model. We denote each calculated statistic from the model as $x_i \in \mathbb{R}$ for $(1 \leq i \leq n)$, with $\vec{x} = (x_i, ..., x_k)$, and call this vector the **model statistics**. The authors of [Buchholz et al., 2013] used a family of stochastic processes known as multi-type Bellman-Harris processes to model the data, which we detail later. From these, they calculated model statistics.

We wish to define a function that measures the distance between $\hat{x}$ and x to see how well model statistics explain observed statistics. This is done with an **objective function** denoted by $f(\hat{x}, x)$, $f : \mathbb{R}^{2k} \to \mathbb{R}$.

There are many candidates for what the objective function could be and selection should be informed by what we think is important in terms of fit, for example, non–negativity and the function should be zero if and only if $x = \hat{x}$. One option is the **mean squared error** (**MSE**) ([Bertsekas and Tsitsiklis, 2002, Ch. 8.3]) defined by

$$\text{MSE}(\hat{x}, \vec{x}) = \frac{1}{n}\sum_{i=1}^{k}(\hat{x}_i - x_i)^2. \tag{1}$$

The squaring of the differences means larger errors are penalised more harshly. There are some cases when this is an inappropriate choice of objective function, such as if statistics have a wide range of magnitudes and variances. This would mean statistics of larger magnitude or variance will have a disproportional effect on the MSE than smaller ones.

To account for this, we may wish to use a similar function that adjusts for the **standard error of the mean** (**SEM**) of the statistics and we denote the sample variance as $\hat{\sigma}_i^2 \in \mathbb{R}$ of each summary statistic $\hat{x_i}$. The sample variance of the observed statistics may be unknown *a priori*, but can be approximated by **bootstrapping** ([Press et al., 1992, Ch. 15.6]). Bootstrapping allows us to estimate the underlying distribution the dataset is selected from by creating new datasets $O'_{m\times n} \in \mathbb{R}^{m\times n}$ by random sampling with replacement from the original data set. Experiments must be i.i.d. (independent and identically distributed) in order for this technique to be appropriate. To allow us to randomly sample from these data we create a series of i.i.d. discrete random variables that take uniform distributions over integers between one and the number of experiments (m) and we denote these random variables Y_i, with $i = 1, 2, 3, ...$ Then $\vec{o}_{Y(i)}$ is a random sample of one experiment from $O_{m\times n}$. For $j = 1, 2, 3, ...$ we define a new dataset $O^j_{m\times n} \in \mathbb{R}^{m\times n}$ as a dataset created by randomly sampling from the original dataset, $O^j_{m\times n} = (\vec{o}_{Y(mj+1)}, ..., \vec{o}_{Y(mj+m)})$. We want a large number of such samples which we denote l. Then the value of σ_i^2 is approximated by

$$\hat{\sigma}_i^2 \approx \frac{1}{l}\sum_{j=1}^{l} s_i(O^j_{m\times n})^2 - \frac{1}{l^2}(\sum_{j=1}^{l} s_i(O^j_{m\times n}))^2.$$

The bootstrapping method is used in [Buchholz et al., 2013] to approximate the value of the sample variance.

A similar measure to the mean squared error is **least squares** or **standard** χ^2 ([Press et al., 1992, Ch. 15.1]) which is calculated as

$$\chi(\hat{x}, \vec{x}) = \sum_{i=1}^{k} \frac{(\hat{x_i} - x_i)^2}{\hat{\sigma}_i^2}. \tag{2}$$

In this case different magnitudes of standard deviation are adjusted for. However, this does not adjust for any covariance between the summary statistics. The statistics used to fit in [Buchholz et al., 2013] had sample variances of differing magnitudes, hence the authors used χ^2 minimization (2) to measure the distance between their observed statistics and their model statistics.

Let us suppose we have a probabilistic model that we suspect captures the underling dynamics that created the data set and this model is parametrised by q parameters ($\vec{\theta} \in \mathbb{R}^q$) and the set of all possible parameters is $\Theta \subset \mathbb{R}^q$. For each $\vec{\theta}$, we denote this model's respective vector of calculated statistics as $\vec{x}(\vec{\theta}) \in \mathbb{R}^k$. For a given objective function f and a given set of observed statistics $O_{m\times n}$, we can use an objective function such as MSE (1) or χ^2 (2) to identify the parameters of the model that give an infimum distance between the observed statistics and the model statistics. The values of $\vec{\theta}$, should

they exist, that return this minimum value objective function, are denoted as

$$\{\vec{\theta}^*\} = \arg\inf_{\vec{\theta}\in\vec{\Theta}} f(\hat{x}, \vec{x}(\vec{\theta})). \tag{3}$$

and this set we call the **best fit parameterisation(s).** Thus, we have a method of choosing a $\vec{\theta}* \in \Theta$ that best approximates the data according to an objective function.

Unless the function being optimised is convex, solving optimisation problems can be a challenge and this is a rich field of research in itself. There are many computer algorithms designed to find a minimum value of a function. These minimising programmes require equivalents of the following inputs: the objective function redefined $f(\hat{x}, \vec{x}(\vec{\theta})) : \mathbb{R}^q \to \mathbb{R}$ where $\vec{\theta}$ is now the variable and $\hat{x}$ is fixed, a starting point $\vec{\theta} \in \vec{\Theta}$, bounds for the variable defining $\vec{\Theta} \subset \mathbb{R}^q$ (optional), and a selected algorithm, for example the authors of [Buchholz et al., 2013] used a trust-region-reflective algorithm (Conn et al. [1987]). There is no guarantee that this function will return the global minimum and not a local one. This is mitigated by the use of multiple start points for $\vec{\theta_0}$ ([Buchholz et al., 2013] used 300 different start points) or using a range of algorithms and choosing the minimum of these. Thus, we have a method for acquiring the best fitting parameters as described in (3) for a given model, objective function and observations.

Supposing we have r different mathematical models labelled i (with $i \in \{1, ..., r\}$) whose parameters take values in their respective spaces $\vec{\Theta}_i$ that may have a different number of dimensions q. One method to determine which model best describes the data is to find the minimum of the objective function $f(\hat{x}, \vec{x}(\theta))$ over $\vec{\theta} \in \vec{\Theta}_i$ for each model and then finding the lowest value of the objective function across all r models. However, it may be the case that we wish to penalize models for higher complexity as this carries the risk of over fitting. It may also be that some sets of models may be a subset of one another. This is known as one set of models being **nested** in another.

One possible method could be to rank the models by using the χ^2 minimization value divided by the **degrees of freedom (d.o.f.)** ([Press et al., 1992, Ch. 15.1]). A degree of freedom is the number of independent scores (observations that make up this estimate), m, minus the number of parameters q. The values of (2) for the different models are ranked only after being divided by their respective d.o.f.

An alternative method would be to use **Akaike information criterion (AIC)** ([Akaike, 1974]). For each model, where q is the number of model parameters that are fitted and L is the maximum value of the likelihood function, i.e. for a given model i, $L(\vec{x}|\vec{\theta})$ is the maximized value of $P(\vec{x}|\vec{\theta})$ for all $\vec{\theta} \in \vec{\Theta}$, the AIC value is given as

$$\text{AIC} = 2q - 2ln(L(\vec{x}|\vec{\theta})). \tag{4}$$

There is also a **corrected Akaike information criterion (AICc)** ([Akaike, 1974]) that takes into

account the size of a finite sample (m) and this is calculated as

$$\text{AICc} = 2q - 2ln(L) + \frac{2q(q+1)}{m-q-1}. \tag{5}$$

Models that return lower values from χ^2 minimization corrected by the degrees of freedom or AICc indicate a better fit and thus models can be ranked. It is important to note that this is a relative method of comparing models rather than an absolute one and does not compare against a null hypothesis.

The authors of [Buchholz et al., 2013] used both AICc on the likelihood function and χ^2 value corrected by the degrees of freedom to rank suggested models of the $CD8^+$ T cell expansion phase. We recreate this and adapt it to new data in Chapter 5.

4.2 Bellman-Harris processes

4.2.1 Description of a Bellman-Harris process

A Bellman-Harris process is a probabilistic model describing how populations of objects reproduce, go extinct or expand over time. Applications include populations of animals reproducing, particles in a nuclear cascade, names in a family tree or cells dividing.

We first give a high level informal description of a Bellman-Harris process and then define it precisely and mathematically. Let us say we have one object alive at time zero, and we assign this object a random lifetime. At the end of the object's lifetime, it has a random number of children, which could be any integer including zero. These new objects are then also assigned random lifetimes and they also have a random number of children upon their death. This process continues, giving rise to a family tree, otherwise known as an age dependent branching process. A core assumption of the Bellman-Harris process is that the lifetime of each of the objects are i.i.d. Throughout this section we say these lifetime random variables are identical to a random variable J. Also, the distribution of the number of offspring that each object has upon death is i.i.d.. We shall call a random variable identical to these offspring random variables as H. In addition, all the distributions of offspring are independent from the lifetime distributions. The term Bellman-Harris process also includes processes with more than one object alive at time zero, but we only look at cases starting with one ancestor object and these results can be easily expanded to cases starting from multiple ancestor objects due to the properties of independence. For a visual realisation of a Bellman-Harris process see Figure 18.

We outline the formal definition described in Bellman-Harris [Harris, 1964] Chapter 6. We firstly define a set of unique identifiers for objects in the branching process. Suppose $\langle x \rangle = \langle x_1, x_2, ..., x_k \rangle$ is a finite collection of ordered integers with $x_1, x_2, ..., x_k \in \mathbb{N}_{>0}$ and $k \in \mathbb{N}_{\geq 0}$. Let us define $\mathbb{X}$ as the collection of all possible $\langle x \rangle$s. Note $\mathbb{X}$ also includes the empty set. Let us denote the ancestor object as $\langle \emptyset \rangle$. Let us denote the ancestor's first child as $\langle 1 \rangle$, its second child as $\langle 2 \rangle$ and so on. In general, indexed object $\langle x_1, ..., x_{k-1} \rangle$ has its $x_k{}^{\text{th}}$ child denoted as $\langle x_1, ..., x_{k-1}, x_k \rangle$. Because all children appear at the

same time, there is no significance of first, second or third child etcetera; these numbers merely act as unique identifiers.

To fully define a specific family history of a branching process, we need to assign lifetimes and numbers of offspring to each object. For each object $\langle x \rangle \in \mathbb{X}$ let us assign it two values: its lifetime $j\langle x \rangle \in \mathbb{R}_{\geq 0}$ and the number of offspring it has upon death $h\langle x \rangle \in \mathbb{N}_{\geq 0}$. We define a family history ω as the collection of all these values for $j\langle x \rangle$ and $h\langle x \rangle$ for each $\langle x \rangle \in \mathbb{X}$, $w = (j\langle \emptyset \rangle, h\langle \emptyset \rangle, j\langle 1 \rangle, h\langle 1 \rangle, j\langle 1,1 \rangle, h\langle 1,1 \rangle, j\langle 2 \rangle, h\langle 2 \rangle, ...)$. Let us denote the complete collection of all family histories as Ω. We note that if, for example, $h\langle x_1, ..., x_l \rangle = k$ then child $\langle x_1, ..., x_l, k+1 \rangle$ is never born and so $j\langle x_1, ..., x_l, k+1 \rangle$ and $h\langle x_1, ..., x_l, k+1 \rangle$ are redundant information contained in ω.

We want to introduce a probability measure P on the space of all family histories Ω. We do this by defining a probability measure on each of its elements. Let us introduce a set of random variable $J_{\langle x \rangle} : \Omega \to \mathbb{R}_{\geq 0}$ for each $\langle x \rangle \in \mathbb{X}$ such that $J_{\langle x \rangle}(\omega) = j\langle x \rangle$. Let us say the set of random variable $J_{\langle x \rangle}$ are i.i.d. and identically distributed to a random variable J having some distribution $P(\omega \in \Omega : J(\omega) \leq t) = P(J \leq t) = F_J(t)$. We assume $F_J(t)$ to be any probability distribution with the properties $F_J(0-) = 0$ and $F_J(0+) < 1$. Likewise, let us introduce a random variable $H_{\langle x \rangle} : \Omega \to \mathbb{N}_{\geq 0}$ for each $\langle x \rangle \in \mathbb{X}$ such that $H_{\langle x \rangle}(\omega) = h\langle x \rangle$. We say all $H_{\langle x \rangle}$s are identically distributed to a random variable H with the distribution $P(\omega \in \Omega : H(\omega) \leq h) = P(H \leq h) = F_H(h)$. The random variables $H_{\langle x \rangle}$ are i.i.d. and independent of the random variables $J_{\langle x' \rangle}$s for all $\langle x \rangle, \langle x' \rangle \in \mathbb{X}$.

Each ω in Ω now corresponds to a denumerable family of independent random variables. [Kolmogorov, 1933, Ch. 4] gives us that a denumerable set of independent, real valued random variables determine a unique countable additive probability measure P on measurable sets in Ω.

To define a random variable for the population at time t under this probability measure, let us first define a function that designates if an object is alive. The function $a(\langle x_1, ..., x_k \rangle, \omega, t) : \mathbb{X} \times \Omega \times \mathbb{R}_{\geq 0} \to \mathbb{Z}$ is defined as

$$a(\langle x_1, ..., x_k \rangle, \omega, t) = \begin{cases} 1, & \text{if } x_1 \leq h\langle \emptyset \rangle \text{ and } x_2 \leq h\langle x_1 \rangle \text{ and } x_3 \leq h\langle x_1, x_2 \rangle ... \text{ and } x_k \leq h\langle x_1, ..., x_{k-1} \rangle \\ & \quad \text{and } j\langle \emptyset \rangle + j\langle x_1 \rangle + ... + j\langle x_1, ..., x_{k-1} \rangle \leq t \\ & \quad \text{and } j\langle \emptyset \rangle + j\langle x_1 \rangle + ... + j\langle x_1, ..., x_{k-1} \rangle + j\langle x_1, ..., x_k \rangle > t. \\ 0 & \text{otherwise} \end{cases}$$

This means $a(\langle x \rangle, \omega, t) = 1$ if object $\langle x \rangle$ is alive at time t for a family history ω. Let us define a function $Z(\omega, t) : \Omega \times \mathbb{R}_{\geq 0} \to \mathbb{Z}$ as

$$Z(\omega, t) = \sum_{\langle x \rangle \in \mathbb{X}} a(\langle x \rangle, \omega, t).$$

$Z(\omega, t)$ is the count of all objects alive at time t for family history ω. As $Z(\omega, t)$ is a measurable function on Ω ([Harris, 1964, Ch. 6.4]), we define $Z(t) : \Omega \to \mathbb{N}_{\geq 0}$ using the probability measure defined by P over Ω, so that $Z(t)$ as a random variable, representing the number of objects alive at time t, with the distribution $P(\omega \in \Omega : Z(t, \omega) \leq z) = P(Z(t) \leq z) = F_{Z(t)}(z)$.

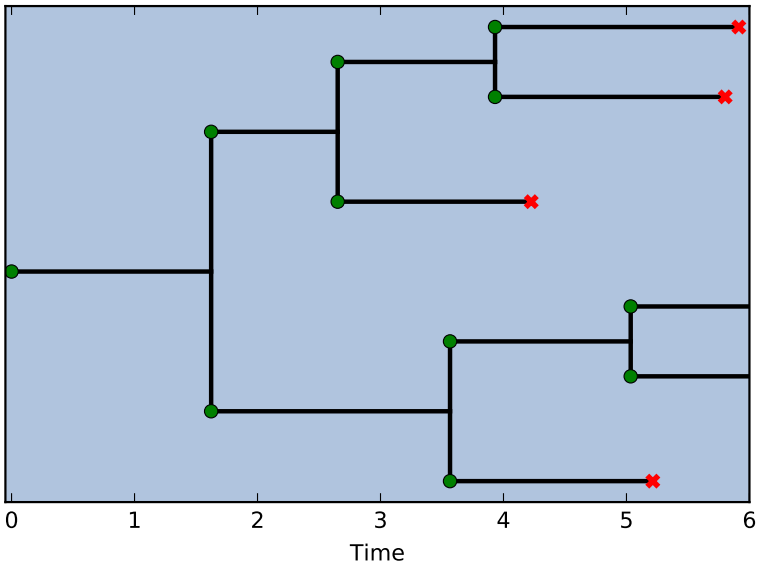


Figure 18: Simulated example of a Bellman-Harris process. Dots indicate an object's birth, horizontal lines indicate a an object's lifetime, vertical lines connect siblings and crosses show death without offspring. In this case, an object's lifetime, J, is uniformly distributed between one and two. For the offspring distribution, H, there is a 0.6 chance for the object to divide into two objects upon death and a 0.4 chance for the object to die without offspring.

4.2.2 Results from a Bellman-Harris process

Supposing we have a defined Bellman-Harris process and want to find summary statistics, such as expected value for the total number of objects at any time t. We do this by deriving relationships for the generating function $G_{Z(t)}(s)$ via the total expectation theorem.

For any specific time t, we partition what may happen into two scenarios: either the original object $\langle\emptyset\rangle$ has ended its lifetime by time t or it has not. We use the law of total expectation to split the generating function into these two cases:

$$G_{Z(t)}(s) = \mathbb{E}(s^{Z(t)}) = sP(J_{\langle\emptyset\rangle} > t) + \mathbb{E}(s^{Z(t)}|J_{\langle\emptyset\rangle} \leq t)P(J_{\langle\emptyset\rangle} \leq t). \tag{6}$$

When an object ends its life at a time u it has a number of new offspring distributed identically to H and the independence property allows us to treat each new object as a fresh Bellman-Harris process with $t-u$ time to divide before time t. Using this property, we adapt the right hand size of Equation 6 and get

$$G_{Z(t)}(s) = sP(J > t) + \int_0^t G_H(G_{Z_{(t-u)}}(s))dP(J \leq u). \tag{7}$$

This gives us an integral equation for the generating function of $Z(t)$ ([Harris, 1964, Ch. 6, Th. 7.1]) and is known as a **non-linear Volterra** integral equation. A proof for the uniqueness and existence

for the solution can be found in [Harris, 1964, Ch. 6.9]. If J is exponential random variable, this can be solved explicitly due to the memoryless nature of the exponential function. There is no closed form in general for other cases, but results can be achieved by numerical analysis and general asymptotic results can be established.

4.2.3 Bellman-Harris process: The exponential case

In the case of exponential lifetimes J with parameter λ so that $F_J(t) = 1 - e^{-\lambda t}$, by differentiating 7 with respect to t, the integral equation is rearranged to a differential equation ([Harris, 1964, Ch. 6, Th. 11.1]):

$$\frac{dG_{Z(t)}(s)}{dt} = \lambda(G_H(G_{Z(t)}(s)) - G_{Z(t)}(s)). \tag{8}$$

We are able solve this for a given offspring distribution H. For example, if every object has probability $1 - p$ to divide into two cells and p to die without offspring, then

$$G_H(s) = (1 - p)s^2 + (p). \tag{9}$$

Substituting (9) into (8), solving the differential equation and using that $G_{Z(0)}(s) = s$ gives the solution for the generating function for $Z(t)$ as

$$G_{Z(t)}(s) = \frac{s}{(1 - p)s + (1 - s - ps)e^{\lambda t}}.$$

Differentiating the generating function with respect to s evaluated at time $s = 1$ gives us

$$\mathbb{E}(Z(t)) = e^{\lambda(1-2p)t}.$$

and we can perform analogous calculations for other moments.

4.2.4 Multi–type Bellman-Harris process

The Bellman-Harris process can be expanded to include multiple types of object with different lifetime distributions and offspring distributions. This is outlined briefly in [Harris, 1964, Ch. 6.28.3], which we explain in this chapter. As in the single case we define the process informally and then formally.

In the previous case, all objects in the process had the same lifetime and offspring distributions, but let us assume in the new process, all objects belong to one of r different types ($r \in N_{>0}$), and each type has different properties. To define the process we must say what objects are alive at time zero, so let us say there is one object alive of type q, with $q = 1, \ldots, r$ (as in the single case this is easily expanded to cases beginning with more than one object). We randomly assign this object a lifetime. At the end of the object's lifetime, it has a random number of children of type one, another number of type two and so on up to type r. We also assign random lifetimes to these new objects and they also

have a random number of children of different types upon their death and so this process continues. The lifetime random variable of each object of type i is independent all other objects' lifetimes and identically distributed to a random variable J_i that takes non negative values on the real line. Also the distribution of number of offspring, which can be thought of as a r length vector of non negative integers for each object type that are independent of each other and identically distributed to some distribution, H^i. These distributions also independent of the lifetime distributions. Figure 19 shows a visual realization of a multi-type Bellman-Harris Process.

We define the set $\langle x \rangle \in \mathbb{X}$ as we did in Section 4.2 and again use these as unique identifiers for objects in families. We then seek to define an array of independent random variables to each object to indicate its lifetime and offspring as we did in the single case. However these random variables now need to be of higher dimensions to represent the cases of the object being any of the r types.

For each object $\langle x \rangle$ we assign it a random lifetime if it is type one, $j_1\langle x \rangle \in \mathbb{R}_{>0}$, another for if it is type two, $j_2\langle x \rangle \in \mathbb{R}_{>0}$, and so on, up to r. We arrange these in a r length vector in $\mathbb{R}^r$ and call it $\vec{j}\langle x \rangle$. We also define $h_{(i,k)}\langle x \rangle \in \mathbb{R}_{\geq 0}$ as the number of children of k type the object $\langle x \rangle$ has if it is as a i type object. We define $\vec{h}_{(i)}\langle x \rangle \in \mathbb{R}^r_{\geq 0}$ as an r length vector of with it's k^{th} component corresponding to $h_{(i,k)}\langle x \rangle$. We assign r of these vectors to each $\langle x \rangle$ for the case of $\langle x \rangle$ parent itself being each of the r types, and we and arrange these in a $r \times r$ matrix in $\mathbb{R}^{r\times r}_{>0}$ and call it $h^{r\times r}\langle x \rangle$. Then for each $\langle x \rangle$, $\vec{j}\langle x \rangle$ and $h^{r\times r}\langle x \rangle$ Thus under this convention a family history is defined by $w = (\vec{j}\langle \emptyset \rangle, h^{r\times r}\langle \emptyset \rangle, \vec{j}\langle 1 \rangle, h^{r\times r}\langle 1 \rangle, \vec{j}\langle 1,1 \rangle, h^{r\times r}\langle 1,1 \rangle, \vec{j}\langle 2 \rangle, h^{r\times r}\langle 2 \rangle, ...)$ with Ω as the space of all ωs. There is redundant information here as only the information relevant to the object $\langle x \rangle$'s actual type will be used.

Let us introduce a set of random variable $\vec{J}\langle x \rangle : \Omega \to \mathbb{R}^r_{\geq 0}$ for each $\langle x \rangle \in \mathbb{X}$ such that $\vec{J}\langle x \rangle(\omega) = \vec{j}\langle x \rangle$. This is as we did for the single case except that this is a vector of independent random variables. We likewise do this for $H^{r\times r}_{\langle x \rangle} : \Omega \to \mathbb{R}^{r\times r}_{\geq 0}$ for each $\langle x \rangle \in \mathbb{X}$ such that $\vec{H}^{r\times r}\langle x \rangle(\omega) = \vec{h}^{r\times r}\langle x \rangle$. So now each ω in Ω corresponds to a denumerable family of independent random variables and so we can define a countable additive probability measure P on measurable sets in Ω.

We wish define a function that maps a object $\langle x \rangle$ to its object type when given a family history ω and a starting object of type q, $n_i(\langle x \rangle, \omega, q) : \mathbb{X} \times \omega \times \mathbb{N}_{>0} \to \mathbb{N}_{>0}$. This function will also return zero if the object does not exist. Let us first define $g(\langle x \rangle) : \mathbb{X} \to \mathbb{N}_0$ to be the number of objects in $\langle x \rangle$, which is the same as the generation and define a function $s(\langle i_1, ..., i_k \rangle) : \mathbb{X} \to \mathbb{X}$ to remove the last ordered object in a sequence, i.e. $s(\langle i_1, ..., i_k \rangle) = \langle i_1, ..., i_{k-1} \rangle$ except in the case of the ancestor object where $s(\langle \emptyset \rangle) = \langle \emptyset \rangle$, thus s returns an object's parent. We define a series of functions

$m_i(\langle x\rangle, \omega, q) : \mathbb{X} \times \Omega \times \mathbb{N}_{>0} \to \mathbb{N}_{>0}$ for $i = 0, 1, 2....$ Define $m_0(\langle x\rangle, \omega, q) = q$. Then for $i > 0$

$$m_i(\langle x\rangle, \omega, q) = \begin{cases} 0, & \text{if } m_{i-1}(s\langle x\rangle, \omega, q) = 0 \text{ or } h_{(v,1)}\langle s\rangle + h_{(v,2)}\langle s\rangle + ... + h_{v,r)}\langle s\rangle < x_k \\ 1, & \text{if } h_{(v,1)}\langle s\rangle \geq x_k \\ 2, & \text{if } h_{(v,1)}\langle s\rangle < x_k \text{ and } h_{(s(x),2)}\langle s\rangle \geq x_k \\ 3, & \text{if } h_{(v,1)}\langle s\rangle + h_{(v,2)}\langle s\rangle < x_k \text{ and } h_{(v,3)}\langle s\rangle \geq x_k \\ & \vdots \\ r, & \text{if } h_{(v,1)}\langle s\rangle + h_{(v,2)}\langle s\rangle ... + h_{(v,r-1)}\langle s\rangle < k \text{ and } h_{(v,r)}\langle s\rangle \geq x_k \end{cases}$$

where $v = m_{i-1}(s\langle x\rangle, \omega, q)$. Then we then define $n(\langle x\rangle, \omega, q) = m_{g(\langle x\rangle)}(\langle x\rangle, \omega, q)$, thus giving us a function that returns the type of a object. The function $m_0(\langle x\rangle, \omega, q)$ is measurable and by induction all $m_i(\langle x\rangle, \omega, q)$ are measurable.

We wish to define random variable for the population of objects of type b at time t under the probability measure we defined over Ω, let us first define a function that designates if an object is alive of type b. The function $a(\langle x_1, ..., x_k\rangle, \omega, q, b, t) : \mathbb{X} \times \Omega \times \mathbb{N}^2_{>0} \times \mathbb{R}_{\geq 0} \to \mathbb{Z}$ is defined as

$$a(\langle x_1, ..., x_k\rangle, \omega, q, b, t) = \begin{cases} 1, & \text{if } n(\langle x\rangle, \omega, q) = b \\ & \quad \text{and } j\langle\emptyset\rangle + j\langle x_1\rangle + ... + j\langle x_1, ..., x_{k-1}\rangle \leq t \\ & \quad \text{and } j\langle\emptyset\rangle + j\langle x_1\rangle + ... + j\langle x_1, ..., x_{k-1}\rangle + j\langle x_1, ..., x_k\rangle > t. \\ 0 & \text{otherwise} \end{cases}$$

This means $a(\langle x\rangle, \omega, q, b, t) = 1$ if object of type b, $\langle x\rangle$ is alive at time t for a family history ω. Let us define a function $Z(\omega, q, b, t) : \Omega \times \mathbb{N}^2_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{Z}$ as

$$Z(\omega, q, b, t) = \sum_{\langle x\rangle \in \mathbb{X}} a(\langle x\rangle, \omega, q, b, t).$$

$Z(\omega, q, b, t)$ is the count of all objects of type b alive at time t for family history ω given a the ancestor object was of type q. Since $Z(\omega, q, b, t)$ is a measurable function on Ω, we again define $Z(q, b, t) : \Omega \to \mathbb{N}_{\geq 0}$ using the probability measure defined by P over Ω, so that $Z(q, b, t)$ as a random variable, representing the number of objects of type b alive at time t when the ancestor object is q, with the distribution $P(\omega \in \Omega : Z(\omega, q, b, t) \leq z) = P(Z(q, b, t) \leq z) = F_{Z(q,b,t)}(z)$.

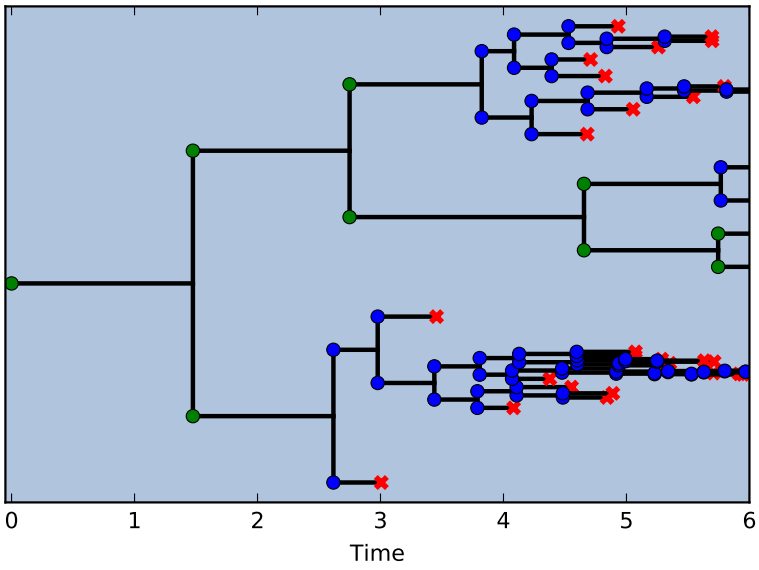


Figure 19: Simulated example of a multi–type Bellman-Harris process. Dots indicate object birth, the colour of dots indicate object type, horizontal lines indicate object lifetime, vertical lines connect siblings, and crosses show death without offspring. Green objects have lifetime uniformly distributed between one and two. At the end of their lifetime, there is a 0.6 chance for the green object to divide and a 0.4 chance for the object to become two blue objects. Blue objects have a lifetime distributed uniformly between 0.25 and 0.5. Upon the end of their lifetime, blue objects can either give birth to two blue objects with 0.5 chance or die without offspring with 0.5 chance. Due to the unidirectional object differentiation in this case, looking only at green objects is an example of a single–type Bellman-Harris process.

4.2.5 Results for the multi-type Bellman-Harris process

Let us suppose we have a multi-type Bellman-Harris process with two types of objects: one and two. Supposing we assume the population starts at $t = 0$ with an object of type one. We want to find the generating function of $Z(1, i, t)$ for $i = 1$ and $i = 2$. We apply the same technique as in the single type case to get an integral equation

$$G_{Z(1,1,t),Z(1,2,t)}(s_1, s_2) = \mathbb{E}(s_1^{Z_{(1,1)}(t)} s_2^{Z_{(1,2)}(t)}) \tag{10}$$

$$= s_1 P(J_1 > t) + \int_0^t G_{H_1}(G_{Z(1,1,t),Z(1,2,t)}(s_1, s_2), G_{Z(2,1,t),Z(2,2,t)}(s_1, s_2)) dP(J_1 \leq u) \tag{11}$$

However, now this equation is dependent on $G_{Z(2,1,t),Z(2,2,t)}(s_1, s_2)$, the generating function for the case when we start with a one object of type two. However, similar to the type one case, we have

$$G_{Z(2,1,t),Z(2,2,t)}(s_1, s_2) = s_2 P(J_2 > t) +$$
$$\int_0^t G_{H_2}(G_{Z(1,1,t),Z(1,2,t)}(s_1, s_2), G_{Z(2,1,t),Z(2,2,t)}(s_1, s_2)) dP(J_2 \leq u)$$

Solving this pair of integral equations will provide the generating function. This technique can be expanded for any r number of object types to get a series of integral equations that need to be solved.

4.2.6 Example of a multi–type Bellman-Harris process

Let us assume that the lifetimes are exponentially distributed for each object type and this distribution is parametrised by λ_i for each object type i, $F_J(t) = 1 - e^{-\lambda t}$. As for the single case, this allows for rearrangement into a series of differential equations. Let us find these differential equations for the moments of a specific example, one similar to that of the optimal model described in [Buchholz et al., 2013]. In this case, for objects of type $i = 1, ..., r-1$, upon the end of their lifetime, each have a $1 - p_i$ chance to divide, that is, become two objects of i type, and a p_i chance to differentiate, with one offspring of $i+1$ type. For the final object type r, division is certain at the end of the object's life.

Under these conditions, the case of starting with an object type r, reduces to the single case. Therefore, the generating function for this case can be found using techniques described in Section 4.2.3. Suppose we already have an equation for the generating function for the case of starting with a single k^{th} object at time zero, that is $G_{Z(k,1,t),...,Z(k,r,t)}(s_1, ..., s_r)$, and we want to calculate the equivalent of this for starting with the $(k-1)^{th}$ object. We have an integral equivalent to (10) for the $(k-1)^{th}$ generating function, but we substitute the solution to the k^{th} object generating function in. We then have an integral equation with only one unknown function, which can be solved. Therefore, we can calculate generating function of a family tree starting from object type one by iteration.

Solving for the generating function itself can be extensive. If we just want to calculate the moments, we can use differentiation and substitution to get a series of differential equations. Below are differential equations that can be solved trivially for r, and then iteratively for the other cases. ($\phi(X) = \mathbb{E}(X^2) - \mathbb{E}(X)$ so requires one extra step to obtain the moment.)

$$\begin{aligned}
\frac{d\mathbb{E}(Z(k,i,t))}{dt} &= \lambda_k(1-2p_k)\mathbb{E}(Z(k,i,t)) + \lambda_k p_k \mathbb{E}(Z(k-1,i,t)) \\
\frac{d\phi(Z(k,i,t))}{dt} &= \lambda_k(1-2p_k)\phi(Z(k,i,t)) + \lambda_k p_k \phi(Z(k-1,i,t)) + 2\lambda_k(1-p_k)\mathbb{E}(Z(k,i,t))^2 \\
\frac{d\mathbb{E}(Z(k,i,t)Z(k,v,t))}{dt} &= \lambda_k(1-2p_k)\mathbb{E}(Z(k,i,t)Z(k,v,t)) \\
&\quad + \lambda_k p_k \mathbb{E}(Z(k+1,i,t)Z(k-1,v,t) + 2\lambda_k(1-p_k)\mathbb{E}(Z(k,i,t)Z(k,v,t))
\end{aligned}$$

These equations allow for a simpler means to calculate the summary statistics of mean, variance and covariance of different object types. However, when looking into models with more than three types, even in this simple case of linear object differentiation, the computations get more laborious with solutions running into tens of lines.

4.3 How the Bellman-Harris process is used in [Buchholz et al., 2013]

The paper [Buchholz et al., 2013] has observed summary statistics of means, variances and covariances for $CD8^+$ T cell families, eight days after infection split by three proposed phenotypes of central memory precursor (TCMp; $CD27^+CD62L^+$), effector memory precursor (TEMp; $CD27^+CD62L^-$) and effector cells (TEF; $CD27^-CD62L^-$). See Section 3.2 for information on how these data are acquired. [Buchholz et al., 2013] looks to see if these summary statistics contain a footprint that gives information as to how the immune response works. To this end, [Buchholz et al., 2013] defines a set of parameterised models as follows (see diagram 20):

- Within each model are four types of cell: Naive, TCMp, TEMp, and TCMp. There is one naive cell alive at time zero.
- Lifetimes of cells are distributed exponentially and are identically distributed according to their cell–type. Each of these exponential distributions need one parameter each and thus there are four parameters needed to describe the distribution of all types of cells' lifetimes.
- Upon the end of a cell's life, there is a chance for each cell–type to divide (except the naive cell which cannot divide) or differentiate into another cell-type (other cell-types cannot differentiate into a naive cell). Possible differentiation between cell-types is known as a path. Different model types have a different combination of paths open between cell types, and paths can be uni- or bi-directional. There must be a direct or indirect path from the naive cell to each cell–type. Thus, each path is another parameter on the model (minus one to account for the naive cell having no probability divide). This describes 304 different models with between seven and 12 parameters.

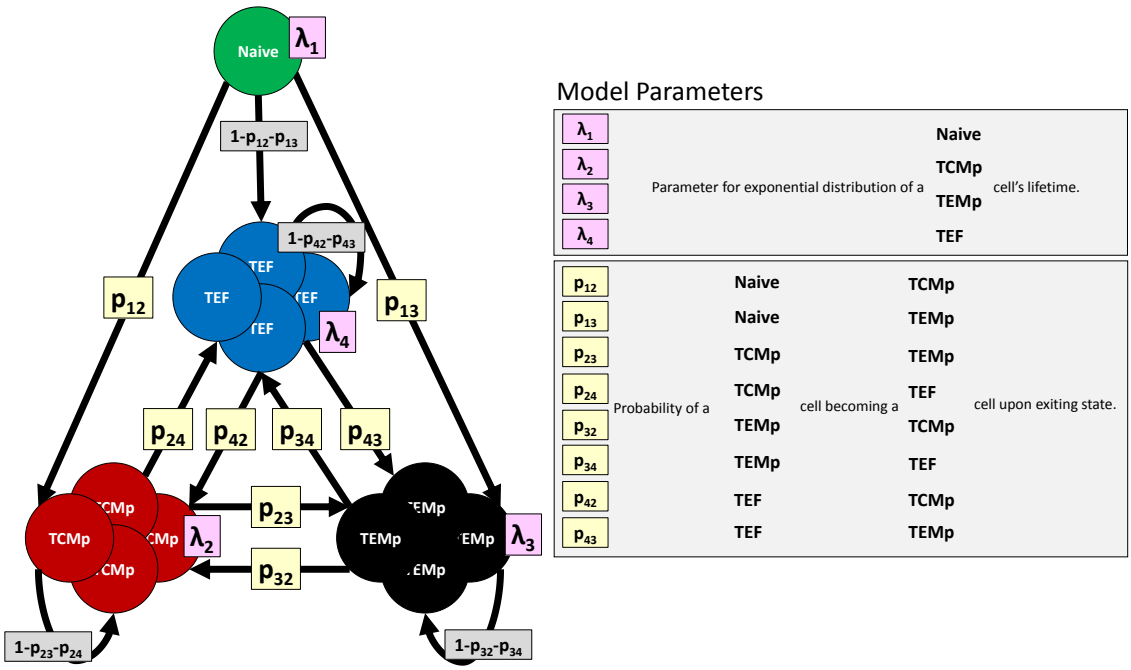


Figure 20: [Buchholz et al., 2013] model scheme allowing for all possible differentiation pathways between phenotype subsets. Of the parameters, four (pink) determine the distribution of the cell's lifetime and eight (yellow) determine the distribution of a cell's fate upon exiting its current state. There are 304 models that can be constructed by closing permutations of different differentiation pathways (i.e. fixing the probabilities of differentiation to zero or one), but leaving at least one path (direct or indirect) from the naive cells to each of the subsets.

For each of the 304 models, [Buchholz et al., 2013] finds the relevant summary statistics for a given parameterisation. [Buchholz et al., 2013] then measures the distance from these expected summary statistics to the observed summary statistics using χ^2 minimization, adjusted for degrees of freedom where bootstrapping was used to acquire a sample variance required. Therefore [Buchholz et al., 2013] has 304 objective functions for each of the models. For each 304 models [Buchholz et al., 2013] then minimises using

> "...both simulated annealing and a local optimization algorithm (trust-region-reflective algorithm using Matlabs optimization toolbox) with at least 300 different random initial values.[Buchholz et al., 2013, supplementary material page. 5]"

[Buchholz et al., 2013] repeats the fitting for different sets of summary statistics: in one case fitting to mean, variance and covariance of phenotype populations on day eight, and in another case fitting to the same data but including proportional phenotype data. They also separately rank each model using both AICc and χ^2 minimization adjusted for degrees of freedom which, in this case, reduce to the same best fitting models. [Buchholz et al., 2013] uses the convention that an AICc of above ten is unsupported by the model and, of those that remain, [Buchholz et al., 2013] excludes those that

predict an absence of the TEMp subset in a certain threshold, 42%-56% in this case. The remaining models are accepted as the best fit. Refer to Section 5.1.1 for a full description of results and our recreation.

5 Novel adaptation the fitting method to other data sets

Objective. The authors of [Buchholz et al., 2013] used a methodology of fitting probabilistic models to observed clonal data in order to rank cell differentiation pathways during $CD8^+$ T cell expansion phase (Section 4.3). In this section we first recreate their results to establish that our implementation of their method is correct (Figures 21, 22, 23). We adapt their method so that it can be applied to statistics reported other papers. We first test this adaptation on the original data from [Buchholz et al., 2013] and then apply it to [Kinjyo et al., 2015] and [Schlub et al., 2010] to see if the resulting best fitting model changes.

Fitting to proportions. The adaptation of the methodology to new data requires the removal of dependency on statistics that can only be determined from clonal data. We recreate the [Buchholz et al., 2013] methodology, but fitting to the reported cohort data (reported proportions of phenotypically defined cell types at different times), which is more readily available in other papers.

We establish that fitting the model to cohort data gives essentially the same result as when fitting to the clonal statistics, the best fitting probabilistic model being Naive $\rightarrow$ $CD62L^+CD27^+$ $\rightarrow$ $CD62L^-CD27^+$ $\rightarrow$ $CD62L^-CD27^-$. We also see a similar best fitting parameterisation (Figures 23, 24, 25, 26). This leads to the deduction that in some circumstances, fitting to cohort data is a reasonable alternative to fitting clonal data, which is less widely available.

Adjusting to data reported in other papers. To be able to apply the [Buchholz et al., 2013] method to other data sets requires further changes. We change the objective function from a χ^2 to a weighted MSE, in order to remove dependency on sample variance. We reduce the number of phenotypes to two, $CD62L^+$ and $CD62L^-$, thus removing dependency on CD27. This reduces the models to two: a Linear Memory First Model or a Linear Effector First Model (Figure 27), representative of those described qualitatively in Section 3.1.

Taking data from the paper as-is and using this adapted method we find, consistent with the conclusions of the respective papers, that [Buchholz et al., 2013] data fits best to a Linear Memory First Model, whereas [Kinjyo et al., 2015] data fits the best to a Linear Effector First Model (Figure 33).

Adjusting for experimental differences. The number of T cells adoptively transferred in these experiments are different: $10^2 - 10^4$ cells transferred in [Buchholz et al., 2013] and 10^6 in [Kinjyo et al., 2015] (see Chapter 3). The papers [Schlub et al., 2010] and [Badovinac et al., 2007] both observed a shortened expansion phase for experiments with high adoptive transferred numbers, such as those in [Kinjyo et al., 2015] (Figures 28, 29, 32). As a result, the naive application of the [Buchholz et al., 2013] methodology, which is conditioned on asking when memory occurs should it occur during the expansion phase, to the [Kinjyo et al., 2015] data is inappropriate as this likely includes a contraction phase. When this is adjusted for, the [Kinjyo et al., 2015], [Buchholz et al., 2013] and [Schlub et al.,

2010] data all fit best to a Linear Memory First Model (Figure 33, 34, 35) under the assumption of the method. The best fitting parameterisations to [Kinjyo et al., 2015] data shows few memory cells appearing during the expansion phase. This, along with evidence presented in the [Kinjyo et al., 2015] paper itself, such as a subset of slowly proliferating cells of memory phenotype appearing after the end of the expansion phase (day four-five in their case), may suggest the assumption memory appears during the expansion phase could be challenged.

As a second, more minor point, we suspect the cells [Buchholz et al., 2013] reported as TCMp type also include cells that would traditionally be described as naive (Baars et al. [2005]) and we adjust for this in the method. This does not change the best fitting model to [Buchholz et al., 2013] data.

Headline result. Initially, the papers [Buchholz et al., 2013] and [Kinjyo et al., 2015] have apparent contradictions, this is both in the conclusion the papers come to themselves and our analysis on their data. However, correcting for the number of adopted cells transferred changes this. If all assumption of the method are true and memory is made during the expansion phase, then the [Buchholz et al., 2013] methodology consistently supports a memory first path of differentiation above an effector first path when fitting to all blood and spleen data. If memory is made after the expansion phase is complete, as the [Kinjyo et al., 2015] data seems to suggest is possible, then [Buchholz et al., 2013] cannot speak for this possibility.

5.1 Recreating [Buchholz et al., 2013]

5.1.1 Amendments to [Buchholz et al., 2013] fitting methodology

We recreated results of the model fitting method as described in [Buchholz et al., 2013] with some simplification and changes.

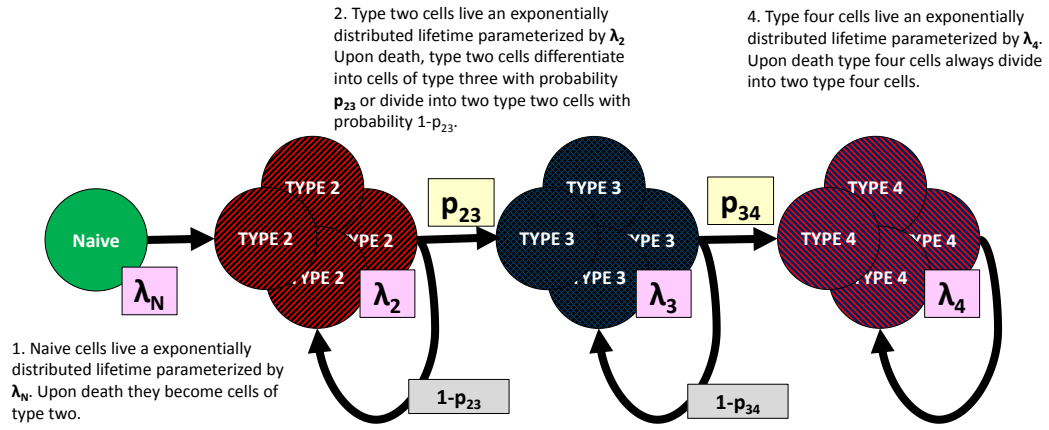
Firstly, [Buchholz et al., 2013] fits to the 304 probabilistic models to find the best fitting model (Section 4.3 for a full description of these models). We fit the data to a small subset of these. Specifically, we only fit linear models, i.e. models were there is a single uni-directional path between phenotypes (Figure 21). We index the six permutations of these models for brevity:

- **Model 1**: Naive $\rightarrow$ CD62L$^+$CD27$^+$ $\rightarrow$ CD62L$^-$CD27$^+$ $\rightarrow$ CD62L$^-$CD27$^-$
- **Model 2**: Naive $\rightarrow$ CD62L$^+$CD27$^+$ $\rightarrow$ CD62L$^-$CD27$^-$ $\rightarrow$ CD62L$^-$CD27$^+$
- **Model 3**: Naive $\rightarrow$ CD62L$^+$CD27$^-$ $\rightarrow$ CD62L$^+$CD27$^+$ $\rightarrow$ CD62L$^-$CD27$^-$
- **Model 4**: Naive $\rightarrow$ CD62L$^+$CD27$^-$ $\rightarrow$ CD62L$^-$CD27$^-$ $\rightarrow$ CD62L$^+$CD27$^+$
- **Model 5**: Naive $\rightarrow$ CD62L$^-$CD27$^-$ $\rightarrow$ CD62L$^+$CD27$^+$ $\rightarrow$ CD62L$^+$CD27$^-$
- **Model 6**: Naive $\rightarrow$ CD62L$^-$CD27$^-$ $\rightarrow$ CD62L$^-$CD27$^+$ $\rightarrow$ CD62L$^+$CD27$^+$

Model 1 is that which [Buchholz et al., 2013] reported as one of the two best fitting of the 304 models they tested.

For each model, the first cell-type that a naive cell differentiates to we denote as cell-type two, the second cell-type in the linear differentiation path is denoted type three and the third is denoted as cell-type four. Each cell-type requires a parameter to define its lifetime distribution which we denote as: $\lambda_N, \lambda_2, \lambda_3, \lambda_4$. The probability a cell-type two differentiates upon exiting its current state is denoted p_{23} and the probability cell-type three differentiates is denoted p_{34}. Cell type four is terminally differentiated cell and can only divide. Thus, for each model we have parameters $\vec{\theta} \in \mathbb{R}^6$, $\vec{\theta} = (\lambda_N, \lambda_2, \lambda_3, \lambda_4, p_{23}, p_{34})$.

We numerically attempt to identify the $\vec{\theta}$ that minimises $\chi^2(\hat{x}, \vec{x}(\vec{\theta}))$ using the method described in Section 4.3. Since all models have the same number of parameters, we ae able to rank them with a χ^2 objective function directly.



MODEL INDEX	TYPE 2 PHENOTYPE	TYPE 3 PHENOTYPE	TYPE 3 PHENOTYPE
MODEL 1	$CD62L^+CD27^+$	$CD62L^-CD27^+$	$CD62L^-CD27^-$
MODEL 2	$CD62L^+CD27^+$	$CD62L^-CD27^-$	$CD62L^-CD27^+$
MODEL 3	$CD62L^-CD27^+$	$CD62L^+CD27^+$	$CD62L^-CD27^-$
MODEL 4	$CD62L^-CD27^+$	$CD62L^-CD27^-$	$CD62L^+CD27^+$
MODEL 5	$CD62L^-CD27^-$	$CD62L^+CD27^+$	$CD62L^-CD27^+$
MODEL 6	$CD62L^-CD27^-$	$CD62L^-CD27^+$	$CD62L^+CD27^+$

Figure 21: Index for six linear differentiation models being fitted to. These models are a subset of the 304 Bellman-Harris processes modelled by [Buchholz et al., 2013] and described in Section 4.3. Each of the six models has six parameters, $\vec{\theta} = (\lambda_N, \lambda_2, \lambda_3, \lambda_4, p_{23}, p_{34})$.

[Buchholz et al., 2013] does not explicitly state if these parameters are bounded during their optimisation routine. We assume $\lambda_N, \lambda_2, \lambda_3, \lambda_4 \in \{x \in \mathbb{R} | 0 < x \leq 3\}$, as it is biologically implausible for a cell to take less than eight hours to divide on average ([Hawkins et al., 2009],[Kinjyo et al., 2015]). We also take $p_{23}, p_{34} \in \{x \in \mathbb{R} | 0 < x < 1\}$.

For each model, when minimizing the objective function we use the python library *scipy.optimize.minimize*. We use 50 starting locations ($\vec{\theta_0}$) randomly selected by uniform distribution between the boundaries of each parameter. For parameters that must be strictly between their bounders and not on them, we allow these parameters to takes values up to three significant figures from that boundary. [Buchholz et al., 2013] uses 300 starting points and does not explicitly state how they are randomly chosen. For each starting point we repeat the data fitting with two different algorithms: **Truncated Newton Algorithm** (TNC,Nash [1985]) and **Sequential Least Squares Programming Algorithm** ($SLSQP$,[Nocedal and Wright, 2006, Ch. 18]) and take the minimum of them both.

5.1.2 Result of the recreation of [Buchholz et al., 2013] fitting

The results matched [Buchholz et al., 2013] when fitting models to mean cell counts, variances and covariances. We find that Model 1 ($CD62L^+CD27^+ \rightarrow CD62L^-CD27^+ \rightarrow CD62L^-CD27^-$) is the best fit model to the data, as the authors of [Buchholz et al., 2013] found. We also find a similar parameterisation for the best fit as seen in Table 5 and Figure 22. The slight differences are explained by the fact that both the observed statistics and the [Buchholz et al., 2013] best fit parameterisation we compared to were read manually off graphs (supplementary material S26, S17 and respectively, Figure 4E reported the proportional data).

Parameter	**A. Value read from [Buchholz et al., 2013]**	**B. Recreated value using [Buchholz et al., 2013] method**
λ_N^{-1} (days)	0.41	0.34
λ_2^{-1} (days)	0.95	1.03
λ_3^{-1} (days)	0.69	0.66
λ_4^{-1} (days)	0.63	0.63
p_{23}	0.19	0.15
p_{34}	0.03	0.03

Table 5: Best fitting parameters described in [Buchholz et al., 2013] compared with parameters we derive for Model 1 ($CD62L^+CD27^+ \rightarrow CD62L^-CD27^+ \rightarrow CD62L^-CD27^-$). [Buchholz et al., 2013] parameters were read from a graph in the supplementary material S17. The parameterisation is similar between the two best fitting models.

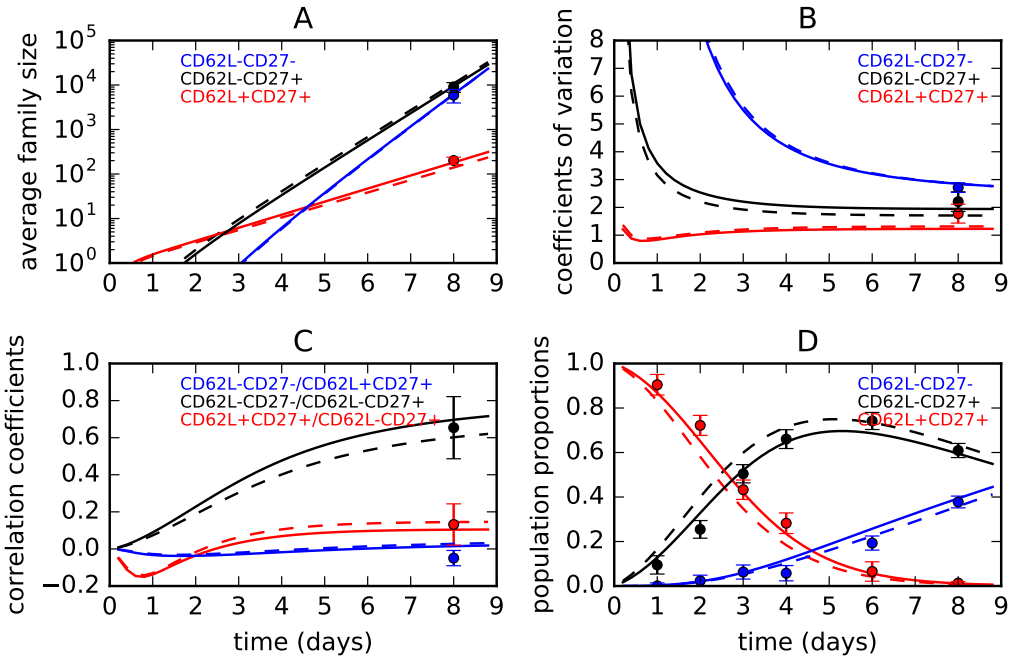


Figure 22: Statistics from the for Model 1 (CD62L$^+$CD27$^+$→CD62L$^-$CD27$^+$→ CD62L$^-$CD27$^-$) using [Buchholz et al., 2013] best fit parameterisation (dashed line) and a recreated best fit parameterisation (solid line). Statistics shown are: A mean cell count, B coefficients of variation, C correlation coefficients and D phenotype proportions. Observed data (dots) are taken from data graph points in [Buchholz et al., 2013] supplementary material S26. Error bars show the standard error of the mean. The statistics are similar between the two best fitting models.

5.2 Adapting the model fitting methodology described in [Buchholz et al., 2013] to work with proportional data

5.2.1 Changing objective function

To transfer the methodology as developed for clonal data to one suitable for cohort data, we repeat the methodology described in Section 5.1, but instead of fitting to the mean cell count, variance and covariance of the three phenotypes, we fit to the proportions of the three phenotypes observed on day one, two, three, four, six and eight post infection reported in [Buchholz et al., 2013], and also fit to the average size of a clonal family at day eight. This last statistic is required as otherwise the fitting will have no concept of scale. Thus the observed statistics will be defined as $\hat{x} \in \mathbb{R}^{19}$.

Because we are looking at proportional data at different times, we have a new data set in addition to that of the data set $O_{47\times3}$ we defined in Section 4.1. Each of these new data points we denote as $q_{t,i,j}$, which are counts of cells of phenotype i ($i \in \{2,3,4\}$), in replicate experiment j ($j \in \{1,...,47\}$) harvesting at time period indexed as t ($t \in \{1,...,6\}$) post infection. The data set of all such points we consider a matrix $Q_{6\times47\times3} \in \mathbb{R}^{6\times47\times3}$. Thus the vector of summary statistics are defined by a function

$\vec{s} : O_{47\times3} \times Q_{6\times47\times3} \to \mathbb{R}^{19}$ and for brevity we denote $\hat{x} = \vec{s}(O_{m\times n} \times Q_{6\times47\times3})$. The observed data $Q_{6\times47\times3}$ is at a population level (from multiple transferred progenitor cells per mouse) and the cell counts are not absolute because they do not take into account cell loss during the counting process.

We only require one statistic using the data $O_{m\times n}$, which is the average total family size at day eight post infection, which we denote as $\hat{x_1}$ given by

$$\hat{x}_1 = s_1(O_{m\times n} \times Q_{6\times47\times3}) = \frac{1}{47}\sum_{i=1}^{47} o_{i,1} + o_{i,2} + o_{i,3}.$$

The 18 average phenotype proportion statistics are defined as

$$\hat{x}_{(3t+i-2)} = s_{(3t+i-2)}(O_{m\times n} \times Q_{6\times47\times3}) = \frac{1}{47}\sum_{j=1}^{47} \frac{q_{t,j,i}}{(q_{t,j,1} + q_{t,j,2} + q_{t,j,3})}$$

$$\text{for } t = 1, ..., 6 \text{ and } i = 1, 2, 3.$$

These 18 statistics do not incorporate elements of $O_{m\times n}$.

We also require the equivalent model statistics, $\vec{x}(\vec{\theta})$, to compare the estimates $\hat{x}$ to. Recall from Section 4.2.4 that if we have in a multi-type Bellman-Harris process and at time zero we have a single cell of type 1 (naive), we denote as the expected number of cells of type j at time t as $\mathbb{E}(Z(1, j, t))$. The model statistic for average family size at day eight, i.e x_1, is

$$x_1 = \mathbb{E}(Z(1, 2, 8)) + \mathbb{E}(Z(1, 3, 8)) + \mathbb{E}(Z(1, 4, 8)). \tag{12}$$

Because we do not have the raw data we cannot calculate the SEM of the total population directly, but can approximate it by assuming independence and summing the SEMs of the individual phenotype populations. The naive cell data are not reported in [Buchholz et al., 2013] and for this calculation they are assumed to be excluded from the observed statistics.

Let t_j be the times post infection observed statistics were measured in [Buchholz et al., 2013] indexed by j, with $j = 1, .., 6$, that is $t_1 = 1, t_2 = 2, t_3 = 3, t_4 = 4, t_5 = 6$ and $t_6 = 8$. Then the proportional model summary statistics we wish to compare to observed statistics are defined as

$$x_{(3k+j-2)} = \frac{\mathbb{E}(Z(1, j+1, t_k))}{\mathbb{E}(Z(1, 2, t_k)) + \mathbb{E}(Z(1, 3, t_k)) + \mathbb{E}(Z(1, 4, t_k))} \text{ for } k = 1, .., 6 \text{ and } j = 1, 2, 3. \tag{13}$$

We then find numerically $\arg\inf_{\vec{\theta}} \chi^2(\hat{x}, \vec{x}(\vec{\theta}))$ via the methods described in Chapter 4.

5.2.2 Naive cell classification in [Buchholz et al., 2013]

The authors of [Buchholz et al., 2013] observed three phenotypes within their experiment and assigned each of these a cell-type: $CD62L^+CD27^+$ cells were labelled as TCMp cells, $CD62L^-CD27^+$ cells were labelled as TEMp cells and $CD62L^-CD27^-$ cells were labelled as TEF cells. However, there were four cell types within the suggested model, the forth being the naive cells, and there were no reported statistics for observations of naive cells after infection. This could be interpreted two ways:

- that naive cells have been excluded from the data and thus ignored during the model fitting,
- or naive cells were included in the count of the other phenotypes.

Equations 12 and 13 assume the first interpretation.

Whichever of these interpretations is correct has negligible impact on statistics taken on day eight. This is because for the average number of naive cells in a family must be less than one and thus would have small impact on the large populations of cells seen on day eight. Also, biologically plausible models should have the majority of naive cells activated at the time of peak immune response. Thus, this ambiguity would have a small effect on the results when using the methodology described by [Buchholz et al., 2013]. However, if naive cells were included in another phenotype, this may have a significant impact on the proportional statistics observed at earlier time points in the immune response, when total populations are lower and many naive cells are still not differentiated.

Naive cells have a high expression of CD62L, however reports of CD27 expression on naive cells have been conflicted (Hendriks et al. [2000], Baars et al. [2005], [Kinjyo et al., 2015]) and papers have reported on now naive cells are often mis-categorised as memory cells (Murali-Krishna and Ahmed [2000]). [Buchholz et al., 2013] themselves define naive cells as CD62L$^+$ CD27$^+$ ([Buchholz et al., 2013] Supplementary Figure 3), but with a different gating strategy than used for the classification of other cell types ([Buchholz et al., 2013] supplementary Figure 22). It is therefore plausible that naive cells are counted within the reported TCMp (CD62L$^+$CD27$^+$) count in [Buchholz et al., 2013].

We wished to see how the results would be affected if we changed the proportional fitting method to account for this. Let us suppose within the linear model for CD8$^+$ T cell expansion described in Section 5.1.1, the TCMp cells are the i^{th} cell-type in the linear differentiation path, with naive cells being cell-type one and with the i^{th} type being either two, three or four. So using the notation as in Equation 13 but with naive cells being counted in the i^{th} phenotype, model proportional summary statistics are calculated as

$$x_{(3k+j-2)} = \frac{\mathbb{E}(Z(1,j,t_k))}{\mathbb{E}(Z(1,1,t_k)) + \mathbb{E}(Z(1,2,t_k)) + \mathbb{E}(Z^(1,3,t_k)) + \mathbb{E}(Z^(1,4,t_k))} \text{ when } i \neq j$$

$$x_{(3k+j-2)} = \frac{\mathbb{E}(Z(1,1,t_k)) + \mathbb{E}(Z(1,j,t_k))}{\mathbb{E}(Z(1,1,t_k)) + \mathbb{E}(Z(1,2,t_k)) + \mathbb{E}(Z(1,3,t_k)) + \mathbb{E}(Z(1,4,t_k))} \text{ when } i = j,$$

$$\text{for } k = 1,..,6 \text{ and } j = 1,2,3.$$

The model statistic for total average family size at day eight, i.e x_1, is also changed to

$$x_1 = \mathbb{E}(Z(1,1,8)) + \mathbb{E}(Z(1,2,8)) + \mathbb{E}(Z(1,3,8)) + \mathbb{E}(Z(1,4,8))$$

although this will have small impact. We then fit these new model statistics to their respective observed statistics as before.

Throughout the rest of this chapter, unless otherwise stated, we assume that naive cells are included in the TCMp count. An analysis of the impact of assuming the alternative are considered in Appendix C.1.

5.2.3 Results of fitting to cohort data

Using the methodology of fitting models to cohort data, we found Model 1 ($CD62L^+CD27^+ \rightarrow CD62L^-CD27^+ \rightarrow CD62L^-CD27^-$) to be the best fit, as was the case when using clonal data to fit.

If we take the data as-is, and adopt the assumption that naive cells were excluded from the reported data, as opposed to naive cells included within TCMp cell counts, the conclusion of Model 1 being the best fit is stronger than the original fit to clonal data, with Model 1 having an objective function for its best fitting parameterisation being fifty times lower than any other model (Appendix Figure 40). For a more detailed analysis on the results using this different assumption see Appendix C.1.

When assuming naive cells are included with the counts of TCMp cells, we reach the same conclusion, with Model 1 being the best fit. Thus, we find Model 1 is clearly the best fit and results are the same regardless of which assumption is taken. This suggested that fitting to cohort data could be an appropriate alternative to fitting to clonal data. The parameterisation for the best fitting model was similar in both the original and adjusted method for five parameters out of six, as shown in Figure 24. The difference in the parameter λ_N, which parametrises the lifetime of naive cells, could be explained by the fact that the adjusted method compares the sum of proportions of naive and TCMp cells to one statistic, the proportion of $CD62L^+CD27^+$ cells, so that while the model is able to fit well to the portion of the sum of the two cell types, it is not able to discriminate between their separate contributions. Models one showed a good fit visually when comparing model and observed phenotype proportions (Figure 25). The returned statistics of the best fitting model were also visually similar between the original and adjusted method as seen in Figure 26, further supporting that fitting to clonal or cohort data produces similar results.

The magnitudes of the objective function for the cohort fit are larger than that of objective function for the clonal fit. One factor that explains this is there are nine observed statistics to fit to when fitting to means variances and covariances against 19 when fitting to cohort data. Also, the sample variances are smaller for the proportion statistics and so errors between observed and model statistics are penalised more. Thus it is not appropriate to compare the absolute values of two different objective functions for the same model, rather we should compare two different models using the same objective function.

Together, these results suggest that for this model the fitting to proportional data with an average family size is a reasonable alternative to fitting to the mean count of cells, variances and covariances.

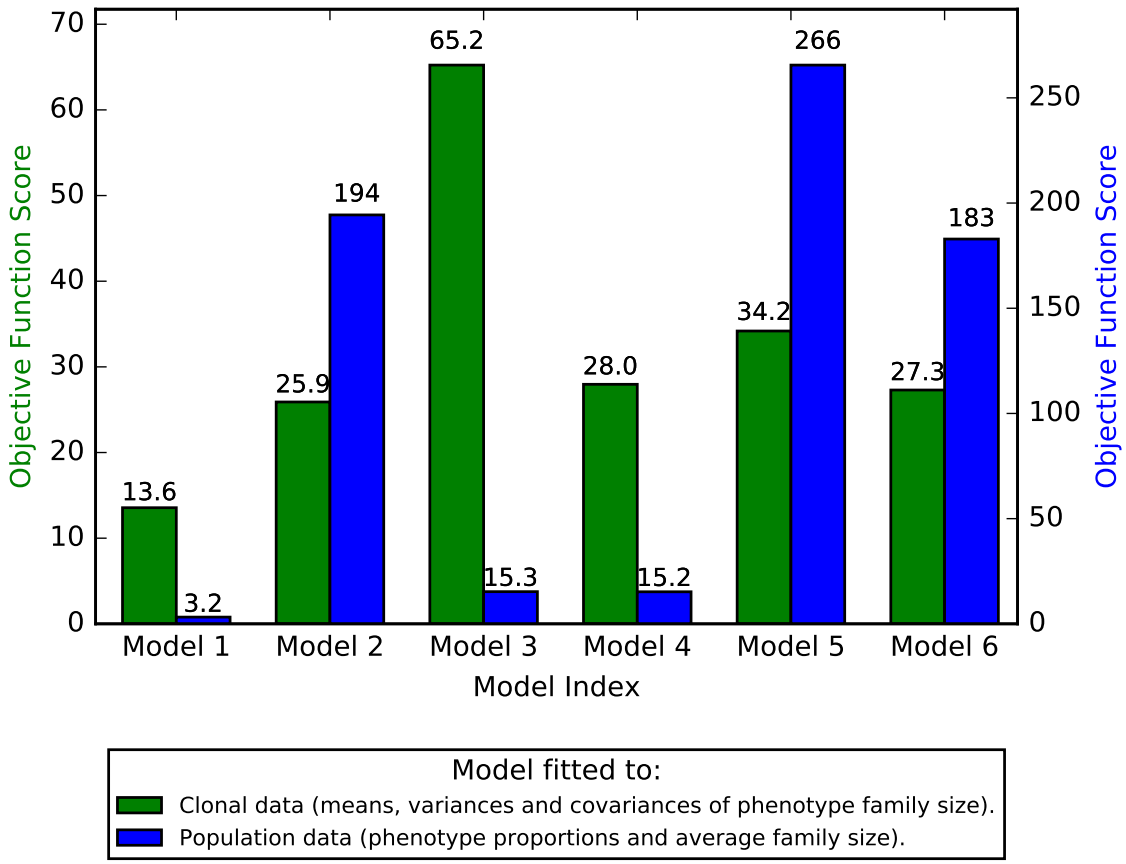


Figure 23: The objective function value for the best fitting parameterisations of the six linear models when fitting to observed statistics of mean cell count, variances and co–variances as in [Buchholz et al., 2013] (green, as in Section 5.1.1) compared with fitting to proportion data and average family size (blue). Smaller values indicate a better fit. Model 1 ($CD62L^+CD27^+ \rightarrow CD62L^-CD27^+ \rightarrow CD62L^-CD27^-$) is the best fitting in both cases, however fitting to population data rather than clonal reduces the strength of this result.

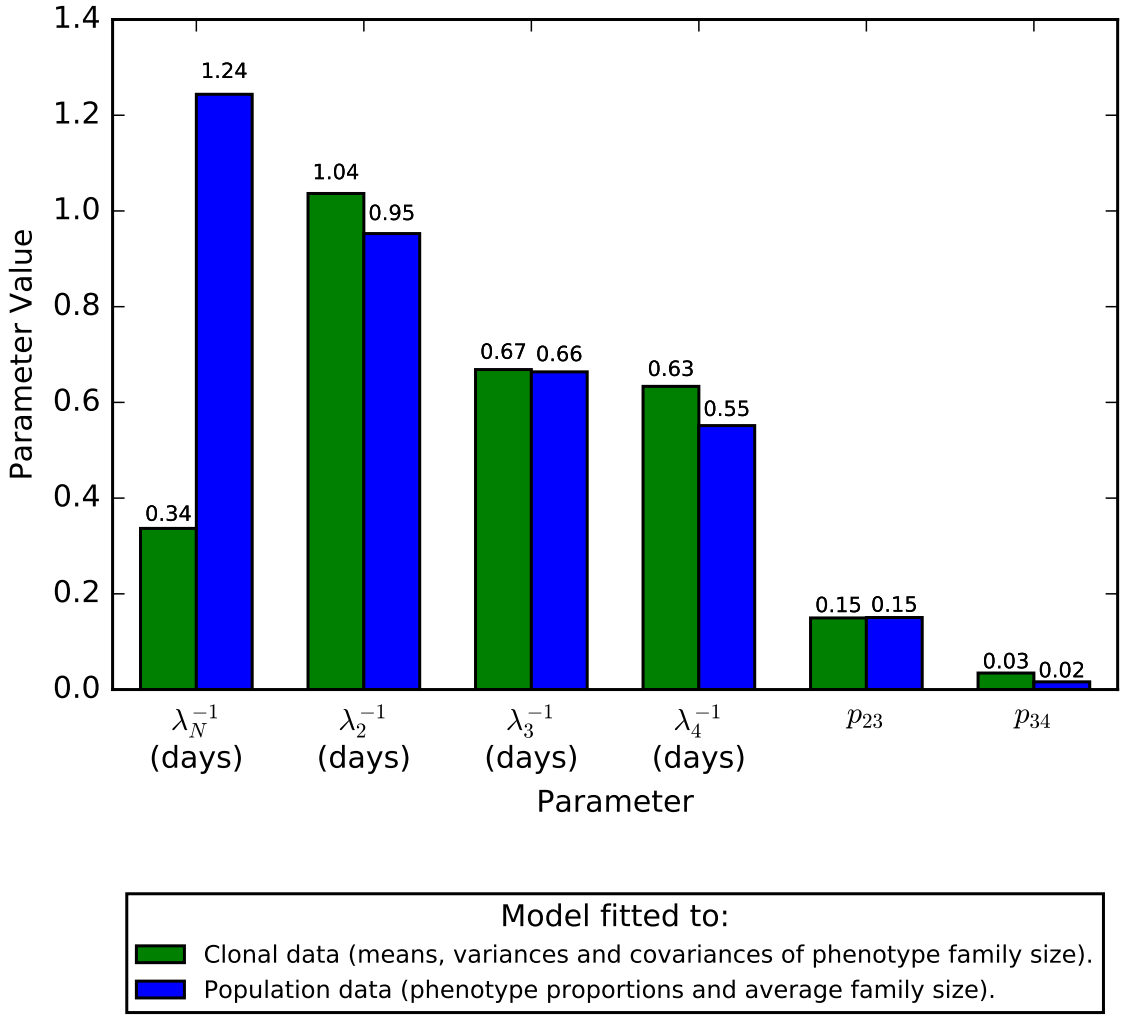


Figure 24: The best fitting parameters derived by fitting against means of cell count, variances and co-variances alongside parameters from fitting to proportional data and average family size for Model 1 (CD62L$^+$CD27$^+$→CD62L$^-$CD27$^+$→ CD62L$^-$CD27$^-$). The parameters are similar between the two best fitting models with the exception of λ_N.

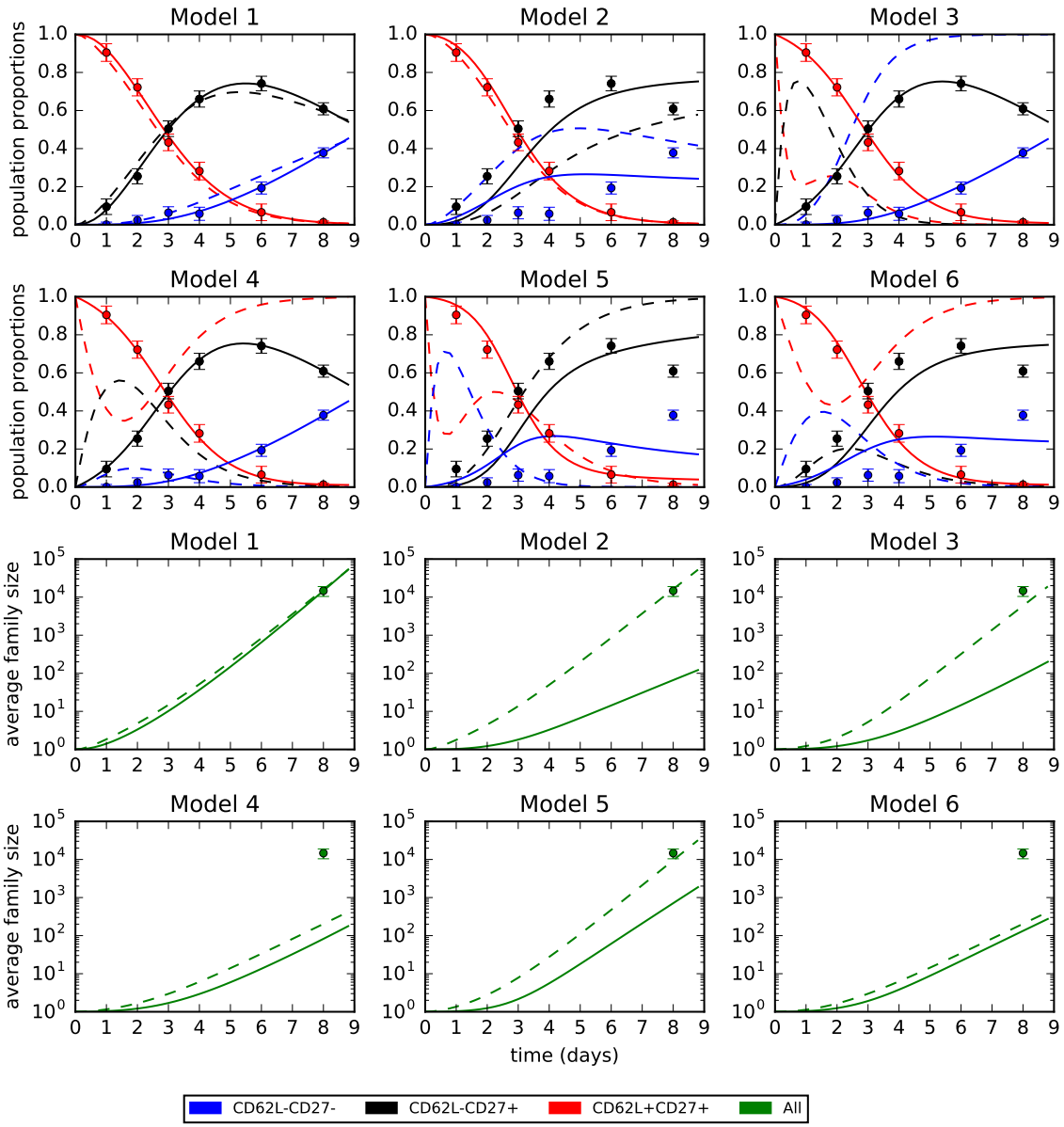


Figure 25: Proportions and average family size of phenotype of best fitting parameterisation for each of the six linear models. Dashed lines show when fitting to mean cell counts, variances and covariances and solid lines shows when against proportional data and average family size. Observed data (dots) was taken from data graphs in [Buchholz et al., 2013] supplementary material S26. Error bars show the standard error of the mean. Only Model one ($CD62L^{+}CD27^{+} \rightarrow CD62L^{-}CD27^{+} \rightarrow CD62L^{-}CD27^{-}$) fits well visually to the observed data when fitting to clonal data.

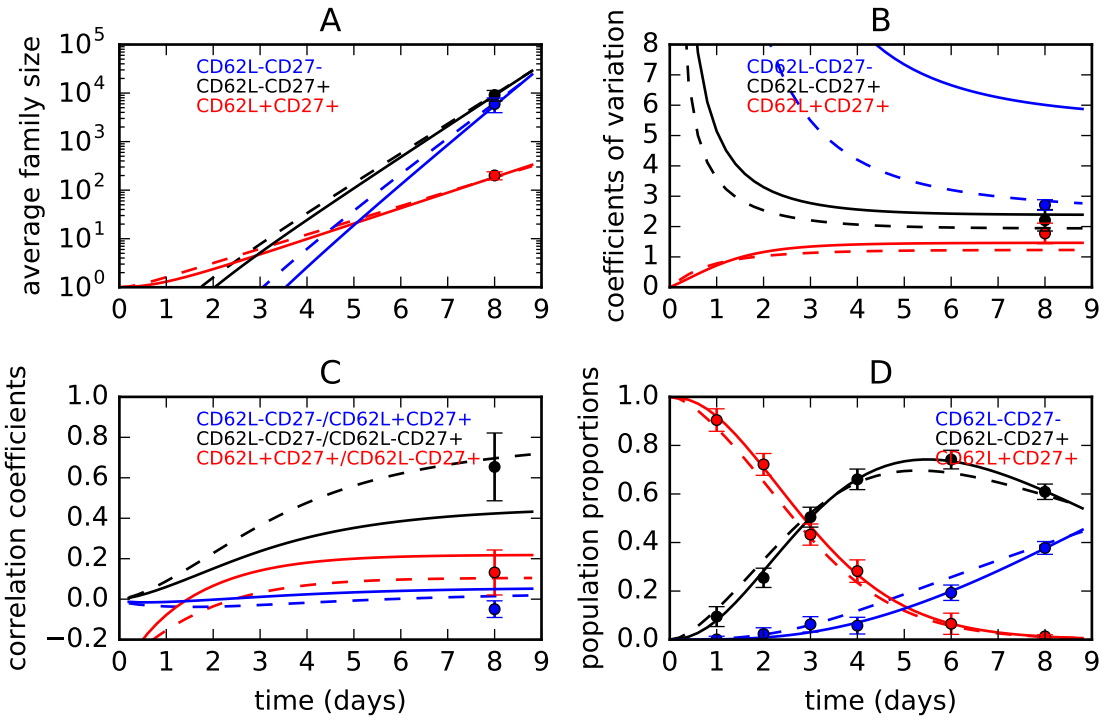


Figure 26: Statistics from Model 1 ($CD62L^+CD27^+ \rightarrow CD62L^-CD27^+ \rightarrow CD62L^-CD27^-$) for the best fit parameterisation fitting to mean cell count, variances and covariances (dashed line) compared with fitting to proportional data and average family size at different times (solid line). Statistics shown are: A mean cell count, B coefficients of variation, C correlation coefficients and D phenotype proportions. Observed data (dots) were taken from data graphs in [Buchholz et al., 2013] supplementary material S26. Error bars show the standard error of the mean. The statistics are similar between the two best fitting models with the exception of the $CD62L^-CD27^-$ variance and its covariance to $CD62L^+CD27^+$. Both fits assume naive included in $CD62L^+CD27^+$.

5.3 Further adaptations to method

5.3.1 Changing the weighting of the objective function

The χ^2 objective function that [Buchholz et al., 2013] used to rank models required estimates of the sample variance ($\sigma^2_{x_i}$) for each observed statistic, as defined in Section 4.1.2. However, the paper [Kinjyo et al., 2015] reports only the proportional statistics and thus there is not sufficient data to make the bootstrapping calculation for their sample variance. Therefore we need to select a different objective function not dependent on this missing data.

A simple MSE objective function is not appropriate. This is because differences in scale between the total population statistics and the proportional statistics will lend heavy weighting the population statistics.

It is possible to apply a weighting to the MSE to correct for this. One possibility is weighting each element of the sum by the inverse of the observed statistic squared. This would weigh the percentage difference between any two statistics equally. We denote the objective function $\mathbf{WMSE_1}$, defined as

$$\text{WMSE}_1(\hat{x}, \vec{x}(\vec{\theta})) = \sum_{i=1}^{19} \frac{(\hat{x_i} - \vec{x_i}(\vec{\theta}))^2}{\hat{x_i}^2}. \tag{14}$$

This objective function has some disadvantages. It is not defined if any element of $\hat{x}$ takes the value zero, thus we redefine the domain $\text{WMSE}_1 : \mathbb{R}_{>0}^{19} \times \mathbb{R}^{19} \to \mathbb{R}$. However this is not such an issue for this thesis as all proportional statistics within in [Buchholz et al., 2013] are greater than zero. This objective function will heavily weigh absolute differences between statistics that are close to zero.

Alternatively, we use another WMSE where we only divide one of the sum's elements by the respective observed statistic squared, the observed populations at day eight post infection. We leave the proportional statistics not weighted. We denote this objective function $\mathbf{WMSE_2}$ as

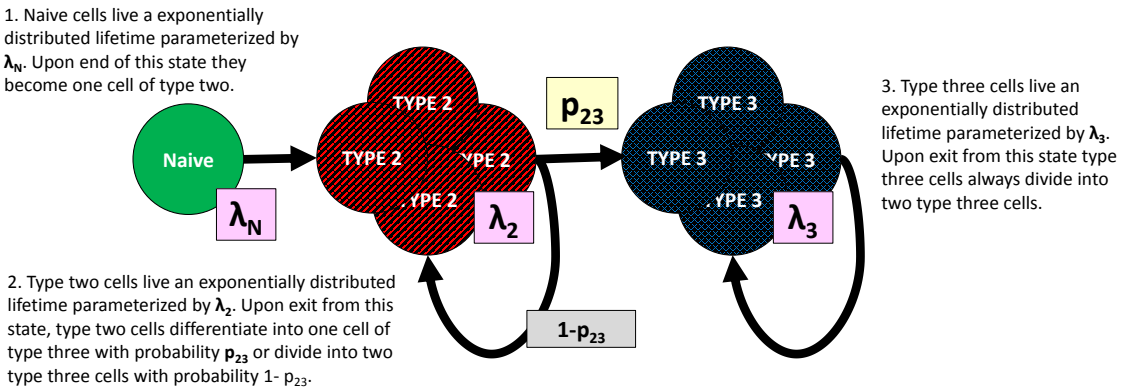
$$\text{WMSE}_2(\hat{x}, \vec{x}(\vec{\theta})) = \frac{(\hat{x_1} - \vec{x_1}(\vec{\theta}))^2}{\hat{x_1}^2} + \sum_{i=2}^{19} (\hat{x_i} - \vec{x_i}(\vec{\theta}))^2 \tag{15}$$

This still has the issue that the total population at day eight cannot be zero, $\text{WMSE}_2 : \mathbb{R}_{>0} \times \mathbb{R}^{37} \to \mathbb{R}$. It also arbitrarily weights statistics in a different ways, but does reduce the issue of proportional statistics having heavier weightings as they approach zero. From henceforth in this thesis unless otherwise stated we use the WMSE_2 objective function. We did an evaluation of the effects of using the WMSE_1 objective function and the same results were reached. Detailed analysis is available in Appendix C.2.

5.3.2 Reducing the number of phenotypes

While the [Kinjyo et al., 2015] does have some $CD27^+$ data it is not in a format we can use for this analysis. What [Kinjyo et al., 2015] does provide is the percentage of $CD62L^+$ cells at different times post infection. We convert the data in [Buchholz et al., 2013] to this format by adding their $CD62L^+CD27^-$ proportional data to their $CD62L^+CD27^+$ proportional data to get the proportion of $CD62L^+$ cells post infection as seen in Figure 16. Thus we have two sets of simplified data of the same format.

These data are not compatible with the current method or objective function. We seek to define a new set of models that do not distinguish between $CD27^-$ and $CD27^+$ cells so that we can fit the observed statistics without this information. We change the assumption that there are four underlying cell types: naive, TCMp, TEMp and TEF used in [Buchholz et al., 2013] to three cell types: naive, memory and effector cells with naive and memory having a $CD62L^+$ phenotype and the effector cells having a $CD62L^-$ phenotype (Figure 27).



MODEL INDEX	NAÏVE PHENOTYPE	TYPE 2 PHENOTYPE	TYPE 3 PHENOTYPE
LINEAR MEMORY FIRST	CD62L$^+$	CD62L$^+$	CD62L$^-$
LINEAR EFECTOR FIRST	CD62L$^+$	CD62L$^-$	CD62L$^+$

Figure 27: Description of the two simplified linear differentiation models being fitted to. These models are a simplified version of those described in Figure 21, with the observed phenotypes reduced from three (CD62L$^+$CD27$^+$, CD62L$^+$CD27$^-$, CD62L$^-$CD27$^-$) to two (CD62L$^+$, CD62L$^-$) and cell types being reduced from four to three. This simpler model has four parameters, $\vec{\theta} = (\lambda_N, \lambda_A, \lambda_B, p_{23})$.

This new model is a multi-type Bellman-Harris process that follows a linear unidirectional differentiation path, analogous to the previous one described in Section 5.1.1. The first cell-type that a naive cell differentiates to we denote as cell-type two, the next cell-type in the linear differentiation path is denoted type three. Each cell-type requires a parameter to define its lifetime distribution which we denote as: $\lambda_N, \lambda_2, \lambda_3$. The probability a cell two differentiates upon exiting is denoted p_{23}. Cell-type three is terminally differentiated cell and can only divide. Thus there are four parameters for each model $(\lambda_N, \lambda_2, \lambda_3, p_{23})$.

There are two possible versions of this model: either cell-type two has a CD62L$^+$ phenotype and cell-type three has a CD62L$^-$ phenotype or the reverse is true. These models are specific mathematical versions of the biological models of Effector First or Memory First described in Chapter 3 and we label them as such:

- **Linear Memory First Model** differentiation path: Naive $\rightarrow$ CD62L$^+$ $\rightarrow$ CD62L$^-$
- **Linear Effector First Model** differentiation path: Naive $\rightarrow$ CD62L$^-$ $\rightarrow$ CD62L$^+$

We calculate the models proportional summary statistics using the method described in Chapter 4 but for a three type Bellman-Harris process. The objective functions are equivalent to those defined earlier as WMSE_1 and WMSE_2 (equations 14 and 15), with the only change being the number of proportional statistics fitted to. In the case of [Buchholz et al., 2013] there is one observed total population statistic to map to at day eight ($\hat{x}_1$) and two proportional statistics of each of the six time

points $x_i, (i = 2, ..., 13)$. Thus the objective functions are defined as before but mapping from fewer statistics $\text{WMSE}_1 : \mathbb{R}^{13}_{>0} \times \mathbb{R}^{13} \to \mathbb{R}$ and $\text{WMSE}_2 : \mathbb{R}^{1}_{>0} \times \mathbb{R}^{25} \to \mathbb{R}$. When using this method on [Kinjyo et al., 2015] data we note that there are two sets of four proportional statistics, so we redefine $\text{WMSE}_1 : \mathbb{R}^{9}_{>0} \times \mathbb{R}^{9} \to \mathbb{R}$ and $\text{WMSE}_2 : \mathbb{R}^{1}_{>0} \times \mathbb{R}^{17} \to \mathbb{R}$.

We then fit these two models to the summary statistics described in the papers [Kinjyo et al., 2015] and [Buchholz et al., 2013] and find the parameterisation that returns the minimum objective function value.

5.3.3 Predicting the day of peak immune response

An average mouse has around 100-1000 cells capable of responding to any specific antigen, meaning large adoptive transfers, such as the 10^6 seen in [Kinjyo et al., 2015], dwarf the endogenous response ([Schlub et al., 2010]). Experiments have shown that the time an immune response peaks is inversely related to the number of adoptively transferred cells in OT-1 experiments ([Schlub et al., 2010] and [Badovinac et al., 2007]). While it is difficult to suggest a one size fits all number for the threshold of cells than can be transferred without an abnormal response, [Badovinac et al., 2007] reports that a 5×10^3 transfer still shows signs of abnormality in OT-1 models but 70 cells transferred did not. It is important to know when the peak of the expansion phase is, as the mathematical model we employ does not cover the contraction phase.

The papers [Schlub et al., 2010] and [Badovinac et al., 2007] both report the proportion of OT-1 lymphocytes at different days post infection in experiments using different numbers of transferred cells. We take the time the proportion OT-1 cells is at its maximum to indicate the peak of the immune response. We use log-linear fitting to find approximate the relationship between the number of transferred cells and an the day of peak response (Figure 28). For an experiment transferring 107 cells, as described in [Buchholz et al., 2013], we estimate the peak to be between seven and eight days post infection. For an experiment transferring 10^6 cells, as described in [Kinjyo et al., 2015], we estimate the peak day to be between day four and five.

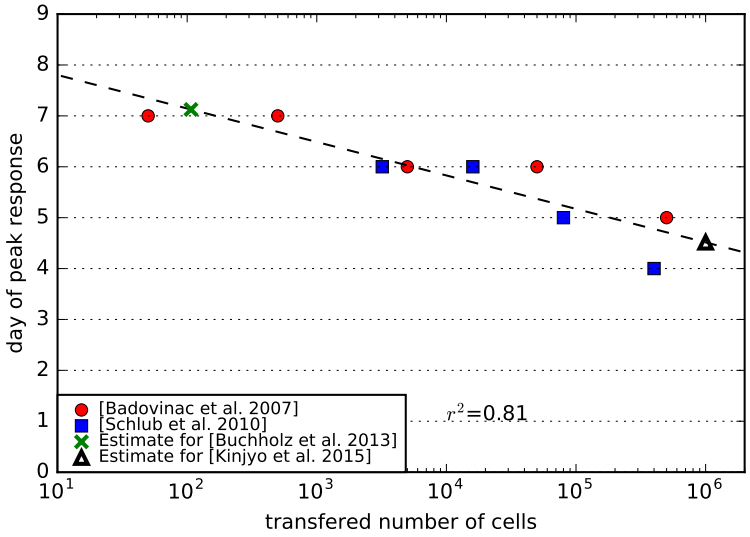


Figure 28: Day of peak response in experiments reported in [Schlub et al., 2010] and [Badovinac et al., 2007] using different numbers of transferred cells. Log-linear regression was used to find an estimate for the day of peak response (y co-ordinate) from the number of adoptively transferred cells (x co-ordinate).

This estimate is based on specific assumptions. The time of peak immune response is taken to mean when the proportion of OT-1 cells is at its maximum, as opposed to absolute cell count. Measurements were not taken at every day post infection and therefore may be subject to censoring. It assumes a log-linear relationship between adoptive transfer numbers and day of peak immune response is a good one. This estimate returns a value for [Buchholz et al., 2013]'s experiment transferring 107 cells closest to day seven, whereas the [Buchholz et al., 2013] model assumes a peak immune response at eight days post infection. Due to the fact the [Buchholz et al., 2013] experiment is under the threshold for number of adoptively transferred cells, and the fact the [Badovinac et al., 2007] experiment did not take samples on day eight or nine, in effect censoring the data, we adopt the value used in [Buchholz et al., 2013] that the expansion phase lasts at least eight days.

5.3.4 Predicting the average family size

The methodology of fitting models to observed cohort data also requires a fit to average family size at peak immune response in order for models to scale correctly, however these data are not reported in [Kinjyo et al., 2015]. Therefore, we need a way to estimate average family size value as a function of the number of cells adoptively transferred. We note that while average family size has been shown to be predictable in the immune response, we do expect individual family sizes to be heterogeneous ([Buchholz et al., 2013],[Gerlach et al., 2013]).

The authors of [Schlub et al., 2010] provide a methodology for calculating the relative difference in the average number of divisions OT-1 transferred cells have taken by the time of peak response between

two experiments (Appendix A). We use a similar method to find the relative number of divisions cells take between the [Kinjyo et al., 2015] experiment and the [Buchholz et al., 2013] experiment, and thus we are able to calculate an estimate for the average number of cells per family in the [Kinjyo et al., 2015] data at the time of the peak immune response.

Let l equal the total number of lymphocytes in a mouse at peak immune response, c be the total number of OT-1 cells transferred (the number of OT-1 families assuming no death), a be the size of an average OT-1 family at peak response and o denote the proportion of OT-1 cells out of all lymphocytes at peak response, then

$$o = \frac{ac}{l}. \tag{16}$$

We would like an estimate for a. While c is known, o and l are not and we would like to derive an estimate for them first. To this end we plot a log-linear scatter plot of the number of transferred cells and the percentage of OT-1 lymphocytes at the peak of the immune response as reported in the data from [Schlub et al., 2010] and [Badovinac et al., 2007]

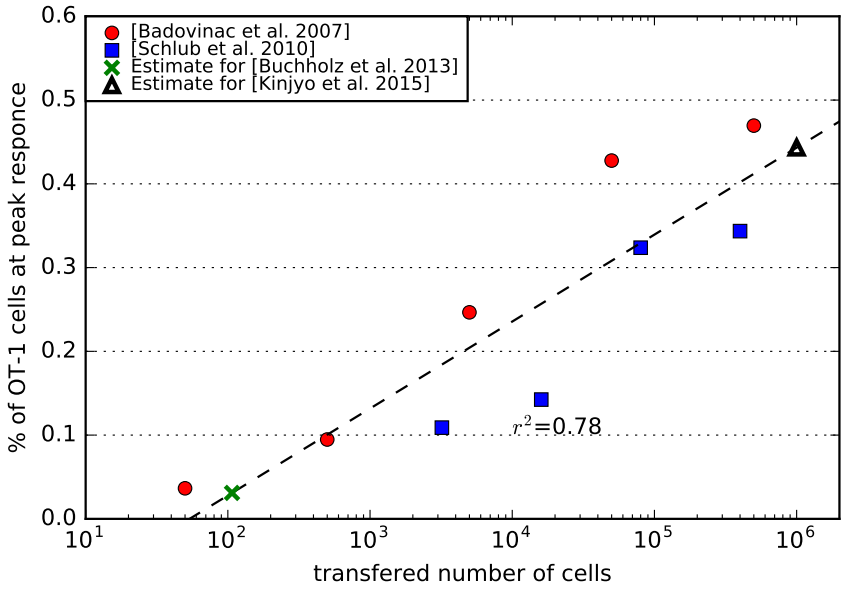


Figure 29: Proportion of OT-1 cells at peak response in experiments reported in [Schlub et al., 2010] and [Badovinac et al., 2007] using different numbers of transferred cells. Log-linear regression was used to estimate the percentage of OT-1 cells out of all lymphocytes at the peak of the immune response. This estimate was used to determine the percentage of OT-1 cells at peak response (y co-ordinate) from the number of adoptively transferred cells (x co-ordinate). We see the numbers of OT-1 cells do not grow relative to the size of the adoptive transfer, thus average family size is shrinking for higher transfers. The fact the log-linear line crosses the x-axis, shows this estimate cannot be used for low transfer numbers, and this has a possible the explanation that low numbers of adoptively transferred cells do not scale in the same way as high ones.

This log-linear fit gives us the equations

$$o = 0.10\log_{10}(c) - 0.18. \tag{17}$$

For the [Buchholz et al., 2013]'s adoptive transfer of 107 cells, we can calculate the expected percentage of OT-1 cells using Equation 17 gives us the value 0.031 or 3.1%. Using this information along with the average family size of OT-1 T cells at the peak of the immune response reported by that paper, we use Equation 16 to calculate an estimate the total number of lymphocytes in a mouse in the [Buchholz et al., 2013] experiment at peak infection,

$$l = \frac{107 \times 15500}{0.031} = 5.3 \times 10^7. \tag{18}$$

We make the assumption that the total number lymphocytes at the peak immune response is the same across all experiments, an assumption supported by spleen data in [Schlub et al., 2010]. However [Badovinac et al., 2007] has spleen counts that suggest this may not always be the case, particularly for low transfer numbers. Based on this assumption we get

$$a = \frac{o \times 5.3 \times 10^7}{c} \tag{19}$$

$$= \frac{(.10 \times \log_{10}(c) - 0.18) \times 5.3 \times 10^7}{c} \tag{20}$$

$$= \frac{5.6 \times 10^6 \times \log_{10}(c) - 9.6 \times 10^6}{c}. \tag{21}$$

Thus we have a estimate of the average size of a cell family at the time of peak immune response from the number of cells adoptively transferred. We can use this to calculate estimates of different average family sizes as reported in Figure 31 and Table 6. This estimate is only useful for adoptive transfer of 100 cells or higher due to the estimate for OT-1 proportions against transfer numbers crossing the x axis, as seen in Figure 29. We expected the method not to be appropriate for these small numbers of transfers and this could be interpreted biologically; [Badovinac et al., 2007] reported the effect of adoptive for 70 OT-1 cells less did not cause the average family size to shrink as did large transfers.

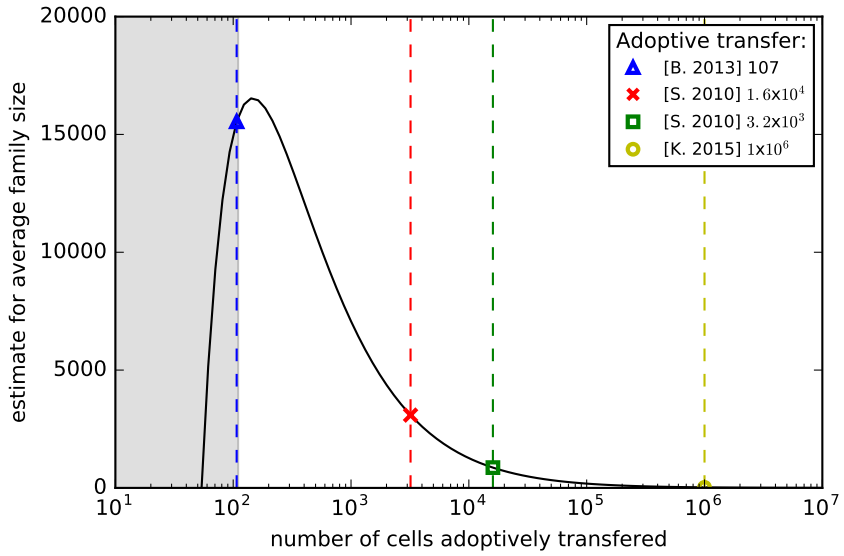


Figure 30: For a given number of Ot-1 transferred cells, an estimate of their average family size at peak immune response. For low numbers of transferred cells (<100) the estimate is not good.

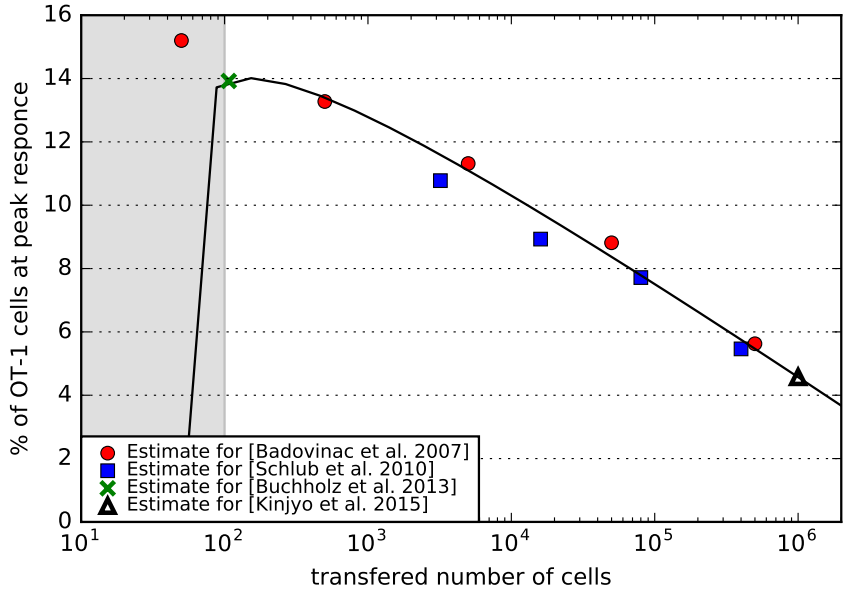


Figure 31: Estimated average number of division per cell at time of peak against adoptive transfer size. Dots show an estimate based on percentage of OT-1 cells observed at a peak assuming there are 5.3×10^7 lymphocytes in total, the black line shows the estimate using equation 21. For low numbers of transferred cells (<100) the estimate is not good.

Paper	**c** (obs)	**l** (est)	**o** (obs)	**o** (est)	**a** (est)
[Buchholz et al., 2013]	1.0×10^2	5.3×10^7	n/a	.03	1.5×10^4
[Buchholz et al., 2013]	107	5.3×10^7	n/a	.03	1.5×10^4 (obs)
[Schlub et al., 2010]	3.2×10^3	5.3×10^7	.11	.18	3.0×10^4
[Schlub et al., 2010]	1.6×10^4	5.3×10^7	.15	.26	8.4×10^3
[Schlub et al., 2010]	8.0×10^4	5.3×10^7	.32	.33	210
[Schlub et al., 2010]	4.0×10^5	5.3×10^7	.34	.40	53
[Kinjyo et al., 2015]	1.0×10^6	5.3×10^7	n/a	.44	23

Table 6: Observation and estimates of values in Equation 16 [Buchholz et al., 2013], [Schlub et al., 2010], [Kinjyo et al., 2015]. c is the total number of cells transferred, l is the total number of lymphocytes at the peak immune response, o is the proportion of all cells at peak immune response, a is the average family size at peak immune response. est indicates an estimation and obs indicates an observation.

This estimate is equivalent to saying [Kinjyo et al., 2015] cells, for a 1×10^6 size transfer, on average went through 9 divisions less than that of [Buchholz et al., 2013], with a 1×10^2 cell transfer (Figure 30). This roughly ties to the conclusion from [Schlub et al., 2010] that states that cells from a 4×10^5 transfer go through approximately five less divisions on average than a 3.2×10^3 transfer. From this, and the fact that [Buchholz et al., 2013] reports a 15k (13.8 average divisions) average family size at peak, we then calculate that in the [Kinjyo et al., 2015] experiment we would expect 23 cells per family (four to five average divisions) at the peak of around day 4.5. This also is reasonable when compared to [Badovinac et al., 2007]'s data, which reported that there is only a 13 fold increase in peak frequency for a 10,000 fold increase in precursor transfer.

While relative numbers are comparable, large discrepancies in reported data for absolute cell counts highlight caveats with these estimates. [Buchholz et al., 2013] reports a value of 15k for the average peak family size (107 cells transfer) while [Badovinac et al., 2007] giving estimates of up to 400k family size estimate for 50 cells transferred. Our estimate for the total number of lymphocyte cells in a mouse (calculated from the [Buchholz et al., 2013] number) is lower than that [Schlub et al., 2010] reported in the spleen by about four fold, and [Masopust et al., 2001] reports that total lymphocytes should be at least double that in the spleen making eight fold difference between estimates. However for high estimates comparisons are better, with [Badovinac et al., 2007] estimating a 40-400 peak average family size for when transferring 5×10^5 cells compared to our estimate of 43. Because of the fragility of this estimate, we do a sensitivity test in Appendix B, which concludes that while average family size does indeed effect the fit, it does not when limited the analysis to only biologically plausible models.

There are three main assumptions to this method: that the number of lymphocytes at the time of

the immune response remains constant; the linear fit displayed in Figure 29 is a good estimate and that our calculation estimating the day of peak of the immune response is a good one. For the first of these assumptions we use evidence from [Schlub et al., 2010] and a similar assumption to theirs, that the observed total lymphocytes counts in the spleen at peak immune response do not vary more that 6%. The log-linear assumption is used for simplicity, and noticeably predicts negative growth for low numbers of OT-1 cells and so can only be useful above a threshold. We note the conflicting data on the average family size make it difficult to reconcile all these problems of the estimate.

5.3.5 Results of adaptation to other papers data

$CD62L^+$ T cell proportions reported in different adoptive transfer experiments show seemingly conflicting trends up to day eight post infection (Figure 32). The data reported in [Buchholz et al., 2013] shows a continual decrease in the proportion of $CD62L^+$ cells up to day eight in the spleen, which is consistent with their conclusion of effector cells out proliferating long lived memory cells late in the expansion phase. The data reported in [Kinjyo et al., 2015] shows an initial decrease in proportions of $CD62L^+$, then a large increase before day seven post infection, consistent with their conclusion of memory appearing later in the differentiation pathway. As seen in Section 5.3.3, a large number of naive cells adoptively transferred can shorten the expansion phase, meaning some of data presented in [Kinjyo et al., 2015] prior to day eight is likely to be after the expansion phase and therefore a comparison should adjust for this. When the [Kinjyo et al., 2015] spleen data are curtailed to only show the expansion phase, using the estimated peak of the immune response described in Section 5.3.3, the trends seen in papers [Buchholz et al., 2013] and [Kinjyo et al., 2015] are consistent, showing percentage of $CD62L^+$ cells to only decrease with time (Figure 32). This is confirmed by the [Schlub et al., 2010] data, where data from experiments with high numbers of adoptively transferred cells show an upturn in the proportion of $CD62L^+$ cells before day eight, but this effect is absent if the data are appropriately curtailed. Data from [Badovinac et al., 2007] follows this trend up to a point, but does show an increase in the proportion of $CD62L^+$ even after curtailing, though it is lessened. This was also true for [Kinjyo et al., 2015] MLN data.

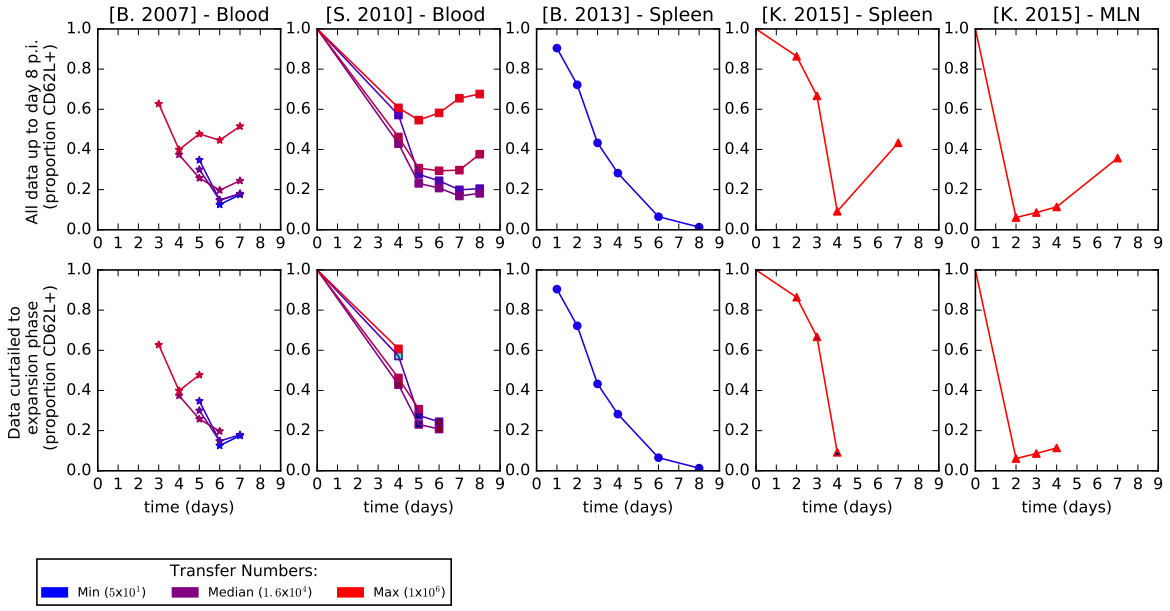


Figure 32: Percentage of CD62L$^+$ cells post infection reported in papers [Badovinac et al., 2007], [Schlub et al., 2010] [Buchholz et al., 2013] and [Kinjyo et al., 2015]. The top row shows all data reported up to day eight post infection, the bottom row shows the data only during the experiments expansion phase. In the top row, experiments with high numbers of adoptively transferred cells show a late increase in the proportion of CD62L$^+$ cells, but when limiting the data to only the expansion phase, as in the bottom row, the proportion of CD62L$^+$ decreases only, with the exception of some data from [Badovinac et al., 2007] and MLN data from [Kinjyo et al., 2015].

When fitting models to the [Buchholz et al., 2013] spleen cohort data, the Memory First Model was marginally the best fit (Figure 33). When fitting to the [Kinjyo et al., 2015] spleen data up to day eight post infection, the method found a Linear Effector First Model to fit best (Figure 33). The Linear Effector First Model was also the best fit for experiments with high adoptive transfer numbers from the blood data reported in [Schlub et al., 2010] (not curtailing data). Thus, when fitting without adjusting for the number of cell adoptively transferred, we find contradicting results. For further breakdown of the fit when not curtailing the data see Appendix C.3.

When we restrict the data from these three papers to the expansion phase, all data sets better fitted the Linear Memory First Model, though some marginally (Figure 33). Some curtailed cohort data (CD62L$^+$ cell proportions harvested at different times post infection and estimated average family size) from [Schlub et al., 2010] had less than four data points and so were not fitted to $(8\times10^4, 4\times10^5)$. While the results of the curtailed data consistently supported the Linear Memory First Model, the strength of some results was low, particularly the result for [Kinjyo et al., 2015]'s curtailed data (Figure 33), where both models could be considered a good fit to the data. This is due to the fact that fewer data points meant models found fitting easier and both models give good visual fits to the data (Figure 34).

The parameterisation of the best fitting models when fitting to [Kinjyo et al., 2015] data showed little memory appearing during the expansion phase: less than one cell per family (Figure 34 and 35). For both the Memory First and Effector First Models, the curve of $CD62L^+$ portions was primarily made up for by the proportion of naive and effector cells only. This could suggest that the model is not sophisticated enough or did not have enough data to fit to. An alternative explanation could come from other [Kinjyo et al., 2015] results (Section 3.3), which showed a homogeneous population effectors on day four post infection make way for two homogeneous groups of effectors and memory cells on day seven. With the insight that day seven is after the expansion phase under the experimental setup described in [Kinjyo et al., 2015], this result suggests the possibility that memory may appear in the contraction phase. However, memory appearing in the expansion phase is an a core assumption of the model, and this alternative is outside of the scope it can evaluate.

While models showed a good fit to the curtailed data, all best fits had biologically implausible parameterisations, with naive cell average lifetimes being over 60 days, the only exception being the Linear Memory First Model fitted to [Buchholz et al., 2013] data (Figure 35). Thus, for all other models, the majority of naive cells were not activated at the end of the expansion phase and the increase population was driven by a very small number of family's that had activated and rapidly expanded (Figure 34). This is at odds with results showing that, even for high adoptive transfers, less than 0.1% of the OT-1 population comprised of unrecruited cells ([Schlub et al., 2010]) especially since average family sizes are small for large transfers. In addition, data from [Kinjyo et al., 2015] showed that no cells had divided on day two post infection yet the majority were preparing to divide, suggesting the majority were recruited already. Other studies have shown that while high doses of pathogens transferred lead to almost universal recruitment of nave cells, low numbers of transferred pathogen can lead to only 30% or 50% recruitment ([Kaech and Ahmed, 2001]). There are reasons why these experiments are not entirely comparable to the experiment we have looked at: the use of the P14 experimental model would have lower affinity than an OT-1 model, the difference in pathogen (Flu vs LM, [Kinjyo et al., 2015]) may affect results, as could the metric used (pfu vs cfu). Nevertheless, even in the low pathogen transfer cases seen in [Kaech and Ahmed, 2001], recruitment was shown to be significantly higher than that seen in our parametrisations.

To see how the model ranking was affected by removing the possibility of the biologically implausible number of non-activated naive cells, we redid the analysis, but upper bounding parameters so that average cell lifetime is five days or less. When fitting with this method further adjusted, the expansion phase the Linear Memory First Model was still was the best fit to all the curtailed data sets, but the result was stronger, with the Linear Effector First Model failing to give a good fit for [Buchholz et al., 2013] and [Kinjyo et al., 2015] data (Figure 36). In addition the parameterisations of the Linear Memory First Model showed memory cells as longer living than effector cells in all cases as expected, but this was not always true in the previous fits. This result showed us that under the assumptions of the model, there was no biologically plausible good fit for the Linear Effector First Model when fitting

against [Buchholz et al., 2013] or [Kinjyo et al., 2015] spleen data. For further breakdown of applying the method with these extra parameter bounds refer to Appendix C.4.

Further CD62L blood data reported in [Badovinac et al., 2007] was also available to fit to (those data sets that had at least three data points were for experiments with $5 \times 10^2, 5 \times 10^4$ and 5×10^5 cells transferred) and the MLN data reported in [Kinjyo et al., 2015] (1×10^6) also had enough data points to be able to fit to. This fit was done curtailing the data to the expansion phase and limiting the parameterisation of the average cell lifetime to five days. The results (Figure 37) show all [Badovinac et al., 2007] blood data continued to support a Linear Memory First Model. However some of these fits had strange parameterisations with memory cells having shorter lifetimes than effector cells (Appendix D). MLN data fitted best to a Linear Effector First Model, however these data are not appropriate to fit to, due to the high migration of T cells to and from the lymph. For further analysis of these data, the parameterisation and statistics, see Appendix D.

Thus when limiting blood and spleen data to the expansion phase, and limiting the parameterisation of models to only biologically plausible cases, under the assumptions the the Linear Memory First Model is the only viable option of the two models.

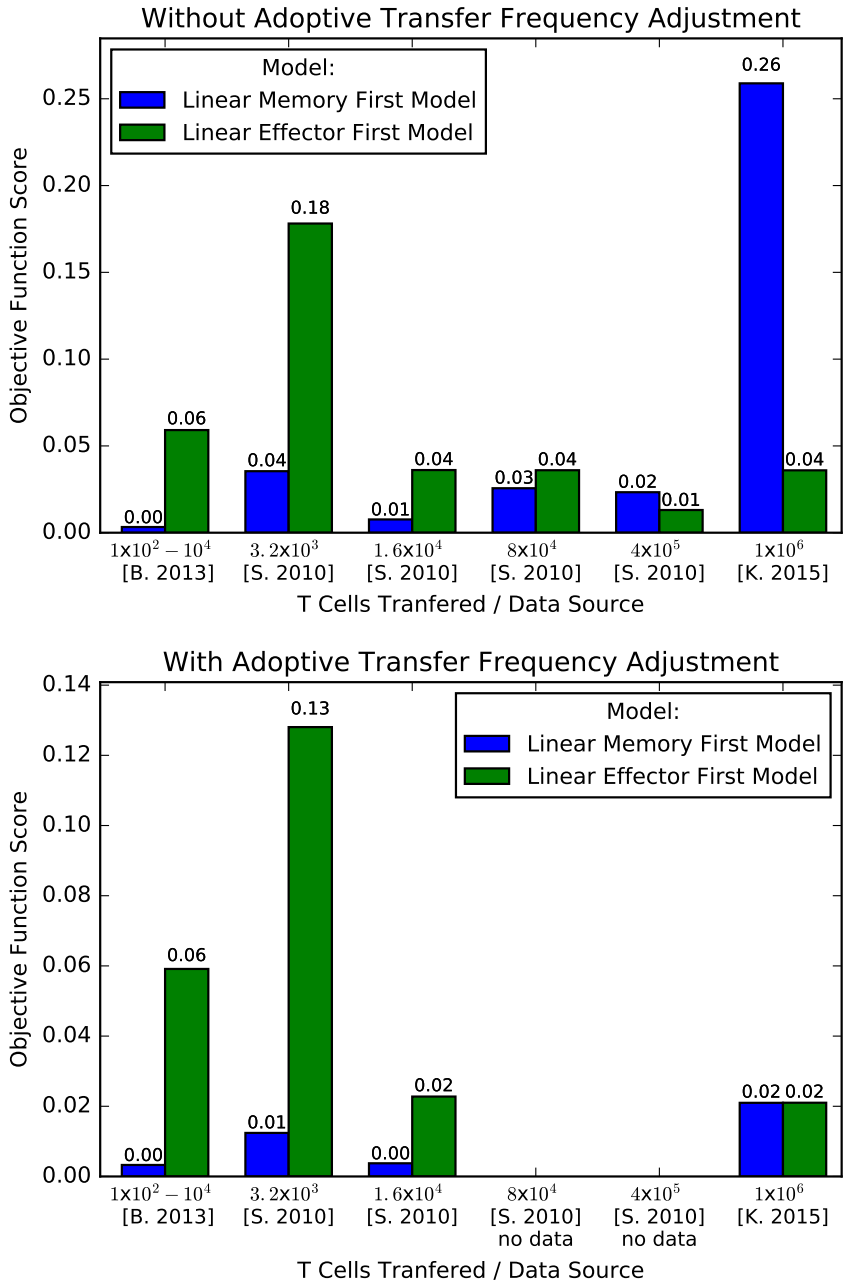


Figure 33: The objective function value for the best fitting parameterisations of the two simplified linear models when fitting to observed cohort statistics from different data sources without assuming a shorted expansion phase (top) and with assuming a shortened expansion phase (bottom). Smaller values indicate a better fit. [Buchholz et al., 2013] and [Kinjyo et al., 2015] data from the spleen, [Schlub et al., 2010] data from the blood. After the data were curtailed, two sets of [Schlub et al., 2010] data had less than four data points (including average family size) and thus were inappropriate for fitting. When fitting to data up to day eight [Buchholz et al., 2013] fits best to a Linear Memory First Model and [Kinjyo et al., 2015] to a Linear Effector First Model. When the shortened expansion phase is accounted for, all models fit best to a Linear Memory First Model. In this second case, the difference between the best fitting models for [Kinjyo et al., 2015] is marginal.

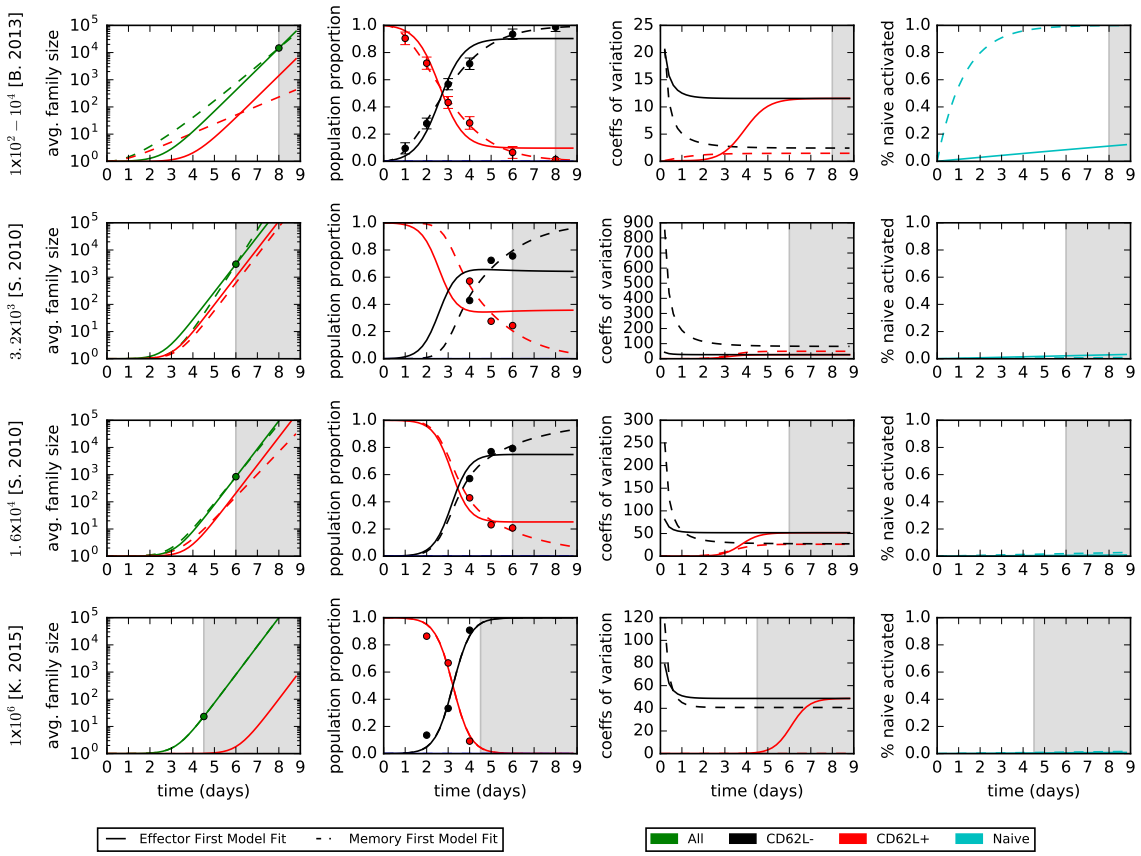


Figure 34: Statistics for the best fitting models when fitting to different sets of cohort data from papers [Buchholz et al., 2013], [Kinjyo et al., 2015] and [Schlub et al., 2010]. Dashed lines show the Linear Memory First Model fits, solid lines show the Linear Effector First Model fits, dots show the statistics from experimental data fitted to. The grey area shows the time after the expansion phase has ended. Visual fits are similar between the two models when looking at proportions of CD62L$^+$ cells during the expansion phase when fitting to [Kinjyo et al., 2015] reported data. Only the [Buchholz et al., 2013] data for the Linear Memory First Model shows the majority of naive cell activated by the end of the expansion phase.

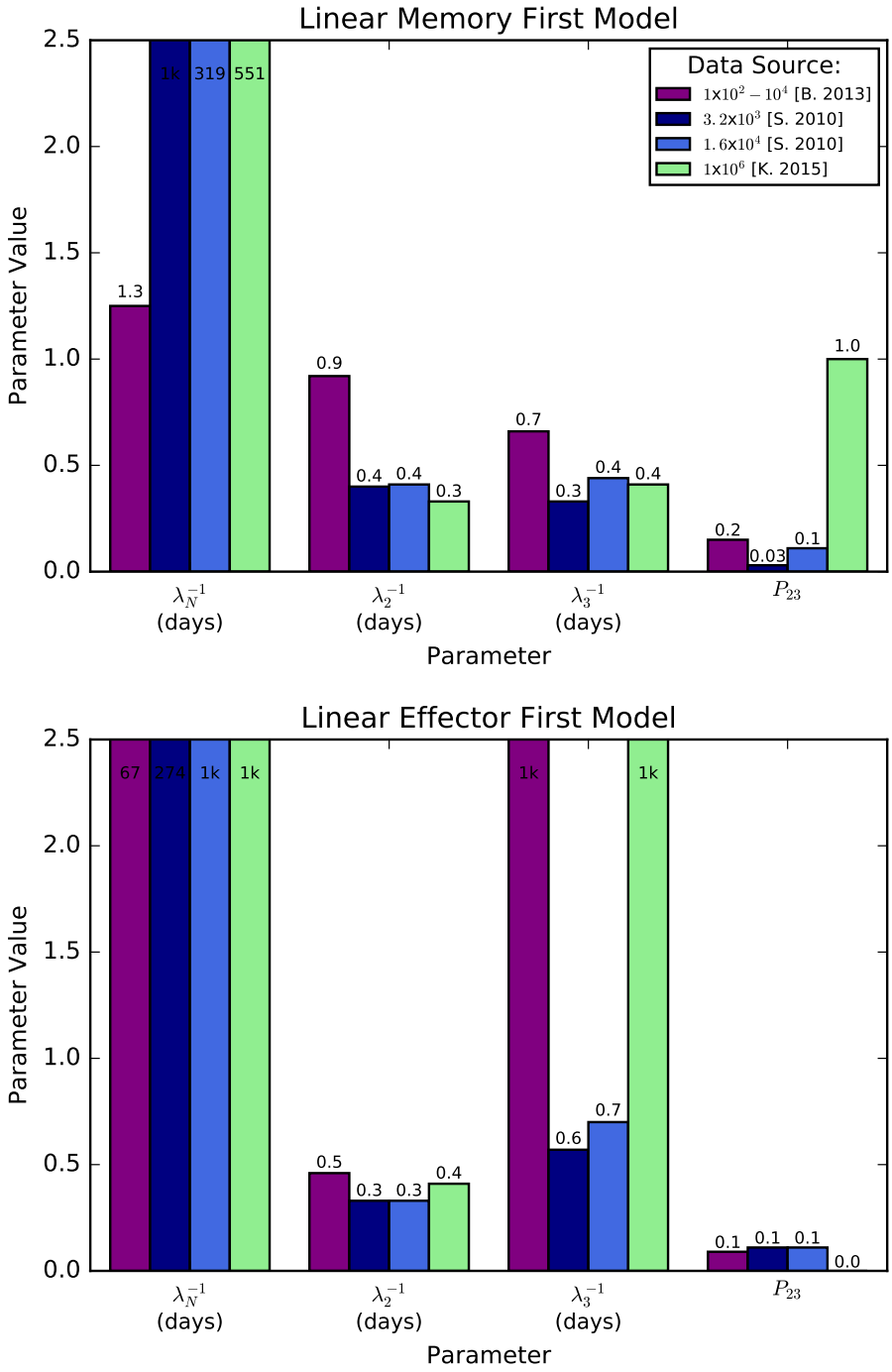


Figure 35: Parametrization for the best fitting models when fitting to different sets of cohort data from papers [Buchholz et al., 2013], [Kinjyo et al., 2015] and [Schlub et al., 2010]. Parameters are as described in Figure 27. Only the memory first model fitted to [Buchholz et al., 2013] data gives a biologically plausible result, the other models parametrise the naive cells average time to activate as 60 days or more, which is not consistent with data indicating substantial recruitment of T cells with cognate receptor ([Schlub et al., 2010],[Kinjyo et al., 2015]). In [Kinjyo et al., 2015]'s case, for the Linear Memory First Model the effector cells live longer average lifetimes than memory cells.

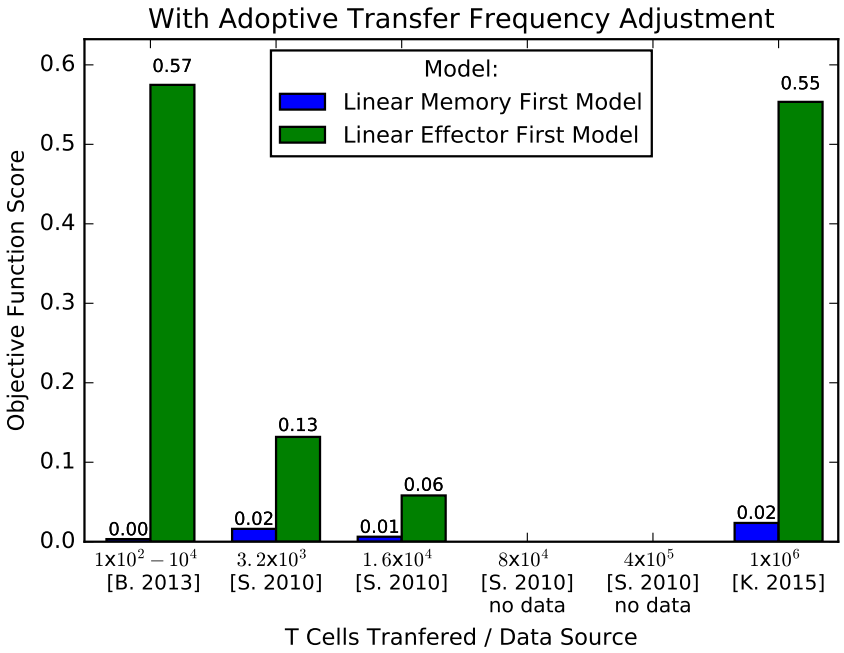


Figure 36: The objective function values for the best fitting parameterisations of the two simplified linear models when fitting to observed proportional phenotype statistics from different data sources and assuming a shortened expansion phase. Smaller values indicate a better fit. As in Figure 33 but upper bounding average cell lifetime to five days.[Buchholz et al., 2013] and [Kinjyo et al., 2015] data from the spleen, [Schlub et al., 2010] data from the blood. This shows strong results in favour of a Linear Memory First Model. After the data were curtailed, two sets of [Schlub et al., 2010] data had less than four data points (including average family size) and thus were inappropriate for fitting.

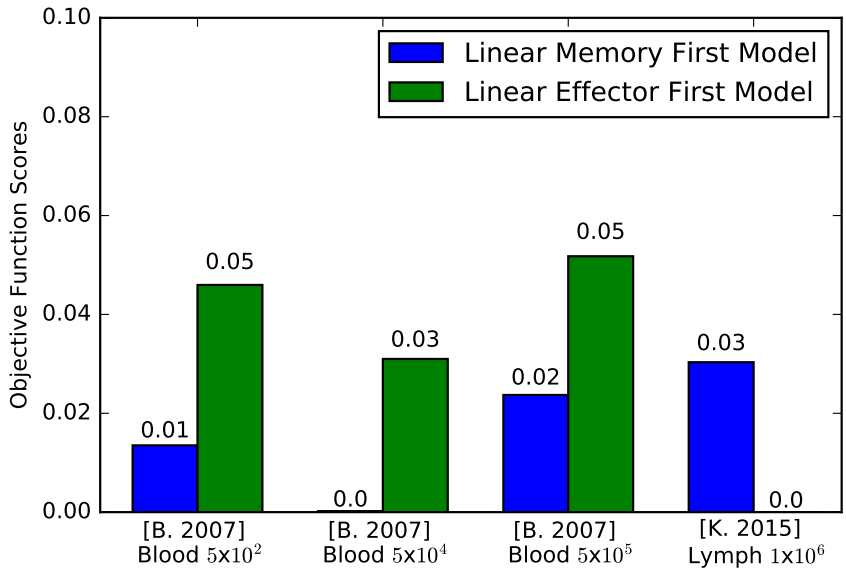


Figure 37: The objective function values for the best fitting parameterisations of the two simplified linear models when fitting to observed proportional phenotype statistics from different data sources and assuming a shortened expansion phase. Smaller values indicate a better fit. As in Figure 33 but looking at [Badovinac et al., 2007] blood data and [Kinjyo et al., 2015] MLN data and upper bounded average cell lifetime to five days. [Badovinac et al., 2007] blood data fitted best to a Linear Effector First Model, however [Kinjyo et al., 2015] MLN data fitted best to a Linear Memory First Model, at odds with other blood and spleen data.

6 Discussion

In this section we review and assess some of the assumptions necessary to the model. We suggest further study to explore the problems tackled outside of these assumptions. Finally we draw together conclusions, both from within the framework of the model and hypothetical ones.

6.1 Caveats of results and assumptions of the method

We reconsider model assumptions based on evidence of their likely truth and impact to core results. Some of the assumptions we inherit from the methodology described in [Buchholz et al., 2013] and others we found necessary as part of the adaptation of the mathematical model in [Buchholz et al., 2013] to make it suitable for the more widely variable cohort data ([Kinjyo et al., 2015], [Schlub et al., 2010]). Quantitative models often require strong assumptions to be mathematically well defined. Assumptions were needed to reduce complexity and model parameters to avoid over-fitting to limited data. Assumptions also come from difficulties in performing analysis across many different papers and distinct data sets. In some cases it is difficult to assess the assumptions truth due to lack of evidence.

T cells' lifetimes are exponentially distributed. Experimental *in vitro* evidence has shown that there is a minimum time to divide and that cell lifetime distributions are closer to a log normal distribution then an exponential one ([Dowling et al., 2014], [Dowling and Hodgkin, 2009]). This was confirmed by the data from [Kinjyo et al., 2015]. However without an exponential distribution the solution to the non-linear Volterra integral equation for the generating function of the population size in a Bellman-Harris process, as in Equation 10, is non explicit and can be expressed in asymptotic approximations. This leaves open further work of fitting the same models but with more realistic lifetime distributions that are not memoryless.

Assumptions of independence and identically distribution for cell lifetime and/or offspring distributions. There is evidence that cell lifetime between sisters is correlated and however lifetimes between cousins and mother-daughter lifetimes are not ([Kinjyo et al., 2015], [Hawkins et al., 2009], [Markham et al., 2010], [Duffy et al., 2012], [Dowling et al., 2014]). There are many alternative assumptions to distributions being i.i.d., for example distributions could change with time or activation theory could mean clonal dependence ([Marchingo et al., 2014]).

Families of T cells behave independently. There are biological factors that could mean information on one T cell families proliferation behaviour contains information on another within the same organism. For example T cells use cytokines to communicate and it is possible environmental factors, such as body temperature, affect T cell behaviour. The result that large adoptive transfers affect T cell behaviour at a population level show that in some circumstances families are dependent.

Cell types are discrete and there are three-four of them. The traditional view, or accepted approximation in the literature is that $CD8^+$ T cells are compartmentalised into four sub-types: naive,

effector, central memory and effector memory ([Sallusto et al., 1999], [Wherry et al., 2003], [Baars et al., 2005]). To create a continuous model or assign new subtypes would be an undertaking outside the scope of the present analysis, but would be of interest. The revised model we used employed an even more simplified partitioning to three cell types: naive, memory and effector. This assumes effector and effector memory cells have similar proliferation behaviour. We do this because of the difficulties in building a more complex model on simple data and due to lack of data overlap when spanning different papers.

Differentiation is not linked to division. Whether cells adopt cell types at the point they divide or at some point mid their lifespan is difficult to determine, though there has been some evidence supporting asymmetry theory ([Chang et al., 2007]). The assumption of the model described in [Schlub et al., 2010] is that differentiation occurs at the point of division. The model described in [Buchholz et al., 2013] took division and differentiation as two different events. It is possible to redefine the mathematical model to have differentiation linked to differentiation. This would also allow us to explore of the asymmetric hypothesis. We note that [Buchholz et al., 2013] reports some modelling around asymmetric division and came to the same conclusion of memory appearing first.

No death during the expansion phase. The model used assumed no cell dies without out offspring during the expansion phase. In general cell death is hard to measure, and particularly in adoptive transfer because of difficulty in separating death during the expansion phase from death due to the trauma of transfer. Some papers have shown evidence of cell death during the expansion phase ([White et al., 2015]). In future study, one could extend models to include a death parameter and use AICc or otherwise to see if this parameter yields a much stronger fit.

Only linear paths considered. While [Buchholz et al., 2013] looked at many different network structures, we only considered linear ones. This adapted model did include one of the two best fitting parameterisations [Buchholz et al., 2013] found. It is possible that had non linear paths been in scope these may have been the best fit to the data. As a future study it would be possible to extend the model to these non linear cases.

Choice of objective function. The choice of objective function is subjective and depends on different values of what it means for a model to be close to observed data. We tested the method under a separate objective function and found the same result (Appendix C.2). Therefore the risk is considered lower but there is always a possibility of another sensible objective function yielding a different result.

Naive cells are CD27$^+$. We assumed in that naive cells are CD27$^+$. This assumption was only necessary in the analysis of the cohort [Buchholz et al., 2013] data under the original model. While reports agree there is some expression of CD27 on naive cells, data varies on whether CD27 in naive cells is universally high or low compared to effectors (Hendriks et al. [2000], Baars et al. [2005], [Kinjyo et al., 2015]). However the impact of this assumption being changed would lead to the same result of

a memory first model fitting to the proportional data described in [Buchholz et al., 2013] data, but more definitively, as seen in Appendix C.1.

The method for assuming day of peak immune response from numbers of adoptive transferred cells is accurate. This method is described more fully in Section 5.3.3 and relies on using a log-linear correlation to predict the day of peak immune response. As we used data from two different experiments [Schlub et al., 2010] and [Badovinac et al., 2007] and found a good correlation between them this gives us confidence that this is a reasonable estimator, but some margin of error is possible. Some studies have also suggested that peak response time may be different between OT-1 cells and endogenous cells ([White et al., 2015]).

The method for estimating average family sizes at peak immune response from adoptive transfer numbers is reliable. This method is described in Section 5.3.4. This method relies on the assumptions that adoptive transfer numbers have a log linear correlation to OT-1 proportions at peak of immune response and total T cells counts at the peak of an immune response are constant between experiments. The first assumption was supported by a good correlation when comparing to two sets of data. The second of these assumptions was evaluated and was shown to have evidence in its favour in the [Schlub et al., 2010] paper, however some spleen data from [Badovinac et al., 2007] may contradict this trend. Sensitivity analysis on the affects of assuming alternative average family sizes is done in Appendix B.

Observations in specific parts of the body represent the whole. Where cells are harvested from could have significant effect on the profile of T cells observed, both because of environmental factors or because T cells are migrating around the body. CD62L itself is a homing receptor for the lymph and so this may significantly effect results. [Buchholz et al., 2013] does a test that showed numbers of OT-1 cells harvested from the spleen correlated to that of the spleen lung and lymph node combined. In addition this test does not report phenotype correlation between the spleen and the rest of the body. [Kinjyo et al., 2015] reports T cells in the lymph node proliferating faster than those in the lungs and the spleen and that the MLN showed a much faster decrease in proportion of $CD62L^+$ cells than the spleen. Therefore this is a significant assumption in our analysis, but one difficult to overcome without a further experiments.

The OT-1 model is representative of a natural immune response. All experiments observe the immune response of OT-1 cells to OVA antigen and it is assumed that these are representative of a endogenous immune response. OT-1 cells react with a very high affinity to OVA and so therefore may represent a stronger response than some endogenous responses. This result was confirmed by the experiment described in the [Badovinac et al., 2007] paper using a P14 model.

CD62L is a good indicator of memory cell-type. CD62L correlating with other memory characteristics is widely accepted in literature and sometimes CD62L even used in the definition of memory itself ([Ahmed et al., 2009]). [Kinjyo et al., 2015] contains experiments showing how CD62L

correlates to slow proliferation, [Buchholz et al., 2013] shows how it negatively correlates to effector markings (KLRG1) and both [Badovinac et al., 2007] and [Schlub et al., 2010] show how it up regulates during the death phase and effectors die off leave behind long lineage memory cells. However, it should be noted that all these remain as correlations and so using one single marker yields a unknown amount of false categorisation. Also, most experiments looking into memory and CD62L phenotype correlation observe it late in the immune response. During the early days post infection (days one to three), when the response is harder to measure *in vivo*, $CD62L^+$ may include some naive cells or activated cells that may be late down regulating their CD62L. it is worth noting $CD62L^-$ subgroup also contains the effector memory subgroup which has some memory characteristics. While using a more in depth method to define cell-type from multiple phenotypic traits seems ideal, this poses increasing difficulties when trying to compare data across papers.

CD27 is a good indicator of effector and effector memory [Buchholz et al., 2013] used CD27 to distinguish effector cell types from effector memory cell types. [Buchholz et al., 2013] shows $CD27^-$ correlating with effector functionality and [Kinjyo et al., 2015] shows it correlating with higher cell division speed. Other papers confirm this and correlate it to CD127 another marker of effector functionality ([Sallusto et al., 1999], [Hendriks et al., 2000], [Baars et al., 2005], [Hikono et al., 2007]).

Experiments are comparable (e.g. the difference in gating strategy or pathogen used are negligible). Both [Buchholz et al., 2013] and [Kinjyo et al., 2015] tested the effects of using different pathogens or different quantities of antigen they found that this change did not strongly effected results. Gating strategies are difficult to compare because they rely on relative setting of individual flow cytometry machines.

All cell types appear in the expansion phase. The model we used assumed that memory appeared during the expansion phase. One support of this assumption is that [Buchholz et al., 2013] data showed all phenotypes had appeared by day eight post infection, though this could be explained by gating. To counter this, [Kinjyo et al., 2015] showed T cell population was made up of a homogeneous group of effector like cells up to the end of their expansion phase and then two heterogeneous groups of cells after the expansion phase, one showing a memory phenotype and one showing a effector phenotype. In addition the parameterisation of the best fitting model to [Kinjyo et al., 2015]'s data showed that the proportion of $CD62L^+$ cells could be explained by naive and effectors cells alone. This leaves the possibility that this assumption is untrue.

Too few data points to fit to. For some data sets the model had few points to fit to. We worked under the convention that the number of data points must equal or exceed the number of parameters in the model. This included average family size which itself was an estimate. While [Buchholz et al., 2013] had seven data points, [Kinjyo et al., 2015] and [Schlub et al., 2010] only had the minimum of four. There was also a similar issue with the original method described in [Buchholz et al., 2013] where the most complex networks had up to 12 parameters and had to fit nine data points.

Human error in reading graphs and figures. The majority of data fitted to was taken visually from graphs in papers and therefore could be subject to human error.

Difficulties in determining significance of results. There were many instances where we were unable to determine significance. This was for two reasons: the data from papers often did not contain enough information to calculate margin of errors; the mathematics of the model made computing the probability density function difficult.

[Buchholz et al., 2013] used two different numbers in their cohort adoptive transfer experiment: 1×10^2(six, eight days p.i.) and 1×10^4 (one-four days p.i.). We simplified this to say that the expansion phase continued up to day eight as was the assumption in the [Buchholz et al., 2013] paper itself.

6.2 Further work

There are opportunities to investigate the method of fitting the model described in [Buchholz et al., 2013] to different data sets outside some of the assumptions listed above. The analysis in this thesis was limited to linear unidirectional networks and it would be possible to expand this to all non-linear networks as [Buchholz et al., 2013] did in the original method. Log-Normal cell lifetimes could be assumed in which case the equations for the generation function of the population of a Bellman-Harris process could be solved analytically. Sisters could be made to correlate in the model (e.g. [Duffy and Subramanian, 2009]). The possibility of cell death could be included. Division could be linked to differentiation and allow the possibility of asymmetric division.

Fitting to summary statistics of clonal data does not use all data extracted from experiments. Theoretically it would be beneficial to fit mathematical models to distributions of clonal size. However, there are obstacles to this approach, one is the computations of likelihood functions for multi-type Bellman Harris processes are extensive and often approximations or assumptions are required ([Chen and Hyrien, 2011]) which is likely why the papers themselves do not use this method. Another issue with this method is that it is not possible to apply the method to experiments gathering evidence at a cohort level such as [Kinjyo et al., 2015], [Badovinac et al., 2007] or [Schlub et al., 2010]. However, opportunities to do some simplified version of distribution fitting in the future could be considered.

While we suggest the number cells adoptively transferred to be the main cause of differences in results between [Kinjyo et al., 2015] and [Buchholz et al., 2013], it is worth considering alternate hypotheses that we did not look at, but could be considered in future work ([Chamberlin, 1890]). For example, the type of infection could play a role, as both experiments use different kinds of pathogens. The flu used by [Kinjyo et al., 2015] is localized and less inflammatory when compared to the Listeria monocytogenes used in [Buchholz et al., 2013]. However, investigation into this possibility would require further experimental setups.

Other possibilities for future work could include fitting a completely different mathematical model. For example a mathematical model based around activation theory, such as the one describing T cell populations in [Deenick et al., 2003]. In these models the behaviour of families is deterministic after activation with heterogeneity of families arising from activation signalling or an internal random variable decided at the start of the cells first division. So in this case random variables are assigned to whole families rather than individual cells, meaning for instance, every single cell in a family has the same time of division. This approach is supported by other recent papers ([Marchingo et al., 2016],[Heinzel et al., 2017])

While the differences in most fits are seen to be large visually, it would be useful to test if they are different in a statistical sense, particularly for more borderline results. This testing of statistical significance is desirable and left for future work.

The method described in this thesis fits mathematical models to the expansion phase, however there are other papers that fit mathematical models to both the expansion and contraction phase ([Antia et al., 2005], [Ganusov, 2007]). The models these papers employed are different in a number of key ways: they are fitted total average family size, not phenotype, and use a different parametrisation, for example giving a decay parameter for the rate at which cells die. The paper [Ganusov, 2007] found the Linear Effector First Model to give a reasonable parametrisation when fitted against observed data, whereas the Linear Memory First Model did not give a reasonable parametrisation. This contrary to our findings under different assumptions. Expanding the model we used to include the contraction phase could be a source of further insight into the differences between methods.

Some models were able to better fit the data than others. It is useful to consider what qualitative properties allow a model to fit well or poorly to different sets of data ([Ganusov, 2016]). We can make certain broad comments, for example, a Linear Memory First Model can only fully match to a decreasing function of $CD62L^+$ over time as there is no way that the expected proportion of $CD62L^+$ can increase, but this is not so with the Linear Effector First Model, which can show a later increase in the proportion of $CD62L^+$ (for example Figure 46). This explains why curtailing the CD62L data when there was an late increase in $CD62L^+$ had the effect of better fitting to the Linear Memory First Model, and why the MLN data still fits best to the Effector First Model. We note the Linear Memory First model may gain flexibility elsewhere, for example it has more ways to parametrize a delay in $CD62L^-$ cells appearing than the Linear Effector First Model, which is more constricted, especially when we put a cap of five days on a cell's maximum average lifetime. Such observations are useful for getting an intuitive view of which models we would expect to fit better than others, however getting a comprehensive idea of what sets of results are likely to fit best to which model structures is left for future studies.

Finally the mathematical model could be applied to analyse other data sets. While the [Buchholz et al., 2013] fitted to clonal data, which is more difficult to acquire *in vivo*, we extended it to proportional data. It is therefore possible apply the model to other data sets that have data on proportions

of cells expressing different phenotype traits. It would be interesting to see if different markers or combinations of markers yielded the same results. One possibility could be using the CD127 data from [Badovinac et al., 2007] or [Schlub et al., 2010].

6.3 Conclusions

Firstly, we established that fitting the model to the cohort data reported in [Buchholz et al., 2013] resulted in the same best fitting model and parameterisation as to clonal data. This showed fitting to cohort data was a reasonable alternative in this instance. This is useful, as cohort data are more readily available that clonal data and this lends credence to other models fitting to similar data sets in the future. This finding also suggested robustness of the [Buchholz et al., 2013] conclusion under their assumptions.

We also found an answer to question arising when questioning the apparent discrepancy in between [Buchholz et al., 2013] and [Kinjyo et al., 2015]; these papers seemed to propose not just differing models but differing results from a similar experiment. These different conclusions included the proportion of CD62L expression and speed of proliferation in days one to seven of infection. Using the information from [Badovinac et al., 2007] and [Schlub et al., 2010] we could explain some of the differences as arising from the number of cells adoptively transferred and the fact this large transfer would cause a shorted expansion phase in [Kinjyo et al., 2015]'s experiment. Under the adapted version of method described in [Buchholz et al., 2013], whether or not we adjusted for the effects of a large adoptive transfers returned fundamentally different conclusions. This might indicate a similar effect for other models, qualitative or quantitative when comparing adoptive transfer experiments with different numbers of cells transferred. Some papers listed in Table 3 used large numbers of transferred cells in experiments. These included papers supporting the Effector First Model ([Opferman et al., 1999]: $5 \times 10^5 - 2 \times 10^6$ cells, [Jacob and Baltimore, 1999]: 2×10^5 cells) and the Asymmetric Model ([Chang et al., 2007]: 5×10^5 cells). This thesis further supports the notion that a large and small adoptive transfer experiments may yield differing conclusions, and that some of the apparent contradictions or ambiguities around questions of differentiation may arise if this difference in experimental set up is not considered.

Under the assumptions of the method, we found, the Linear Memory First Model is a stronger fit to all blood and spleen data ([Badovinac et al., 2007], [Schlub et al., 2010], [Buchholz et al., 2013] and [Kinjyo et al., 2015]) than the Linear Effector First Model. In addition, the best fit for the Linear Effector First Model was biologically implausible with the majority of naive cells unactivated by the peak immune response and was unable to give a reasonable fit when this possibility was not allowed. This inference has some caveats: assumptions were strong and data points fitted to were sometimes few ([Kinjyo et al., 2015] and [Schlub et al., 2010] after curtailing the data).

The MLN data set reported in [Kinjyo et al., 2015] supported the Linear Effector First Model under the assumptions of the method. This conclusion could be drawn intuitively by looking visually at the

proportion data fitted to (Figure 32), as there is sudden decrease, then later increase in CD62L$^+$ after day two post infection and it is difficult to explain this late increase within a Linear Memory First Model. This is at odds with the results in from the blood and spleen. This difference could be explained a number of ways, it could be due to cell migration in the MLN, it may be the cause is that the fitting method is not robust enough, that T cells behave fundamentally differently in depending on their locations or that the large adoptive transfer size has cause strange differentiation dynamics in the MLN. In addition we only had one data set from the MLN.

While these are the conclusions under the assumptions of the model, if we step outside those assumptions we can suggest other possibilities.

One possibility is that memory arises after the expansion phase. This is given credence by the fact that [Kinjyo et al., 2015] showing a memory subset appearing on day seven, which would be after the expansion. In addition the parameterisation of the best fitting models suggested the curves of CD62L during the expansion could be explained purely by effector and naive cell populations.

Another possibility is that while CD62L may correlate strongly with memory late in the immune response, in the early days of an immune response it may not indicate memory as defined in the classical sense and is mixed in with naivety.

Another possibility is the actual path is not a linear differentiation pathway, such as that seen in asymmetry theory, and so was not evaluated.

In summary we find fitting mathematical models to CD8$^+$ cohort expansion data a possible alternative to fitting to clonal data, we suggest cell transfer numbers to be a key difference between some of the results in [Buchholz et al., 2013] and [Kinjyo et al., 2015]. Under the assumptions of the model, our analysis support a Linear Memory First Model above a Linear Effector First Model for CD8$^+$ differentiation pathways. However evidence from other papers and the parameterisation challenges these assumptions and leaves open other possibilities.

Appendices

A [Schlub et al., 2010]'s model for division estimation

In Section 5.3.4 we describe an estimate which we use to determine the average family size at the peak of the immune response. In this appendix we discus the calculation this was adapted from, described in the paper [Schlub et al., 2010]. The paper [Schlub et al., 2010] investigates whether there is a link between the number of times a $CD8^+$ T cell divides during the expansion phase and the amount of CD62L it expresses. They do this via an adoptive transfer experiment described in Section 3.5. Using different frequencies of adoptively transferred cells they limit transferred colony growth. They then measure the proportion of OT-1 cells out of total lymphocytes in a mouse (OT-1 being donor cells progeny) and use these data to estimate the relative number of divisions an average T cell takes under different assumptions of lymphocyte populations. They then see if they can find an exponential correlation between the number of divisions and the proportion of $CD62L^+$ cells.

Let us denote the population count of OT-1 T cells (donor cells) in the mouse for experiment i at time t as $o_i(t)$, where $t = 0$ at the time of adoptive transfer. Let us also say $n_i(t)$ is the number of non-OT-1 lymphocyte T cells in experiment i at time t (i.e. the endogenous cells). Let us denote the average number of divisions an average T cell have gone through between two times t_1 and t_2 in experiment i as $d_i(t_1, t_2)$. For brevity we denote.

$$q_i(t) = \frac{o_i(t)}{o_i(t) + n_i(t)}. \tag{22}$$

i.e as $q_i(t)$ as the proportion of OT-1 T cells at time t in experiment i. Assuming no cell death, we can say

$$2^{d_1(t_1,t_2)} o_1(t_1) = o_1(t_2), \tag{23}$$

which can be rearranged to give

$$d_1(t_1, t_2) = log_2 \left(\frac{o_1(t_2)}{o_1(t_1)} \right). \tag{24}$$

However the absolute number of OT-1 T cells at a particular time in an experiment is often not known (except at $t = 0$), but the proportion of OT-1 cells out of all T cells is reported in [Schlub et al., 2010] and [Badovinac et al., 2007]. It is possible to get an estimate for $d_1(t_1, t_2)$ in terms of observed proportions under different assumptions.

The authors of use a number of different methods to estimate the number of divisions between two time points. We focus on that which method we adopt and adapt further. If we want to measure the difference in number of divisions between two different experiments from the time of adoptive transfer, we can use Equation 24 to get

$$d_1(0, t_1) - d_2(0, t_2) = log_2(\frac{o_1(t_1)}{o_1(0)}) - log_2 \left(\frac{o_2(t_2)}{o_2(0)} \right). \tag{25}$$

Then, assuming $o_1(t_1) + n_1(t_1) = o_1(t_2) + n_1(t_2)$, i.e. the total number of lymphocytes at the two times is a constant we say,

$$d_1(0, t_1) - d_2(0, t_2) = log_2 \left(\frac{o_2(0)p_1(t_1)}{o_1(0)p_2(t_2)} \right). \tag{26}$$

Thus the relative difference in average number of division between two experiments can be calculated from observable values. The assumption $o_1(t_1)+n_1(t_1) = o_1(t_2)+n_1(t_2)$ is justified for the time of peak response (not necessarily other time) from spleen data reported in [Schlub et al., 2010]. The authors of [Schlub et al., 2010] used similar assumptions in the calculations to work out relative divisions within or between experiments, and other assumptions include assuming the number of endogenous cells remains constant between two time points ($n_1(t_1) = n_1(t_2)$).

B Sensitivity analysis on average family size

For all data sets other than that of [Buchholz et al., 2013], absolute count of average family size was not reported in a way that we could incorporate into the fitting. We therefore used the estimate described in Section 5.3.4 to determine of average family size at peak immune response. Because this was only an estimate, we sought to investigate how sensitive the results in Section 5.2.3 were to this calculation changing. Figures 38 show the results for the best fitting objective functions for different values of the average population estimate, both for the Linear Memory First Model and Linear Effector First Model. Figure 38 shows that average family size does effect results. For the [Kinjyo et al., 2015] data, the Linear Memory First Model fits best excepting a window around 20 average family size (approximate), where both models are extremely close in terms of there objective function. The data from [Schlub et al., 2010] showed that different assumptions for average family size could return different best fit models. However, from results seen in Section 5.3.4, we suspected that some of these best may parametrised with extremely long cell lifetimes that were biologically infeasible leading to most naive cells being unactivated at the peak of the immune response. We repeated the test, but capping the peak average cell lifetime to five days (Figure 39). In this case the linear memory first model was stronger regardless of the assumption of family size, and this showed there was no biologically plausible Linear Effector First Model that could fit well to these data sets.

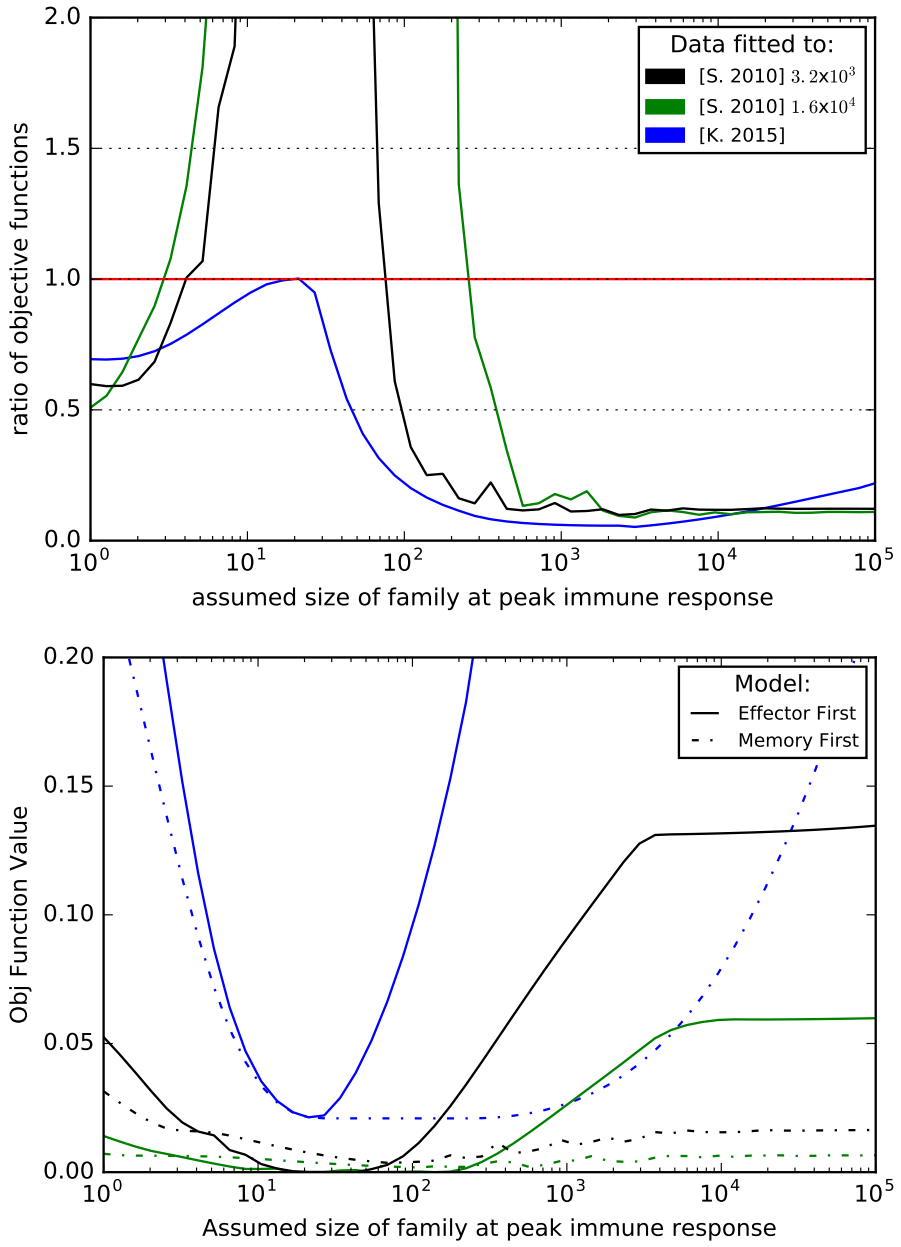


Figure 38: Objective function values when fitting models to data sets and varying the the average family size assumed. The top graph shows the ratio of the the Linear Memory First Model to the Linear Effector First Model (values under one indicate a better fit to the Linear Memory First Model), and the bottom graph shows the absolute values for these two models (smaller values indicate a better fit). Changing the assumed average family size does effect which model returns the better fit.

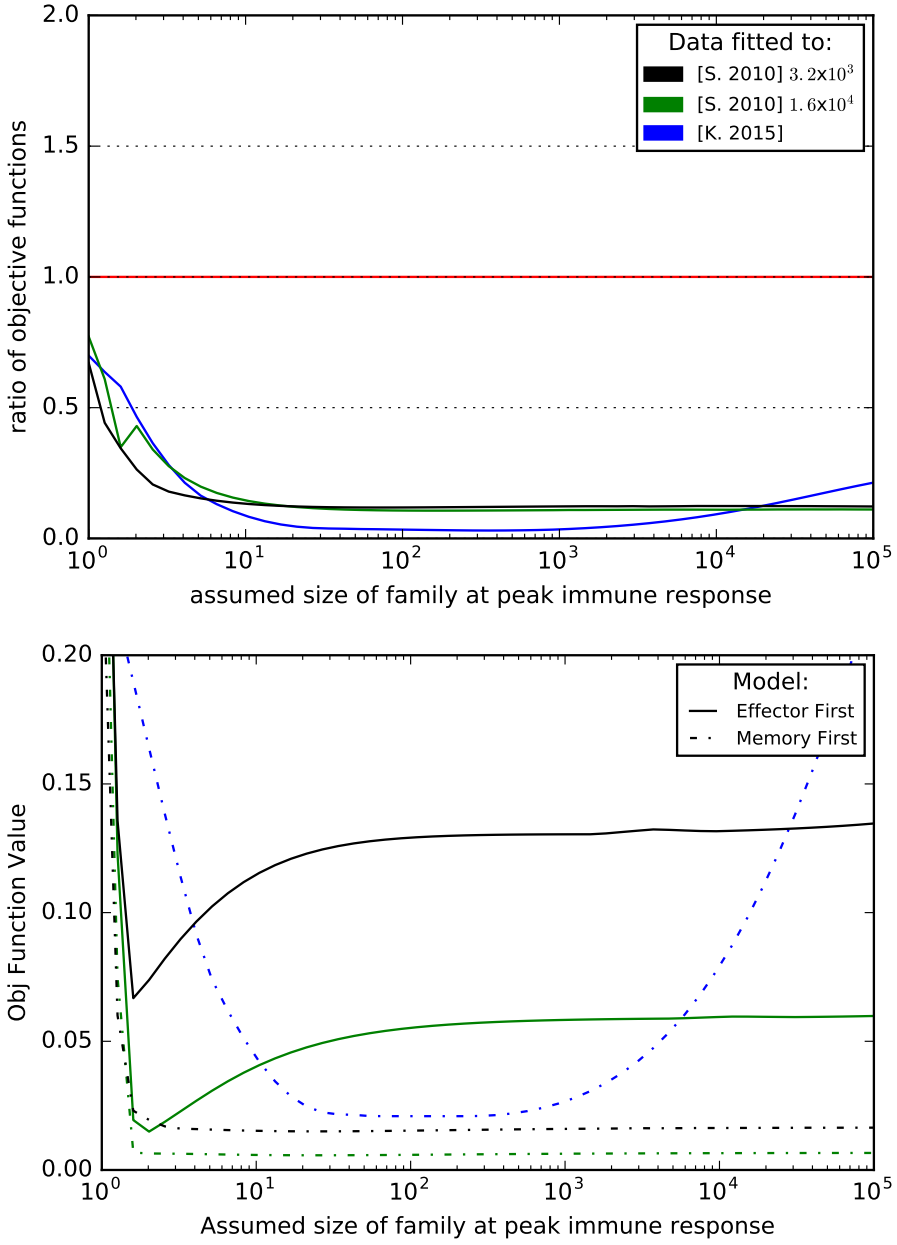


Figure 39: Objective function values when fitting models to data sets and varying the assumption of the average family size and assuming the average cell lifetime cannot exceed five days. The top graph shows the ratio of the the Linear Memory First Model to the Linear Effector First Model (values under one indicate a better fit to the Linear Memory First Model), and the bottom graph shows the absolute values (smaller values indicate a better fit). The Linear Memory First Model fits better in all cases regardless of the assumed family size.

C Changing assumptions in modelling

In Section 5 we adapted of the methodology described in [Buchholz et al., 2013] so that it could be applied to cohort data. During this process we made a number of choices for our adaptations. This appendix explores some of the results from using alternative assumptions when adapting the model. We find the core result remains the same when these assumptions are changed: the Linear Memory First Model fits best to all data sets when adjusting for the number of naive cells adoptively transferred. However the strength of the result can vary, and under some assumptions both the Linear Memory First Model and Linear Effector First Model give good fits to the data.

C.1 Changing classification of naive cells

The adaptation of the [Buchholz et al., 2013] method assumed the reported TCMp data in [Buchholz et al., 2013] include cells that would classically be described as naive. This appendix looks at the equivalent results if naive cells were assumed to be excluded from the data entirely.

In the case best fitting model is the same as in Section 5.2.3, i.e. Model 1 fits best ($CD62L^+CD27^+ \rightarrow CD62L^-CD27^+ \rightarrow CD62L^-CD27^-$) (Figure 40). Under this new assumption Model 1 is also more clearly the best fitting model than the original method. The parameterisation, visual fit and summary statistics of Model 1 were are similar to the original method as seen in Figure 41, Figure 42 and Figure 43.

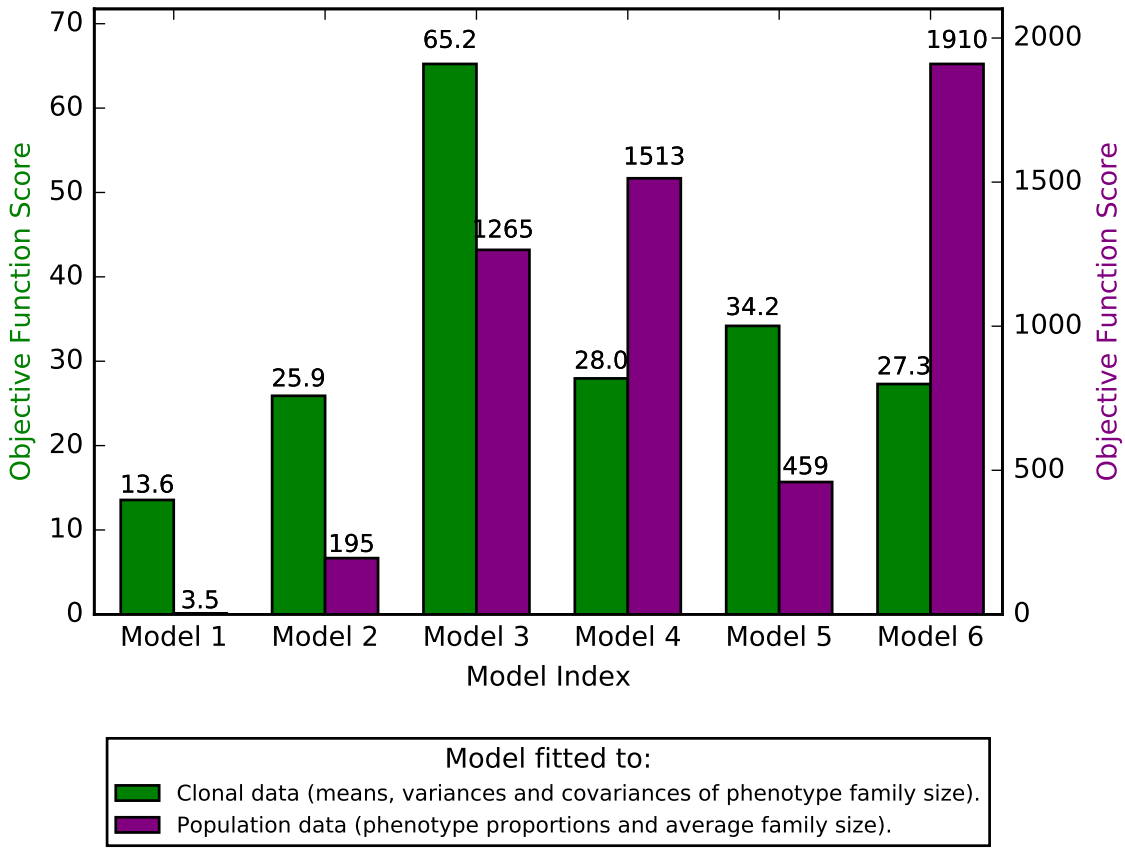


Figure 40: The objective function value for the best fitting parameterisations of the six linear models when fitting to observed statistics of mean cell count, variances and covariances as in [Buchholz et al., 2013] (green, as in Section 5.1.1) compared with fitting to proportion data and average family size (blue). Smaller values indicate a better fit. This is as in Figure 23 but assuming naive cells are not included in TCMp cells within [Buchholz et al., 2013] reported data. Model 1 ($CD62L^{+}CD27^{+}\rightarrow CD62L^{-}CD27^{+}\rightarrow CD62L^{-}CD27^{-}$) is the best fitting in both cases.

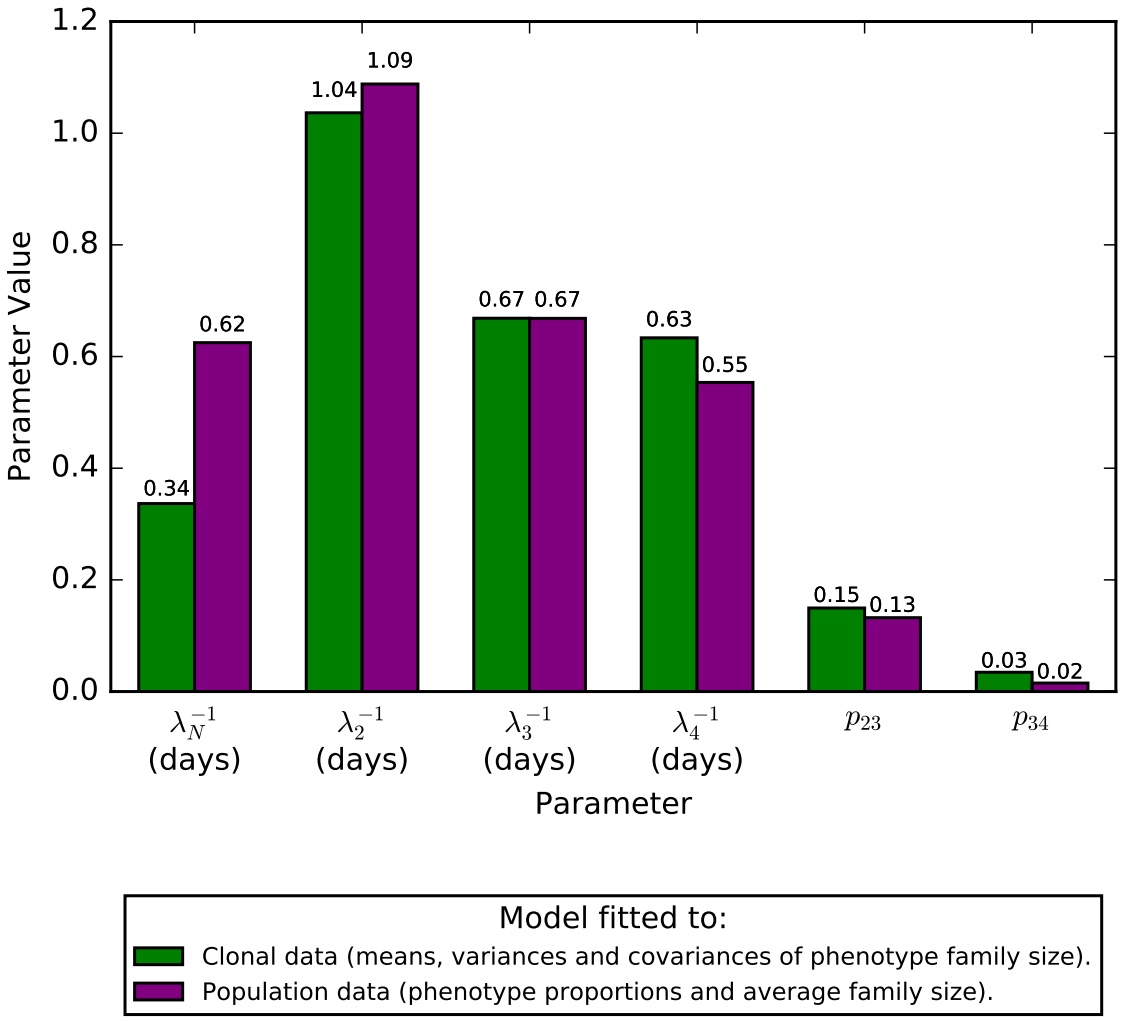


Figure 41: The best fitting parameters derived by fitting against means of cell count, variances and co-variances alongside parameters from fitting to proportional data and average family size for Model 1 (CD62L$^+$CD27$^+$→CD62L$^-$CD27$^+$→ CD62L$^-$CD27$^-$). As in Figure 24 but assuming naive cells are not included in TCMp cells within [Buchholz et al., 2013] reported data. The parameters are similar between the two best fitting models with the exception of λ_N and λ_2.

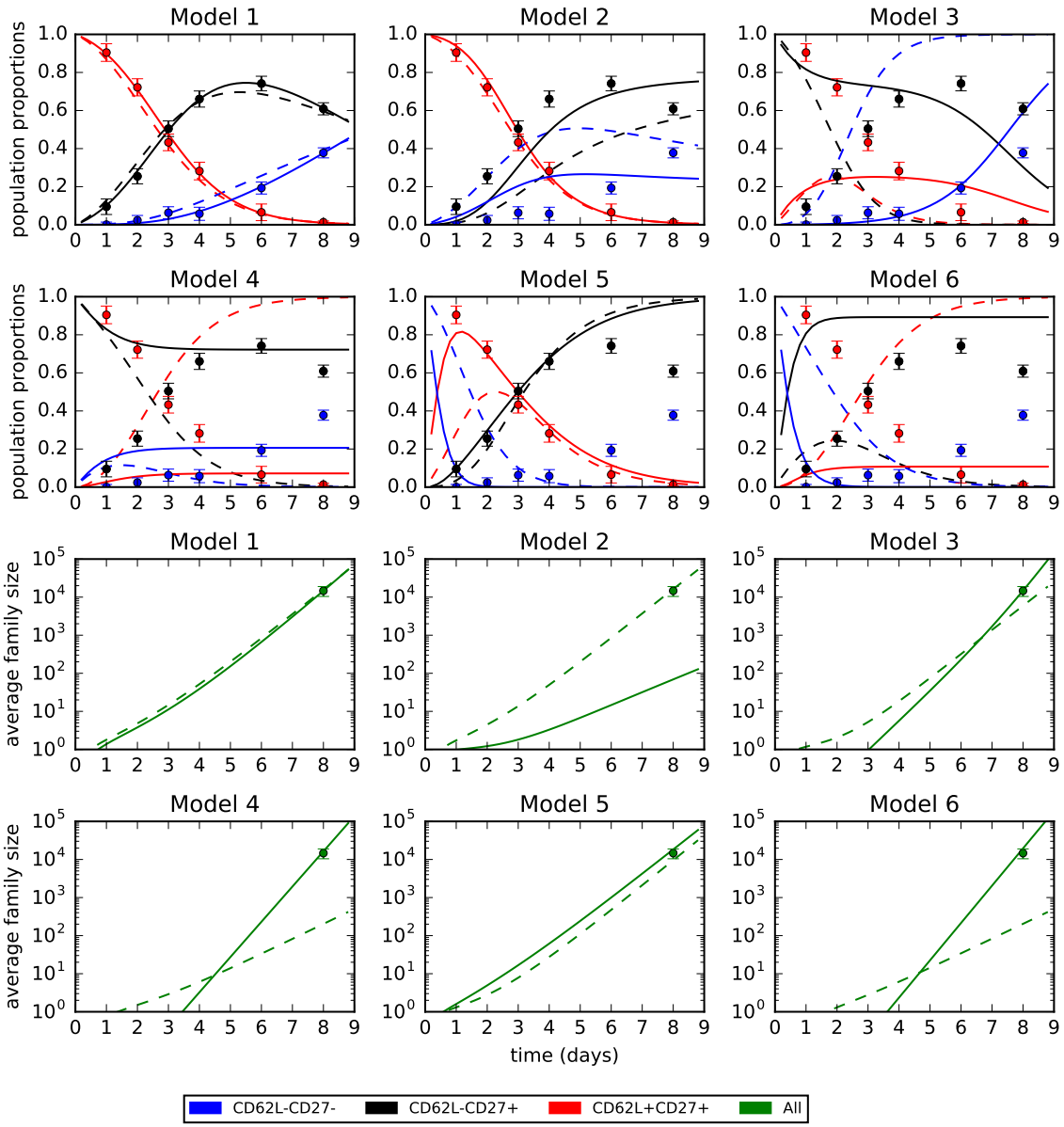


Figure 42: Proportions of phenotype of best fitting parameterisations for each of the six linear models. Dashed lines show when fitting to mean cell counts, variances and covariances and solid lines shows when against proportional data and average family size. Observed data (dots) was taken from data graphs in [Buchholz et al., 2013] supplementary material S26. Error bars show the standard error of the mean. This is as in Figure 24 but assuming naive cells are not included in TCMp cells within [Buchholz et al., 2013] reported data. Only Model 1 ($CD62L^+CD27^+ \rightarrow CD62L^-CD27^+ \rightarrow CD62L^-CD27^-$) gives a good visual fit to the data.

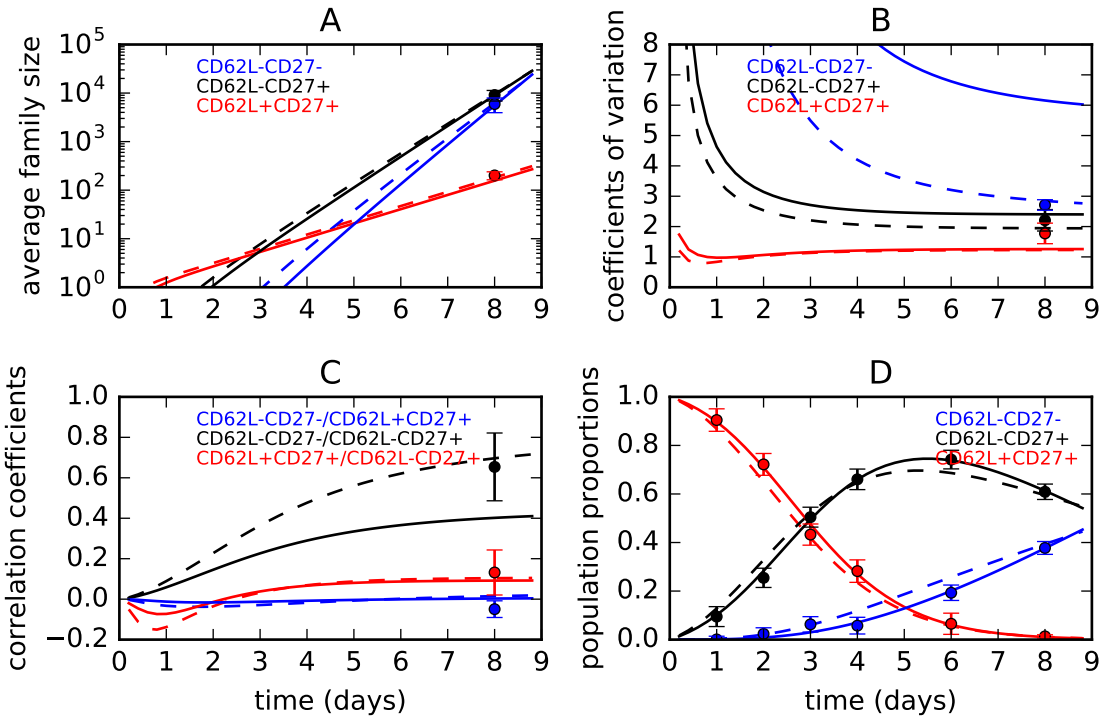


Figure 43: Statistics from Model 1 ($CD62L^+CD27^+ \rightarrow CD62L^-CD27^+ \rightarrow CD62L^-CD27^-$) for the best fit parameterisation fitting to mean cell count, variances and covariances (dashed line) compared with fitting to proportional data and average family size at different times (solid line). Statistics shown are: A mean cell count, B coefficients of variation, C correlation coefficients and D phenotype proportions. Observed data (dots) was taken from data graphs in [Buchholz et al., 2013] supplementary material S26. Error bars show the standard error of the mean. As in Figure 26 but assuming naive cells are not included in TCMp cells within [Buchholz et al., 2013] reported data. The statistics are similar between the two best fitting models.

C.2 Changing objective functions

Because choice of objective function is always somewhat subjective we sough to investigate the effects of a different objective function. Firstly, we reviewed the changes from the objective function described in [Buchholz et al., 2013] to the adapted method described in Section 5. Graph 44 show the ratio of the objective function of Model 1 to the other models for each step change of the adaptation. It shows for all steps, Model 1 is the best fittings.

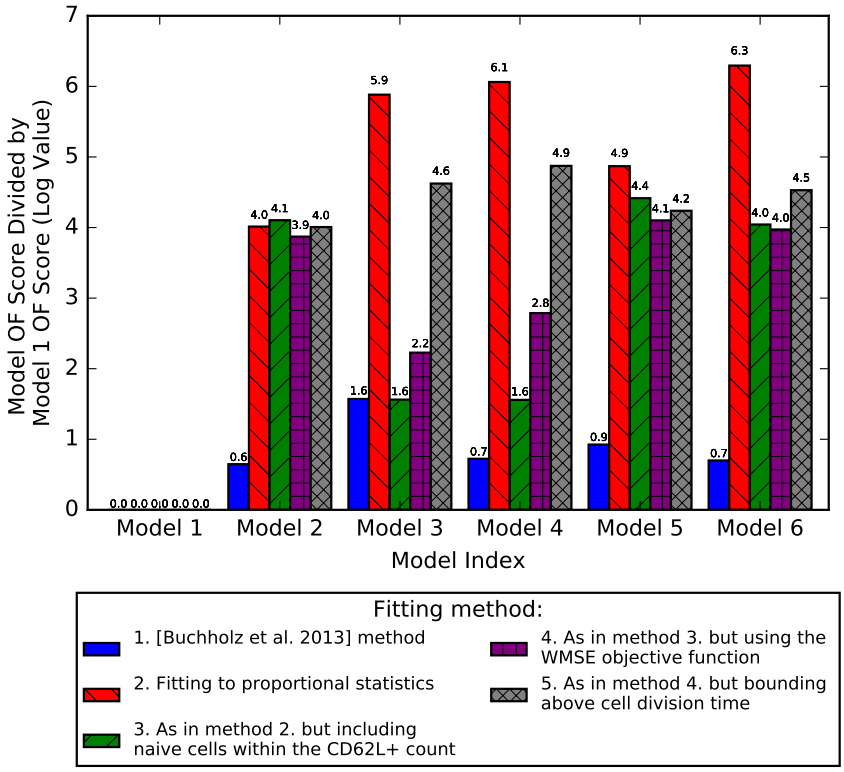


Figure 44: Step changes to the objective function from the original method described in [Buchholz et al., 2013] to the adapted version in Section 5.1.1. The above shows the ratio of best fitting objective function values of the respective models to the best fitting objective function for Model 1. Model indexing is described in Section 5.1.1. Model 1 is the best fitting in all cases. Smaller values indicate a better fit and negative values (of which there are none) would indicate a better fit than Model 1.

In Section 5.3.1 we defined the objective function WMSE_1 as an alternative to the one we used to do the analysis throughout detailed Section 5.3.5. We repeated the analysis using this alternative objective function and came to the same conclusions: the Linear Memory First Model was the best fitting, though it was in some cases close (Figure 45). This shows the robustness of results under at least one different objective function.

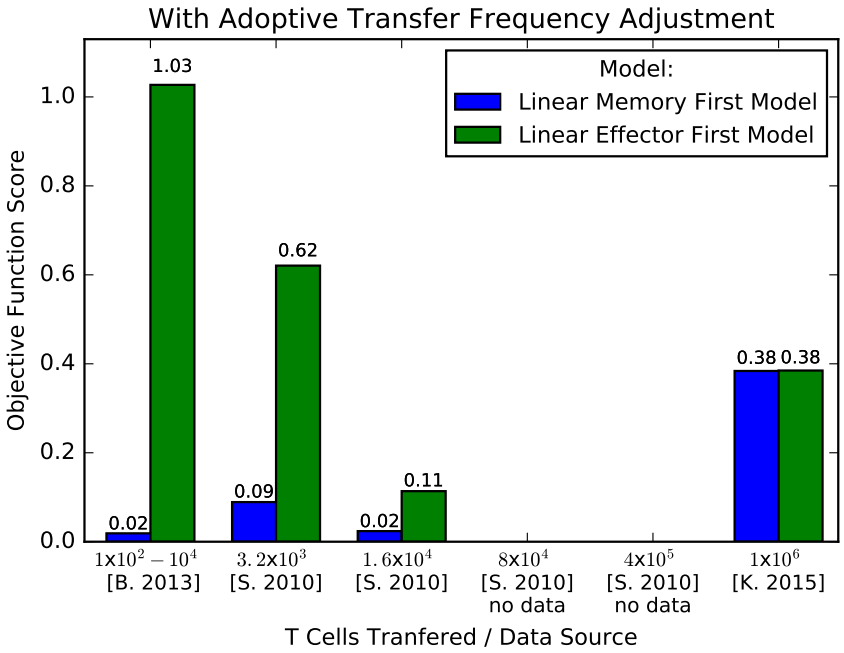


Figure 45: The objective function value for the best fitting parameterisations of the two simplified linear models when fitting to observed cohort statistics from different data sources with assuming a shortened expansion phase. Smaller values indicate a better fit. As in Figure 33 but using the alternative objective function WMSE_1 (defined in Section 5.3.1). All models fit best to a Linear Memory First Model. In this second case, the difference between the best fitting models for [Kinjyo et al., 2015] and [Buchholz et al., 2013] is marginal.

C.3 Removing transfer size adjustment

This appendix reports extra analysis on Figure 33 from Section 5.3.5. We fitted to the data as in the adapted method, but without curtailing the proportional data in the expansion phase. In this case the Linear Effector First Model fitted best for experiments with high numbers of adoptively transferred cells and the Linear Memory First Model for low numbers (Figures 33, 46, 47). This highlights the importance of curtailing the data to the expansion phase.

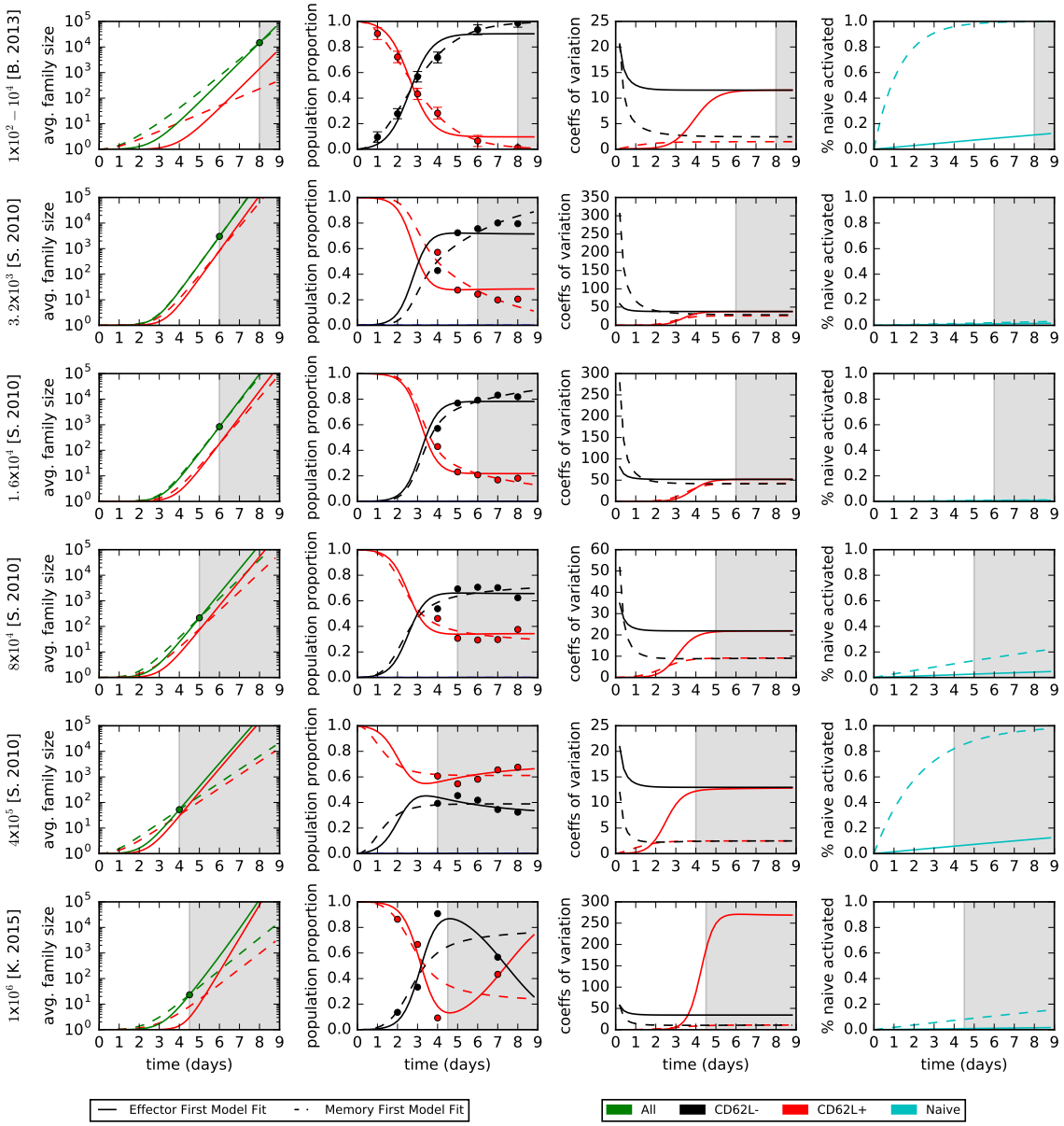


Figure 46: Statistics for the best fitting models when fitting to different sets of data from papers [Buchholz et al., 2013], [Kinjyo et al., 2015] and [Schlub et al., 2010] As in Figure 34 but not curtailing the data for the number of cells adoptively transferred. Dashed lines show the Linear Memory First Model fits, solid lines show the Linear Effector First Model fits, dots show the statistics from experimental data fitted to. The grey area shows the time after the expansion phase has ended.

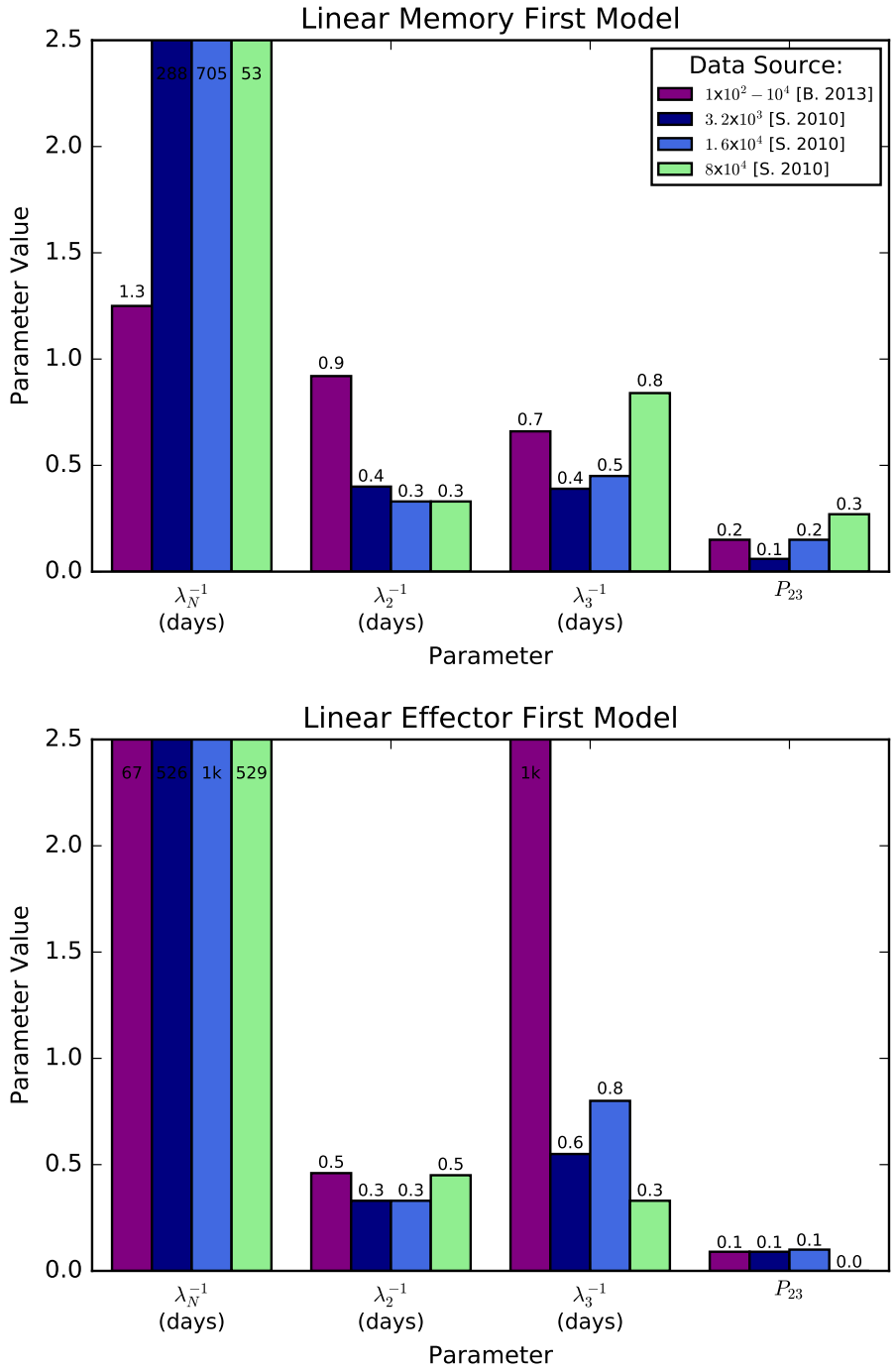


Figure 47: Parametrization for the best fitting models when fitting to different sets of data from papers [Buchholz et al., 2013], [Kinjyo et al., 2015] and [Schlub et al., 2010] As in Figure 35 but not curtailing the data for the number of cells adoptively transferred. Parameters are as described in Figure 27. Only the memory first model fitted to [Buchholz et al., 2013] data gives a biologically plausible result. The other models parametrise naive cells living for an average of over 60 days or more.

C.4 Changing the parameter boundaries

This appendix reports extra analysis on Figure 36 from Section 5.3.5. We repeated the fitting method described in this section, but in order to remove the biologically implausible fits where most naive cells had not activated by the end of the expansion phase we set a cap on average maximum cell lifetime at five days. When curtailing the data to the expansion phase, the Liner Memory First Model was best in all cases and had a much stronger result. This showed there was no biologically plausible Linear Effector First fit to these data. All Linear Memory First Models had a good visual fit (Figures 48, 49). In all cases memory cell lifetimes were parametrised to be longer than effector cells.

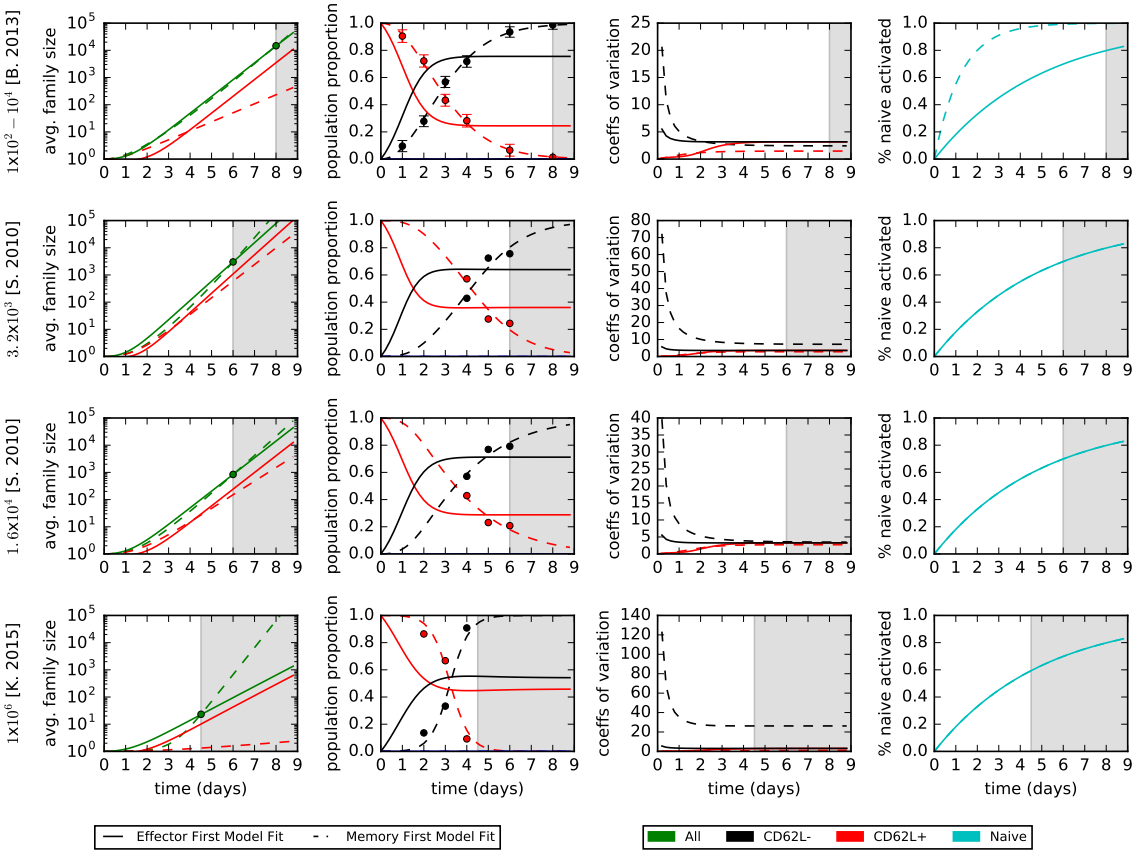


Figure 48: Statistics for the best fitting models when fitting to different sets of data from papers [Buchholz et al., 2013], [Kinjyo et al., 2015] and [Schlub et al., 2010]. As in Figure 34 but upper bounding average cell lifetime to five days. Dashed lines show the Linear Memory First Model fits, solid lines show the Linear Effector First Model fits, dots show the statistics from experimental data fitted to. The grey area shows the time after the expansion phase has ended.

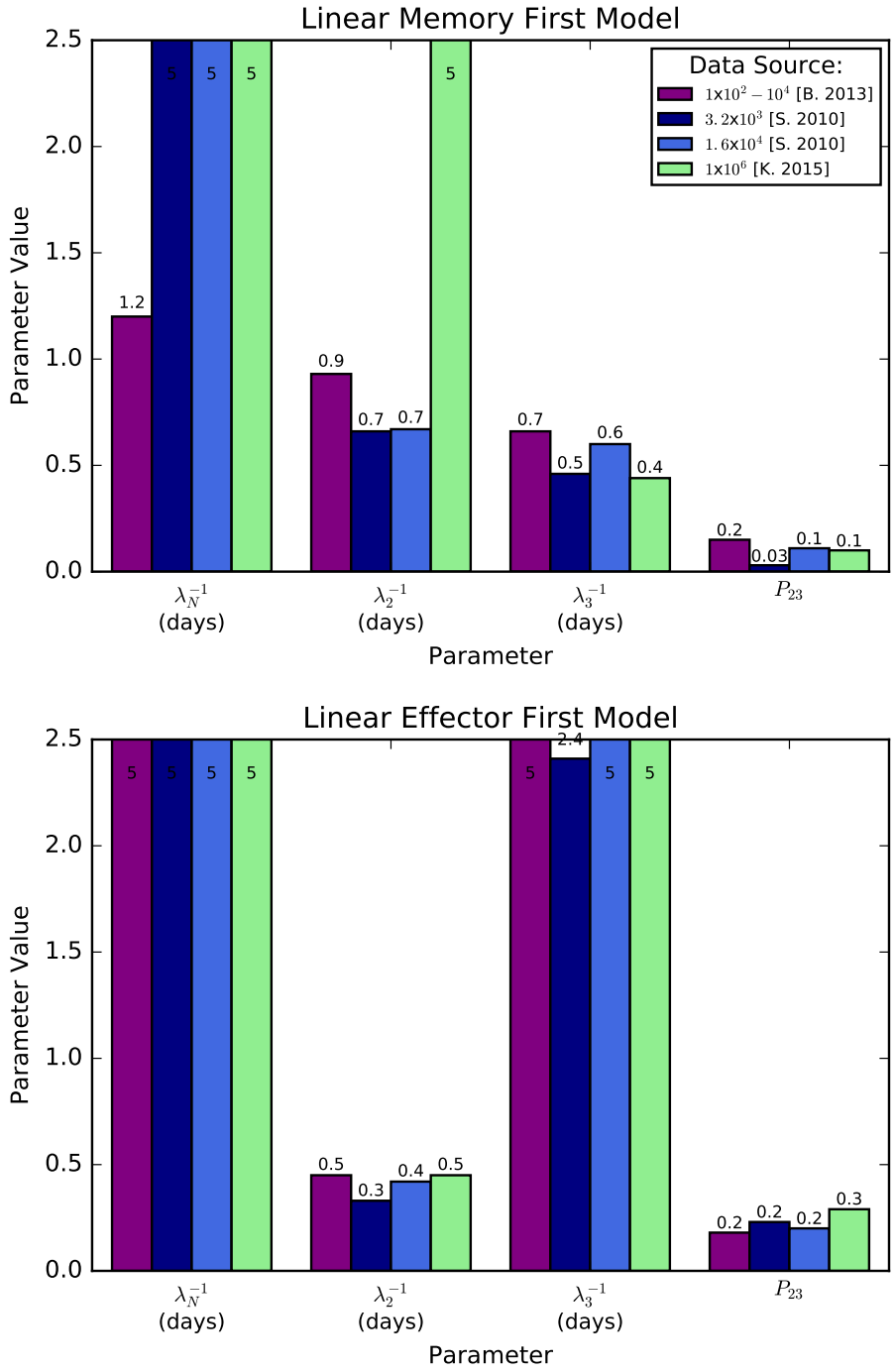


Figure 49: Parametrization for the best fitting models when fitting to different sets of data from papers [Buchholz et al., 2013], [Kinjyo et al., 2015] and [Schlub et al., 2010] As in Figure 35 but upper bounding average cell lifetime to five. The parameters are as described in Figure 27.

D Fitting to other data sets

This appendix reports extra analysis on Figure 37 from Section 5.3.5. Initially we fitted models to spleen data reported in [Buchholz et al., 2013] and [Kinjyo et al., 2015] and blood data reported in [Schlub et al., 2010]. After curtailing the data some of the [Schlub et al., 2010] data could not be fitted to because of too few data points (less data points that parameters) so we fitted to other data sources available.

[Badovinac et al., 2007] performed an adoptive transfer transfer experiment using $1\times10^5, 2\times10^5, 3\times10^5, 4\times10^5$ and 5×10^5 OT-1 cells and took blood samples at different days post infection. After curtailing these data to the expansion phase using the method described in Section 5.3.3 we found three of these data sets had the minimum number of data points to fit to. Another additional set of data was reported in [Kinjyo et al., 2015], but for cells harvested from the MLN rather than the spleen (transfer 5×10^5). This was not included in the main analysis initially because of migration. We fitted models to these two data sets using the adapted method described in Section 5.3, both curtailing the data and limiting average cell lifetime to a maximum of five days to avoid the non biologically viable parameterisations seen in other data sets. This gave us the results reported in Figure 37. The blood data from [Badovinac et al., 2007] fits best to a Linear Memory First Model, in keeping with the other fittings to blood and spleen data seen in Section 5.3.The the 5×10^5 fit is less strong and parametrises for fast memory cells. The MLN data fits best to a Linear Effector First Model.

Figure 50 shows the summary statistics for these fits and Figure 51 shows their best fitting parameterisations. The [Kinjyo et al., 2015] lymph data fitting shows an increase in numbers of $CD62L^+$ cells before the end of the expansion phase. The difference between MLN results and blood and spleen results could be due to migration of T cells around the body or differences in the development of cells at those locations. The high level of migration of the MLN data meant we did not include it in our main analysis.

It is noticeable that the results on the [Badovinac et al., 2007] 5×10^5 data set did not give a good visual fit to the data for either model and also returned a strange parameterisation, where memory cells have shorter lives than effectors. This could be due to problems with the model, or problems with the data which can be seen to oscillate (Figure 32).

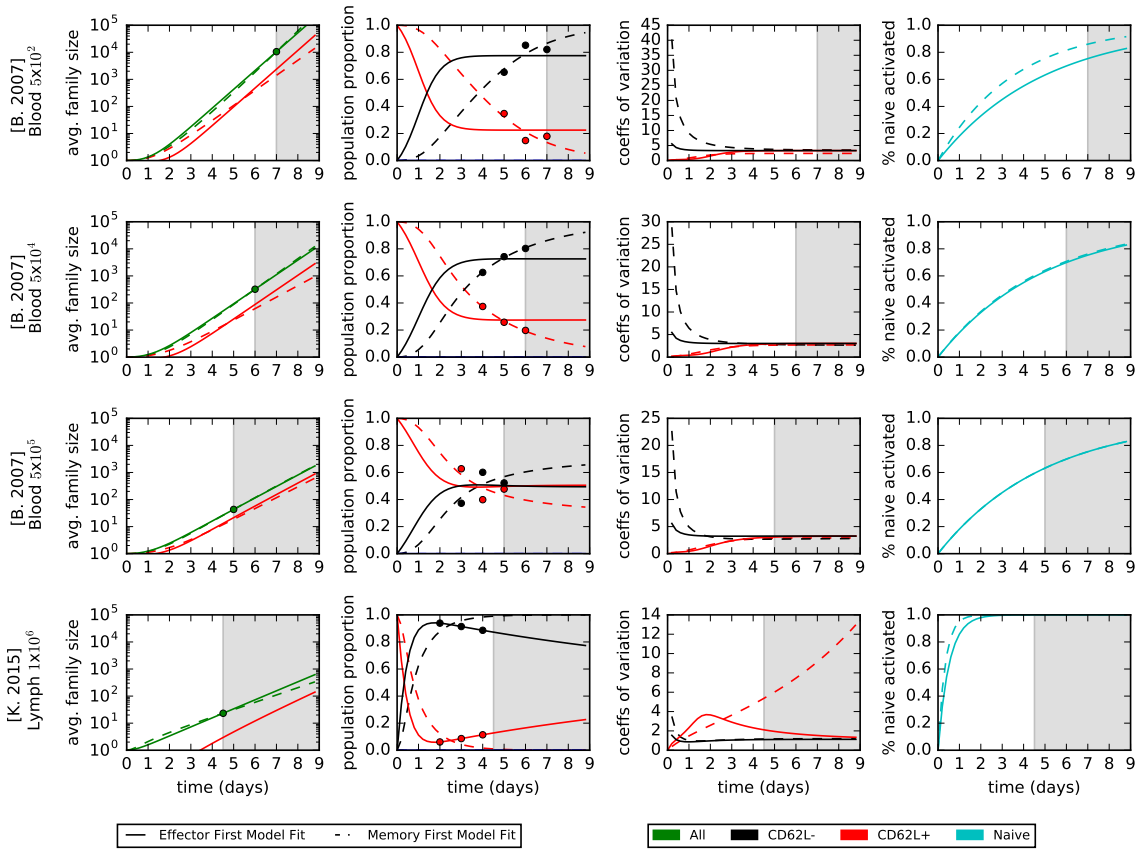


Figure 50: Statistics for the best fitting models when fitting to different sets of data from papers [Badovinac et al., 2007] blood data and [Kinjyo et al., 2015] MLN data. Data curtailed to the expansion phase and parameterisation of average cell lifetime is upper bounded to five days. Dashed lines show the Linear Memory First Model, solid lines show the Linear Effector First Model, dots show experimental data fitted to. The grey area shows the time after the expansion phase has ended. [Badovinac et al., 2007] data shows a better fit to the Linear Memory First Model whereas [Kinjyo et al., 2015] MLN data better fits to a Linear Effector First Model.

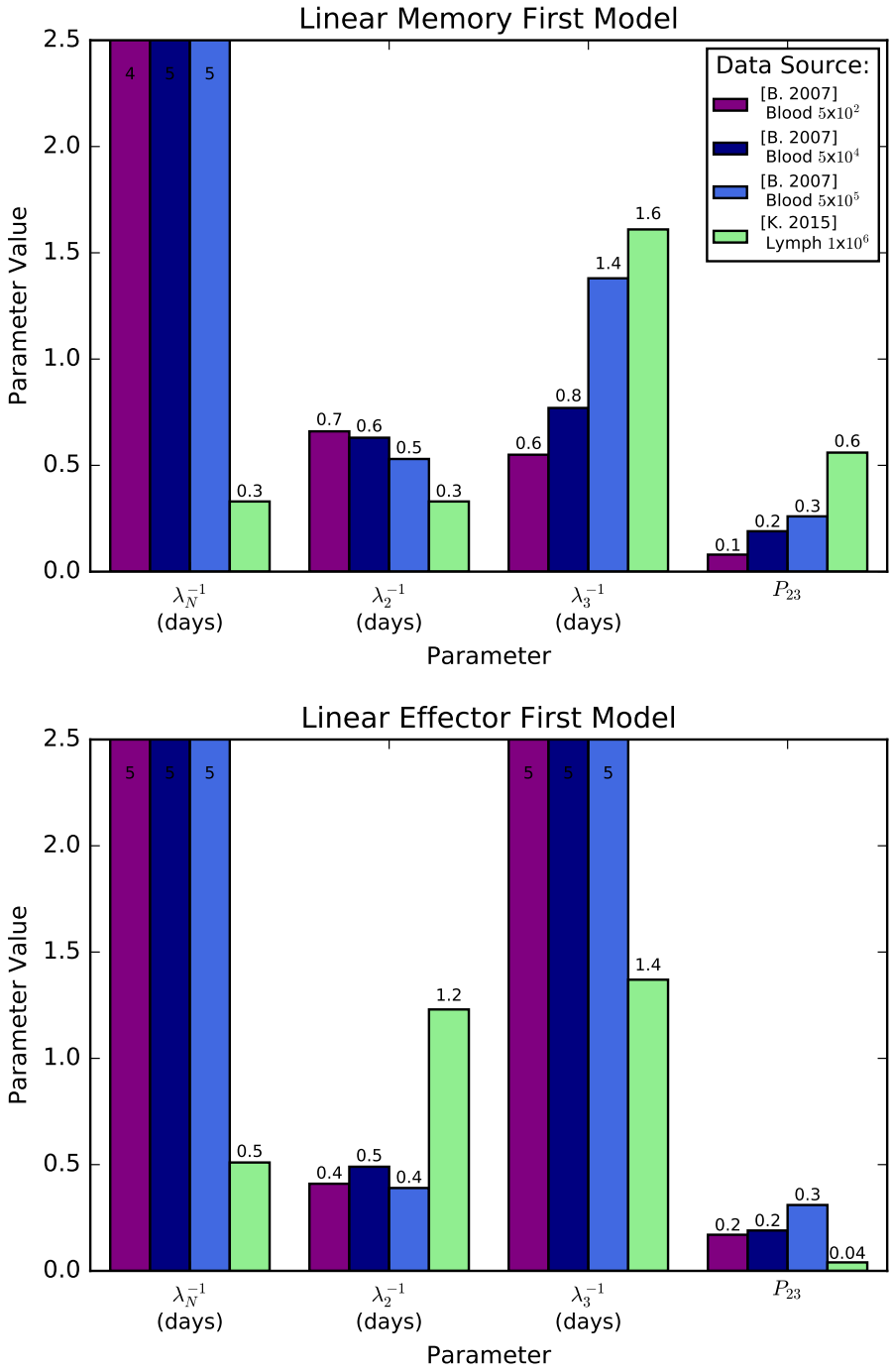


Figure 51: Parametrization for the best fitting models when fitting to different sets of data from papers [Badovinac et al., 2007] blood data and [Kinjyo et al., 2015] MLN data. The parameters are as described in Figure 27. Some parameterisations show memory having a shorter average lifetime than effector cells. Lymph data are likely heavily effected by migration.

References

Rafi Ahmed, Michael J. Bevan, Steven L. Reiner, and Douglas T. Fearon. The precursors of memory: models and controversies. *Nature Reviews Immunology*, 9:662–668, 2009.

Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

Bruce Alberts, Alexander Johnson, Julian Lewis, and et al. *Molecular Biology of the Cel, fourth edition.* New York: Garland Science, 2002.

Matthew N. Alder, Igor B. Rogozin, Lakshminarayan M. Iyer, Galina V. Glazko, Max D. Cooper, and Zeev Pancer. Diversity and function of adaptive immune receptors in a jawless vertebrate. *Science*, 310(5756):1970–1973, 2005.

Rustom Antia, Vitaly V. Ganusov, and Rafi Ahmed. The role of models in understanding CD8+ T-cell memory. *Nature Reviews Immunology*, 5:101–111, 2005.

Paul A Baars, Sophie Sierro, Ramon Arens, Kiki Tesselaar, Berend Hooibrink, Pau Klenerman, and RenAW l van Lier. Properties of murine CD8+CD27 T cells. *European Journal of Immunology*, 35 (11):3131–3141, 2005.

Martin F. Bachmann, Petra Wolint, Katrin Schwarz, Petra Jager, and Annette Oxenius. Functional properties and lineage relationship of CD8+ T cell subsets identified by expression of IL-7 receptor and CD62L. *The Journal of Immunology*, 175(7):4686–4696, 2005.

Vladimir P. Badovinac, Jodie S. Haring, and John T. Harty. Initial T cell receptor transgenic cell precursor frequency dictates critical aspects of the CD8+ T cell response to infection. *Immunity*, 26(6):827841, 2007.

Dimitri P Bertsekas and John N Tsitsiklis. *Introduction to Probability.* Athena Scientific books. Athena Scientific, 2002.

Ccile Bouneaud, Zacarias Garcia, Philippe Kourilsky, and Christophe Pannetier. Lineage relationships, homeostasis, and recall capacities of central and effector memory CD8 T cells in vivo. *The Journal of Experimental Medicine*, 201(4):579590, 2005.

Veit R. Buchholz, Michael Flossdorf, Inge Hensel, Lorenz Kretschmer, Bianca Weissbrich, Patricia Gräf, Admar Verschoor, Matthias Schiemann, Thomas Höfer, and Dirk H. Busch. Disparate individual fates compose robust CD8+ T cell immunity. *Science*, 340(6132):630–635, 2013.

Frank M Burnet. A modification of jerne's theory of antibody production using the concept of clonal selection. *The Australian Journal of Science*, 20:6769, 1957.

Richard Cammack, Teresa Atwood, Peter Campbell, Howard Parish, Anthony Smith, Frank Vella, and John Stirling. Oxford dictionary of biochemistry and molecular biology, 2006.

Thomas Chrowder Chamberlin. The method of multiple working hypotheses. *Science (old series)*, 15: 92–96, 1890.

John T. Chang, Vikram R. Palanivel, Ichiko Kinjyo, Felix Schambach, Andrew M. Intlekofer, Arnob Banerjee, Sarah A. Longworth, Kristine E. Vinup, Paul Mrass, Jane Oliaro, Nigel Killeen, Jordan S. Orange, Sarah M. Russell, Wolfgang Weninger, and Steven L. Reiner. Asymmetric T lymphocyte division in the initiation of adaptive immune responses. *Science*, 315(5819):1687–1691, 2007.

Rui Chen and Ollivier Hyrien. Quasi- and pseudo-maximum likelihood estimators for discretely observed continuous-time markov branching processes. *Journal of Statistical Planning and Inference*, 141(7):22092227, 2011.

Andrew R. Conn, Nicholas I. M. Gould, and Philippe L. Toint. *Trust-Region Methods (MPS-SIAM Series on Optimization)*. Society for Industrial and Applied Mathematics, 1987.

Elissa K. Deenick, Amanda V. Gett, and Philip D. Hodgkin. Stochastic model of T cell proliferation: A calculus revealing IL-2 regulation of precursor frequencies, cell cycle time, and survival. *The Journal of Immunology*, 170(10):4963–4972, 2003.

Mark R Dowling and Philip D Hodgkin. Modelling naive T-cell homeostasis: consequences of heritable cellular lifespan during ageing. *Immunol Cell Biol*, 87(6):445– 456, 2009.

Mark R. Dowling, Andrey Kan, Susanne Heinzel, Jie H. S. Zhou, Julia M. Marchingo, Cameron J. Wellard, John F. Markham, and Philip D. Hodgkin. Stretched cell cycle model for proliferating lymphocytes. *Proceedings of the National Academy of Sciences*, 111(17):6377–6382, 2014.

Warren N. D'Souza and Stephen M. Hedrick. Cutting edge: Latecomer CD8 T cells are imprinted with a unique differentiation program. *The Journal of Immunology*, 177(2):777–781, 2006.

Ken R. Duffy and Vijay G. Subramanian. On the impact of correlation between collaterally consanguineous cells on lymphocyte population dynamics. *Journal of Mathematical Biology*, 59(2):255–285, 2009.

Ken R. Duffy, Cameron J. Wellard, John F. Markham, Jie H. S. Zhou, Ross Holmberg, Edwin D. Hawkins, Jhagvaral Hasbold, Mark R. Dowling, and Philip D. Hodgkin. Activation-induced B cell fates are selected by intracellular stochastic competition. *Science*, 335(6066):338–341, 2012. doi: 10.1126/science.1213230.

Vitaly V. Ganusov. Discriminating between different pathways of memory CD8+ T cell differentiation. *The Journal of Immunology*, 179(8):5006–5013, 2007.

Vitaly V. Ganusov. Strong inference in mathematical modeling: A method for robust science in the twenty-first century. *Frontiers in Microbiology*, 7:1131, 2016.

Carmen Gerlach, Jan C. Rohr, Leïla Perié, Nienke van Rooij, Jeroen W. J. van Heijst, Arno Velds, Jos Urbanus, Shalin H. Naik, Heinz Jacobs, Joost B. Beltman, Rob J. de Boer, and Ton N. M. Schumacher. Heterogeneous differentiation patterns of individual CD8+ T cells. *Science*, 340(6132): 635–639, 2013.

Laurie E Harrington, Karen M. Janowski, James R. Oliver, Allan J. Zajac, and Casey T. Weaver. Memory CD4 T cells emerge from effector T-cell progenitors. *Nature*, 452(7185):356–360, 2008.

Theodore E Harris. *The Theory of Branching Processes*. The RAND Corporation, 1964.

E. D. Hawkins, J. F. Markham, L. P. McGuinness, and P. D. Hodgkin. A single-cell pedigree analysis of alternative stochastic lymphocyte fates. *Proceedings of the National Academy of Sciences*, 106 (32):13457–13462, 2009.

Susanne Heinzel, Tran Binh-Giang, Andrey Kan, Julia M Marchingo, Bryan K Lye, Lynn M Corcoran, and Philip D Hodgkin. A Myc-dependent division timer complements a cell-death timer to regulate T cell and B cell responses. *Nature Immunology*, 18:96–103, 2017.

Jenny Hendriks, Loes A Gravestein, Kiki Tesselaar, Rene A W van Lier, Ton N M Schumacher, and Jannie Borst. CD27 is required for generation and long-term maintenance of T cell immunity. *Nature Immunology*, 1(5):433–440, 2000.

Hirokazu Hikono, Jacob E Kohlmeier, Shiki Takamura, Susan T Wittmer, Alan D Roberts, and David L Woodland. Activation phenotype, rather than central or effectormemory phenotype, predicts the recall efficacy of memory CD8(+) T cells. *J Exp Med*, 204(7):1625–1636, 2007.

Philip D Hodgkin, Mark R Dowling, and Ken R Duffy. Why the immune system takes its chances with randomness. *Nature Reviews Immunology*, 14(10):711–711, 2014.

Joshyi Jacob and David Baltimore. Modelling T-cell memory by genetic marking of memory T cells in vivo. *Nature*, 399:593–597, 1999.

Nikhil S Joshi, Weiguo Cui, Anmol Chandele, Heung Kyu Lee, David R Urso, James Hagman, Laurent Gapin, and Susan M Kaech. Inflammation directs memory precursor and short-lived effector CD8+ T cell fates via the graded expression of T-bet transcription factor. *Immunity*, 27(2):281295, 2007.

Susan M. Kaech and Rafi Ahmed. Memory CD8+ T cell differentiation: initial antigen encounter triggers a developmental program in nave cells. *Nature Immunology*, 2:415 – 422, 2001.

Susan M. Kaech, Scott Hemby, Ellen Kersh, and Rafi Ahmed. Molecular and functional profiling of memory CD8 T cell differentiation. *Cell*, 111(6):837–51, 2002a.

Susan M Kaech, E. John Wherry, and Rafi Ahmed. Effector and memory T-cell differentiation: implications for vaccine development. *Nature Reviews Immunology*, 2-4(1474-1733):251–262, 2002b.

Katherine Kedzierska, John Stambas, Misty R. Jenkins, Rachael Keating, Stephen J. Turner, and Peter C. Doherty. Location rather than CD62L phenotype is critical in the early establishment of influenza-specific CD8+ T cell memory. *Proceedings of the National Academy of Sciences*, 104(23): 9782–9787, 2007.

Carolyn G. King, Sabrina Koehli, Barbara Hausmann, Mathias Schmaler, Dietmar Zehn, and Ed Palmer. T cell affinity regulates asymmetric division, effector cell differentiation, and tissue pathology. *Immunity*, 37(4):709–720, 2012.

Ichiko Kinjyo, Jim Qin, Sioh-Yang Tan, Cameron J. Wellard, Paulus Mrass, William Ritchie, Atsushi Doi, Lois L. Cavanagh, Michio Tomura, Asako Sakaue-Sawano, Osami Kanagawa, Atsushi Miyawaki, Philip D. Hodgkin, and Wolfgang Weninger. Real-time tracking of cell cycle progression during CD8+ effector and memory T-cell differentiation. *Nature Communications*, 6:6301 EP, 2015.

Karin M. Knudson, Curtis J. Pritzl, Vikas Saxena, Amnon Altman, Mark A. Daniels, and Emma Teixeiro. Nfb-pim-1-eomesodermin axis is critical for maintaining CD8 T-cell memory quality. *Proceedings of the National Academy of Sciences*, 114(9):E1659–E1667, 2017.

Andrey N Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin, 1933.

Bruce A Lyons and Christopher R Parish. Determination of lymphocyte division by flow cytometry. *Journal of Immunological Methods*, 171(1):131 – 137, 1994.

Kevin Man and Axel Kallies. Synchronizing transcriptional control of T cell metabolism and function. *Nature Reviews Immunology*, 15(9):574–584, 2015.

Julia M. Marchingo, Andrey Kan, Robyn M. Sutherland, Ken R. Duffy, Cameron J. Wellard, Gabrielle T. Belz, Andrew M. Lew, Mark R. Dowling, Susanne Heinzel, and Philip D. Hodgkin. Antigen affinity, costimulation, and cytokine inputs sum linearly to amplify T cell expansion. *Science*, 346(6213):1123–1127, 2014.

Julia M Marchingo, Giulio Prevedello, Andrey Kan, Susanne Heinzel, Philip D Hodgkin., and Ken . Duffy. T-cell stimuli independently sum to regulate an inherited clonal division fate. *Nature Communications*, 7:13540, 2016.

John F. Markham, Cameron J. Wellard, Edwin D. Hawkins, Ken R. Duffy, and Philip D. Hodgkin. A minimum of two distinct heritable factors are required to explain correlation structures in proliferating lymphocytes. *Journal of The Royal Society Interface*, 7(48):1049–1059, 2010.

David Masopust and Jason M. Schenkel. The integration of T cell migration, differentiation and function. *Nature Reviews Immunology*, 13:309–320, 2013.

David Masopust, Vaiva Vezys, Amanda L. Marzo, and Leo Lefrançois. Preferential localization of effector memory cells in nonlymphoid tissue. *Science*, 291(5512):2413–2417, 2001.

David Masopust, Sang-Jun Ha, Vaiva Vezys, and Rafi Ahmed. Stimulation history dictates memory CD8 T cell phenotype: Implications for prime-boost vaccination. *The Journal of Immunology*, 177 (2):831–839, 2006.

Javier Mestas and Christopher C. W. Hughes. Of mice and not men: Differences between mouse and human immunology. *The Journal of Immunology*, 172(5):2731–2738, 2004.

Dina Montufar-Solis, Alexander Williams, Nadarajah Vigneswaran, and John R. Klein. Involvement of Ly6C, 4-1BB, and KLRG1 in the activation of lamina propria lymphocytes in the small intestine of sanroque mice. *Biochemical and Biophysical Research Communications*, 483:590–595, 2017.

Scott N. Mueller, Thomas Gebhardt, Francis R. Carbone, and William R. Heath. Memory T cell subsets, migration patterns, and tissue residence. *Annual Review of Immunology*, 31(1):137–161, 2013.

Kaja Murali-Krishna and Rafi Ahmed. Cutting edge: Naive T cells masquerading as memory cells. *The Journal of Immunology*, 165(4):1733–1737, 2000.

Kenneth Murphy, Paul Travers, and Mark Walport. *Janeways's immunobiology*. Taylor and Francis Group LLC, 2008.

Stephen G Nash. Preconditioning of truncated-newton methods. *SIAM Journal on Scientific and Statistical Computing*, 6(3):599–616, 1985.

Jorge Nocedal and Stephen J Wright. *Numerical Optimization*, pages 529–562. Springer New York, New York, NY, 2006.

Joseph T. Opferman, Bertram T. Ober, and Philip G. Ashton-Rickardt. Linear differentiation of cytotoxic effectors into memory T lymphocytes. *Science*, 283(5408):1745–1748, 1999.

Youdong Pan, Tian Tian, Chang Ook Park, Serena Y. Lofftus, Shenglin Mei, Xing Liu, Chi Luo, John T. OMalley, Ahmed Gehad, Jessica E. Teague, Sherrie J. Divito, Robert Fuhlbrigge, Pere Puigserver, James G. Krueger, Gkhan S. Hotamisligil, Rachael A. Clark, and Thomas S. Kupper. Survival of tissue-resident memory T cells requires exogenous lipid uptake and metabolism. *Nature*, 543:252– 256, 2017.

W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. Press Syndicate of the University of Cambridge, 1992.

Asako Sakaue-Sawano, Hiroshi Kurokawa, Toshifumi Morimura, Aki Hanyu, Hiroshi Hama, Hatsuki Osawa, Saori Kashiwagi, Kiyoko Fukami, Takaki Miyata, Hiroyuki Miyoshi, Takeshi Imamura, Masaharu Ogawa, Hisao Masai, and Atsushi Miyawaki. Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell*, 132:487–498, 2008.

Federica Sallusto, Danielle Lenig, Reinhold Forster, Martin Lipp, and Antonio Lanzavecchia. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature*, 401(6754), 1999.

Surojit Sarkar, Volker Teichgräber, Vandana Kalia, Antonio Polley, David Masopust, Laurie E. Harrington, Rafi Ahmed, and E. John Wherry. Strength of stimulus and clonal competition impact the rate of memory CD8 T cell differentiation. *The Journal of Immunology*, 179(10):6704–6714, 2007.

Surojit Sarkar, Vandana Kalia, W. Nicholas Haining, Bogumila T. Konieczny, Shruti Subramaniam, and Rafi Ahmed. Functional and genomic profiling of effector CD8 T cell subsets with distinct memory fates. *Journal of Experimental Medicine*, 205(3):625–640, 2008.

Timothy E. Schlub, Vanessa Venturi, Katherine Kedzierska, Cameron Wellard, Peter C. Doherty, Stephen J. Turner, Ruy M. Ribeiro, Philip D. Hodgkin, and Miles P. Davenport. Division-linked differentiation can account for CD8+ T-cell phenotype in vivo. *European Journal of Immunology*, 39(1):67–77, 2009.

Timothy E. Schlub, Vladimir P. Badovinac, Jaime T. Sabea, John T. Harty, and Miles P. Davenport. Predicting CD62L expression during the CD8+ T cell response in vivo. *Immunology and Cell Biology*, 88(2):157–164, 2010.

Ton N. Schumacher and Robert D. Schreiber. Neoantigens in cancer immunotherapy. *Science*, 348 (6230):69–74, 2015.

Sander Sieuwerts, Frank A M De Bok, E. Mols, Willem M De-Vos, and Johan van Hylckama-Vlieg. A simple and fast method for determining colony forming units. *Letters in Applied Microbiology*, 47 (4):275–278, 2008.

Kelan G Tantisira, Eun Sook Hwang, Benjamin A Raby, Eric S Silverman, Stephen L Lake, Brent G Richter, Stanford L Peng, Jeffrey M Drazen, Laurie H Glimcher, and Scott T Weiss. Tbx21: A functional variant predicts improvement in asthma with the use of inhaled corticosteroids. *Proc Natl Acad Sci U S A*, 101(52):18099–18104, 2004.

Emma Teixeiro, Mark A. Daniels, Sara E. Hamilton, Adam G. Schrum, Rafael Bragado, Stephen C. Jameson, and Ed Palmer. Different T cell receptor signals determine CD8+ memory versus effector development. *Science*, 323(5913):502–505, 2009.

E. John Wherry, Volker Teichgräber, Todd C. Becker, David Masopust, Susan M. Kaech, Rustom Antia, Ulrich H. von Andrian, and Rafi Ahmed. Lineage relationship and protective immunity of memory CD8 T cell subsets. *Nature Immunology*, 4:225 – 234, 2003.

Michael J. V. White, David Roife, and Richard H. Gomer. Galectin-3 binding protein secreted by breast cancer cells inhibits monocyte-derived fibrocyte differentiation. *The Journal of Immunology*, 195(4):1858–1867, 2015.