MODELLING THE SPREAD OF THE HUMAN PAPILLOMAVIRUS ON THE CERVIX

MODELLING THE SPREAD OF THE HUMAN PAPILLOMAVIRUS ON THE CERVIX

BY

SPENCER HUNT, B.SC. HON.

A THESIS

SUBMITTED TO THE DEPARTMENT OF MATHEMATICS

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

Master of Science (2015) McMaster University

(Mathematics) Hamilton, Ontario, Canada

TITLE:	Modelling the Spread of the Human Papillomavirus on the Cervix
AUTHOR:	Spencer Hunt, B.Sc. hon. B.Sc. hons., (Mathematics) University of British Columbia, Kelowna, Canada
SUPERVISOR:	Dr. Jonathan Dushoff
NUMBER OF PAGES:	xvii, 158

Lay Abstract

The human papillomavirus (HPV) is a sexually transmitted infection that is known to cause cervical cancer in women along with other genital cancers. Cervical cancer is the fourth most common cancer in women, and thus researchers are looking to reduce the number of cervical cancer cases and the number of HPV infections. In order for HPV to cause cervical cancer, the infection must persist for a long time. Most individuals clear the infection without any complication; however, some individuals develop persistent infections. By using mathematical and computation models, we hope to understand why and how HPV infections spread in the host. We develop a criterion for when the infection may be able to establish in the host, and explore conditions that could lead to clearance. Understanding when and how infections will persist could inform treatment and monitoring of cervical cancer development.

Abstract

Cervical cancer is the fourth most common cancer in women. It is caused by the human papillomavirus (HPV). There are many different types of HPV, some of which are high-risk, highly associated with cancer, and low-risk. While HPV is very common—most sexually active individuals will contract some sexually transmitted HPV infection in their lifetime—most infections are cleared without any complication. However, persistent infections may establish and develop into cancerous lesions. Two vaccines have been developed against the two most high-risk types, and have shown high levels of efficacy thus far. However, infections are still occurring and it is not clear why some individuals develop persistent infections while others do not. In this thesis, we develop a model to describe how the infection spreads within the host. We express the basic reproduction number $\mathcal{R}_0$, a threshold for the establishment of an infection. We solve for the diseased equilibrium, providing insight about whether an infection will persist or not. We develop a spatial model to examine how spatiality of the infection process affects the establishment or clearance. Lastly, we develop a multi-type HPV model to examine whether competitive HPV types are able to coexist in the host for different levels of competition. Ultimately, this work provides groundwork for within-host modelling of HPV and can provide direction for future research.

Acknowledgments

I would first like to thank Dr. Jonathan Dushoff, my supervisor during my masters. The influential guidance and direction he has provided has been invaluable for my success during my Master's. Dr. Jonathan Dushoff consistently pushes me and others to examine and analyze the world with a critical lens, asking the difficult questions—such as, who am I? A warm thank you to the other members of my committee, Dr. Ben Bolker and Dr. Brian Lichty: I would like to extend my gratitude for taking the time to review my thesis work and provide insight to my research. I would be remiss to not thank the rest of the Theoretical Biology Lab: I would not have made it through this process without the assistance and camaraderie of Lindsay Keegan, Chyun Shi, David Champredon, and Dave Leaman. All your help and support has helped shaped me into who I am today.

I would like to thank the graduate students in Mathematics Department who have built a warm and friendly community, making McMaster feel like home. Trivia Nights will never be the same without the crew: Tyler Meadows, Peter Sinclair, Diogo Poças, Ramsha Khan, James Rooney, and Niky Hristov. I would also like to extend a warm thank you to the biology graduate students who adopted me as one of their own. A special thank you to Irena Papst—roommate, fellow student, and friend. She has seen me at my worst, and deserves me at my best. Irena has provided support, guidance,

and friendship. She made me feel like part of her family, when my family was so far away. To all my friends abroad who supported me starting and completing my masters. I wouldn't have been able to do it without your ever-present albeit remote friendship. Thank you Nikki, Rachel, Jodie, Kelsey, and Morgan. Of course, it would be neglectful to not thank all the fantastic staff at the Main and Emerson Starbucks. They were always so hospitable to let me stay there for many hours working on my thesis. When all the baristas learn your name, you know that you have really settled in nicely. Keep the coffees coming!

And lastly, I would like to thank my family for their unending support during this process. They have always encouraged me what ever path I was on, even when they did not exactly know what it was I was doing. Don't worry, I can't catch any diseases from modelling them. Without the amazing and continuous love and support from my family, I would not have been able to get this far. For all this, I'd like to say "Thank you."

Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

Abbreviations

APC	Antigen presenting cell.
B-cell	A type of lymphocyte matured in human bone marrow.
CD4+/8+	cluster of differentiation 4/8, a glycoprotein found on the surface of the immune cells, such as T-cells.
Const.	Constant value.
E6, E7	Early viral proteins expressed in infected cells. Both have cell immortalization properties.
Est.	Estimated value.
HPV	Human papillomavirus.
IRQ	Inter-quantile range.
L1	Late viral protein, the major capsid protein.
p53	Tumour supressor protein.
Rb	Retinoblastoma protein, a tumour supressor protein.
T-cell	A type of lymphocyte matured in the thymus.
VLP	virus-like particles.

Symbols

H, S, E, I	The number of sites in the epithelium of a given state: healthy H, susceptible S, exposed E or infectious I.
V	The amount of virus in the system.
Z	The immune response in the system.
M	The number of memory immune cells in the system.
N	The size of or number of sites in the epithelium.
χ	The abrasion rate of healthy sites in the epithelium.
ρ	The recovery rate of abraded sites in the epithelium.
β	The rate of infection by virus particles.
σ	The maturation rate of exposed sites to infectious sites.
f	The rate of viral production by infectious sites.
δ	The rate of viral degradation.
ζ	The baseline rate of immune response without the presence of pathogen.
γ	The pathogen induced rate of immune cell propagation.
μ	The natural death rate of immune cells.
α	The rate of clearance of infected sites by the immune system.
ε	The propagation rate of memory immune cells.
K_M	The carrying capacity of memory immune cells.
HE	The healthy equilibrium

DE	The diseased equilibrium
$\overline{X}$	The equilibrium value of $X \in \{H, S, E, I, V, Z, M\}$.
$\mathcal{R}_0$	The basic reproduction number.
$\mathcal{R}_e$	The effective reproduction number.
$s(t)$	The immune switch function in the delayed immune response model. Take the value of 0 if the immune system is "off" and 1 if the immune system is "on".
$c_j(i)$	The rate of event j occurring at site i.
c_j	The rate of event j occurring over all sites.
c	The total rate of any event occurring.
$\mathcal{N}(i)$	The number of sites in the neighbourhood (global or local) surrounding site i.
$\mathcal{N}_X(i)$	The number of sites with state X in the neighbourhood (global or local) surrounding the site i. $X \in$ {H,S,E,I}.
$Z_i(t)$	The immune response at site i at time t.
E_1, E_2	Exposed sites of either strain 1 or strain 2, respectively.
I_1, I_2	Infected sites of either strain 1 or strain 2, respectively.
V_1, V_2	Virus particles of strain 1 or strain 2, respectively.
Z_1, Z_2	Cell-mediated immune response primed against strain 1 or strain 2, respectively.
q	The level of cross-reactivity between the two strains, $q \in [0, 1]$

$\overline{\rho}$	Spectral radius
DF_x	The Jacobian evaluated at x.
x	Susceptible patches in generic super-infection patch model.
y_i	Singly infected patch with strain i, $i \in \{1, 2\}$.
z	Coinfected patches
Λ_i	Overall force of infection of strain i.
β_i	Rate of infection by virus particle of strain i.
p_i	Proportion of virus of strain i produced in coinfected patches.
$\overline{\mathcal{R}}$	Boundary equilibrium reproduction number.
z_i	Coinfected patches with the same strain i twice.
z_{ij}	Coinfected patches first infected with strain i then infected by stain j.

Chapter 1

Introduction

Cervical cancer is a major health concern worldwide. With over 500,000 cases and approximately 260,000 deaths each year, the burden of cervical cancer is significant [6, 18, 20]. In particular, women in developing countries are at a higher risk for developing cervical cancer and may have limited access to health care for detection and treatment [6, 18, 20]. Since the 1990s, it has been known that all cases of cervical cancer are caused by persistent infection with the human papillomavirus (HPV) [2,25]. This virus is also highly associated with other cancers affecting the anogenital and oropharyngeal tracts [49]. Furthermore, HPV infections are very common in almost all populations. HPV is the most common sexually transmitted infection, and it is believed that most sexually active men and women will have at least one HPV infection at some point in their life [49]. At a 80-90% clearance rate, most infections are naturally cleared without any complications or symptoms [2, 33, 38, 57]. It is those with persistent infections who are at risk of developing cervical cancer [25, 47]. Therefore, researchers are trying to learn more about HPV to prevent infection and to prevent the progression to cervical cancer and other cancers.

The human papillomavirus is a virus that infects the epithelium. In particular it infects the basal cells in the bottom-most layer of the epithelium [24,51,52]. The virus is able to utilize the DNA replication procedures of these basal cells to replicate its own viral DNA. HPV is only able to infect these cells after they have become uncovered by some abrasion to the epithelium. In the case of sexually transmitted HPV, this may occur during intercourse or sexual activity. Once the HPV virus particle infects the basal cell, it begins replicating its viral DNA. As the infected epithelial cells move up the epithelium, virus particles are assembled [51, 52]. When the cell reaches the top of the epithelium, it naturally undergoes cell death and flattens to form the top-most layer of the epithelium [51, 52]. During this cell death process, virus particles are released into the surrounding milieu and the infection cycle continues.

In fact, HPV is able to delay the natural cell death process in the epithelium. Two viral proteins E6 and E7 have been shown to inhibit proteins p53 and Rb, respectively. The protein p53 is a tumour suppressant. It halts the DNA replication process of damaged DNA until the damage is fixed. If it isn't fixed, it can also induce apoptosis, destroying the cell that would produce the damaged DNA [48]. The viral protein E6 promotes the degradation of p53, which inhibits the tumour suppressant ability [63]. It has also been shown that the E6 protein can inhibit p53 without actually destroying it [30]. The other protein Rb has a major responsibility for initiating the first check point in the cell cycle [22]. It is able to repress the expression of replication enzyme genes which suppresses tumour development [63]. The viral protein E7 binds to the Rb protein and prevents the protein from initializing the check point. This supports excessive cell growth [63]. Thus, these viral proteins E6 and E7 have an important role in immortalizing infected cells, thus producing more viral copies overall. This

process also may lead to pre-cancerous or cancerous lesions. As the life span of the cell increases, the probability for a cancerous mutation to occur also increases. HPV types that produce proteins which interact effectively with these tumour suppressant proteins p53 and Rb are said to be cancer causing and high-risk, while those that do not interact well are considered low-risk. There is a lot of variability between HPV types and the impact they have on the host.

There are over 100 different HPV types, over 40 of which are sexually transmitted and infect the anogenital tract [11,25,57]. Types are categorized into different species and genera. Types in the α-9 and α-7 species are highly associated with malignant lesion formation on the cervix [2,11,25,57]. In particular, HPV-16 (α-9) and HPV-18 (α-7) are associated with 70% of all cervical cancer cases in women [6, 25]. Types highly associated with cancer formation are considered "high-risk." Other high-risk types include 31(9), 33(9), 35(9), 39(7), 45(7), 51(5), 52(9), 56(6), 58(9), 59(7), 68(7), 73(11), and 82(5) [2,11]. The numbers in the parentheses refer to the species in which the type belongs; all species are within the α genus. To combat the effects of HPV on cervical cancer, two vaccines were developed to protect against the two most high-risk types.

Since 2006, two vaccines Gardasil$^{\text{TM}}$ (GlaxoSmithKline) and Cervarix$^{\text{TM}}$ (Merck & Co.) have been administered to protect against the two most high-risk HPV types [25]. Cervarix$^{\text{TM}}$ is a bivalent vaccine. It protects against two HPV types: HPV-16 and HPV-18. Gardasil$^{\text{TM}}$ on the other hand is a quadravalent vaccine. In addition to providing protection against HPV-16 and -18, it also confers protection against HPV-6 and -11, two low-risk types associated with genital warts. Thus far, both vaccines have shown exceptional efficacy, up to 95% over eight years, and participants have

developed high antibody responses to the vaccine [31, 43]. While these vaccines are still in their infancy, preliminary longitudinal results show that antibody levels remain high. There is even evidence that vaccination against these types provides protection for unvaccinated individuals through herd protection [29]. Given these findings, it is believed that these HPV vaccines will have a significant effect in the reduction of HPV caused cancers, specifically cervical cancer.

The advent of the HPV vaccines has led many researchers to examine the impact the vaccine will have on cancer cases and the burden of HPV overall. Because certain data are limited or because experiments and studies may be impractical or impossible, many researchers employ mathematical models to analyze the effects the vaccine and to make informed decisions on the vaccine coverage and implementation. Some models examine epidemiological and ecological benefits of the vaccine [9, 15, 16, 26] while others examine the effects of vaccinating certain populations, such as boys and men [14, 26]. Some models even aim to understand factors for vaccine acceptance and uptake in the population [3]. These models suggest that the vaccination of HPV will provide numerous benefits epidemiologically and economically. Many scientists are in consensus that the vaccines will provide positive outcomes to control of HPV-related cancer. However, some researchers have also been examining the potential negative effects of the vaccine—in particular, type replacement.

Type replacement is the ecological phenomenon wherein the removal of one strain of a pathogen, such as through vaccination, can increase the niche space for other strains, thus increasing the prevalence of the non-vaccine strains. In the context of HPV, these researchers set out to explore if the removal of HPV-16 and -18 through

vaccination can increase the prevalence of other high-risk types. If these other high-risk types do indeed increase in prevalence, then the benefits of the vaccine may not be as substantial as previously thought. In order to examine if type replacement is even possible, mathematical models have been used to analyze conditions for type replacement to occur.

Mathematical models are important tools for examining concerns such as type replacement because information can be gathered about the potential for and impact of type replacement before it occurs in the population. Compartmental models have been developed and analyzed specifically to address the potential for type replacement in HPV [16, 40, 41]. Findings from these models suggest that type replacement is contingent upon the interactions between HPV types. Specifically, competition between the HPV types is a necessary condition for type replacement [16, 28]. If HPV types have neutral or facilitative interactions, then type replacement will not occur. In fact, if HPV types are facilitative, in that the presence of one supports another type, then a further decrease in the prevalence of the non-vaccine type may accompany vaccination. Therefore, in order to determine if type replacement will be a concern, interactions between HPV types will have to be examined further.

Interactions between HPV types within the host are complex and not fully understood. Population studies have shown that infections with multiple types of HPV are not uncommon. It is estimated that between 5-43% of individuals with HPV infections are infected with multiple types [7, 25]. This high rate of type coexistence within the host has led some researchers to suspect that HPV types are not competitive [7]. If HPV types occupy the same niche space, then by the Competitive Exclusion Principle two competitive organisms should not be able to occupy

the same niche space. Because multiple HPV types are found in a high number of HPV infections, some researchers have concluded that HPV types are predominantly independent [56]. Therefore, they believe that type replacement is unlikely given this scenario for type interactions. While there is limited data that would suggest type replacement is occurring, there is one study [29] which showed that there was an increase in non-vaccine HPV types in vaccinated young women. This may be a sign of type replacement; however, it may also be confounding effects due to their sample population. A consensus has not yet been reached in terms of the potential or severity of type replacement, but it has been shown that type interactions play an important role and should be considered more carefully.

To begin exploring HPV type interactions, we first discuss what is known about how HPV types compete with one another, whether for space or indirectly. Two HPV types may have indirect apparent competition between them through immune system cross-protection. That is, if the immune system has cross-protective effects between two types, then the presence of one may drive the elimination of the other through increased immune activity. It has been shown that type specific antibodies have cross-protective effects on HPV types in the same species [46]. While less is known about whether cross-recognition of similar HPV types by the CD4+ and CD8+ T-cells exists and if it confers noticeable levels of cross-protection, evidence for some cross-recognition is supported [36]. There is some evidence for coinfection within the same host cell. An *in vitro* study showed that coexistence is possible between HPV-18 and other HPV types [32]. However, in coinfection scenarios, HPV-18 often dominated the viral production of the host cell. This suggests resource competition. It seems there are still many questions that are unanswered surrounding HPV type

interactions, and it is important to consider the within-host dynamics that are going on between HPV types.

To help elucidate some of the questions surrounding HPV interactions and the potential for type replacement, a mathematical within-host patch model was built and analyzed by Murall et al. [35]. In their paper, they summarize what is currently known and speculated about HPV interactions. Then they analyze long-term results of the mathematical models under a variety of different HPV interaction scenarios. They then compare these results with qualitative characteristics of population level studies looking at multi-type infections. They determined that even in scenarios of competition, coexistence is possible in the host. This suggests that even though multiple type HPV infections are not uncommon, HPV types may still be competitive within the host. From these findings, HPV type replacement should not be yet ruled out.

However, a traditional multi-type patch model does not fully capture the complex infection cycle of HPV. In this thesis, we start by discussing some of the important biological factors at play in the infection process in Chapter 2. We then implement these important factors in the formation of our base within-host HPV model. In particular we introduce an abrasion process into our model which may drive or limit the spread of the infection in the host. We solve for the equilibria of the model and examine their stability. We also formulate an expression for the basic reproduction number $\mathcal{R}_0$, which provides a condition for the effective establishment of an infection within the host and also informs potential for clearance.

In Chapter 3 of this thesis, we focus on how spatiality may affect the spread of HPV within the host. We develop a stochastic, spatial model derived from the base

model established in Chapter 2. We consider two different neighbourhood structures: a global and local neighbourhood. In the global neighbourhood model, we consider that infection sites in the epithelium are able to equally interact with all other sites in the epithelium. This scenario is considered to compare the stochastic simulation to the deterministic base model, and to confirm the rates of the events in the model. The local model considers sites only interacting with the four closest sites surrounding the focal site. This is to examine how locality on the epithelium may play a role in the establishment and clearance of HPV interactions.

Lastly, in Chapter 4 we discuss in more detail some of the different ways HPV may compete within the host. We adapt our base model to consider two different HPV types. We examine a scenario of space competition and cross-reactivity. We consider independent immune responses primed against each of the two strains linked together by a factor of cross-reactivity. In this scenario, we find that coexistence between the two competing types is possible for certain levels of cross-reactivity. This further suggests that coexistence of multiple HPV types within the host is possible even when HPV types are competing with one another.

In conclusion, we show that there are complex dynamics in within-host HPV modelling. Coexistence is possible in within-host models when considering two separately primed but cross-reactive immune responses. We also highlight unintentional asymmetry in super-infection patch models and some techniques in avoiding this asymmetry. This research aims to discuss and highlight particular considerations in within-host modelling and provide a basis for further research in within-host dynamics and the potential for type replacement of HPV.

Chapter 2

Within Host Models for HPV

2.1 Introduction

Cervical cancer is the fourth most common cancer in women and the seventh overall [18]. Fatality rates due to cervical cancer vary nationally but are higher in developing nations [6, 18, 20]. Persistent infections with the human papillomavirus (HPV) have been linked with 99% of all cervical cancer cases [2, 25]. HPV infects the epithelium and can cause warts or lesions. There are over 100 different types of HPV, over 40 of which are sexually transmitted and infect the anogenital tract. The human papillomvirus is also ubiquitous in most populations: 75% of Canadians will have an HPV infection in their lifetime [37]. However, approximately 80-90% of people naturally clear HPV within two years of infection [2, 33, 38, 57]. It is the remaining 10-20% of individuals with persistent HPV infections who are most at risk of developing cervical cancer or other types of cancer.

Natural infection with HPV is marked by a relatively weak immune response.

Many of those who naturally clear a transient HPV infection do not produce a detectable antibody response [2,5]. Of those who do present detectable antibody levels, concentrations are often low [2,5,51]. Because transient infections are cleared without any noticeable antibody levels, it is believed that the cell-mediated immune response accounts for the clearance of HPV infections [51, 52, 54]. It is still not fully understood why certain cases result in transient infection and others result in persistent infection [24]. Understanding the complex interactions between the host, virus, and the host's immune response can be helpful for elucidating these open problems and potentially developing treatments. Mathematical models can be useful for examining these questions. Insights from current biological knowledge of HPV dynamics can be used to develop theoretical models, which can be analyzed in a rigorous fashion.

In this chapter, we review the biological factors that drive HPV infections and from this develop a within host model for HPV infection. We analyze the values of the equilibria and their stability from this model. Importantly, we formulate an expression for the basic reproduction number $\mathcal{R}_0$, which can be used to better understand factors that lead to clearance, persistence, and prevention. We make two alterations to the model by considering the development of memory cells and by introducing a delay in the immune response. Lastly, we examine how these alterations affect the dynamics of the model.

2.1.1 Biological Factors

To build a model of HPV infection within the host, we first consider the biological factors that drive the infection process. We will discuss how HPV infects the host, how viral DNA is replicated and released in the host, how HPV is detected by the

immune system, and how HPV is cleared by the immune system. We then use this information to inform the model and underlying assumptions, discussed in section 2.2.

HPV solely infects the basal layer of the epithelium, the bottom-most layer [24, 51, 52]. These cells are responsible for cell DNA replication, and thus HPV is able to capitalize on the DNA replication processes in the cell to replicate its viral DNA. The cuboidal cells that make up the basal layer are normally protected by the squamous layer (the top-most layer) and stratum spinosa (the middle layer) of the epithelium, but they can be uncovered by micro-abrasions to the epithelium [24, 51]. These micro-abrasions are caused by a trauma to the epithelium; in the case of sexually transmitted HPV infections this occurs during sexual intercourse. HPV then infects these uncovered basal cells, and begins the infection cycle.

After HPV has infected these cells, it begins to produce viral proteins necessary for DNA replication and the construction virus particles. The HPV infection cycle is intrinsically linked to the epithelial life cycle [51, 52]. Different viral proteins are expressed at different times in the epithelial cell cycle. Early in the cycle, while the infected cells are still near the basal layer, early proteins (E) are produced. As the infected basal cells move up the epithelium, late proteins (L) are expressed in higher numbers [51, 55]. There are two important late proteins L1 and L2, which are the major and minor capsid proteins, respectively. They build the outer capsid of viral particles and are responsible for the implantation of viral DNA into uninfected, uncovered basal cells [51,54,55]. Virus particles are released at the end of the epithelial cell cycle. When epithelial cells approach the squamous layer, they die and flatten into the scale-like cells that make up the top-most layer of the epithelium. As these infected scale-like cells flake off, viral particles are released into the surrounding milieu

and can infected other susceptible sites, completing the infection cycle of HPV.

In order to adequately explain how the immune system interacts with HPV, we will first briefly discuss a simplified description of the immune system. The human immune system is divided into two main responses: the innate immune response and the adaptive immune response. The innate immune response is the first line of defence against potential pathogens [51]. It has some protective effects; however, a main function of the innate immune response is to activate the adaptive immune response, which is done in part via inflammatory responses. Antigen presenting cells (APCs) are responsible for detecting, processing, and presenting antigen to the immune system in order to elicit antigen-specific immune responses [51]. APCs are triggered and activated during an inflammatory response. After being activated, APCs interact with antigen epitopes, migrate to the lymph nodes, and present them to naive T-cells there [51]. This initiates the adaptive immune response.

The adaptive immune response develops antigen specific strategies to eliminate an infection and prevent reinfection. It is separated into to cell-mediated immune responses and humoral immune responses. We briefly discuss the main roles of each response in the clearance and prevention of infection. Primed T-cells are the main effector cells of the cell-mediated immune response and have various functions based on how they differentiate. Once T-cells become primed against a particular antigen, some differentiate into helper T-cells, which aid the immune response in a variety of ways, and into killer T-cells [51]. Killer T-cells are primed to detect viral proteins expressed on the surface of infected cells, and then subsequently destroy these infected cells [51]. In this way, primed killer T-cells eliminate the current infection.

The other arm of the adaptive immune response, the humoral response, has B-cells as its effector cells. Unlike T-cells, which require antigen to be presented via ABCs, naive B-cells are able to interact directly with antigen. B-cells, which have been primed against a particular antigen, produce and release antibodies into the surrounding milieu [51]. These antibodies then interact with virus particle, deactivating them and preventing further infection. In this way, the humoral response disallows the infection from spreading further. Another role of the B-cells is to differentiate into memory cells. These memory cells live for a long time within the host and help to jump start the immune process when the host comes into contact with the same antigen at a later time. Together, the cell-mediated and humoral responses of the adaptive immune system are often very effective at clearing and preventing further infection. Understanding how the immune system works in general can help provide insight to how HPV interacts with the immune system.

The first way that HPV interacts with the immune system is avoiding it. HPV is quite effective at evading the immune system, which in turn results in a fairly weak immune response. As discussed previously, HPV virus particles are released when the epithelial cells flatten and die at the end of the epithelial cycle. In this way, HPV does not need to induce cell death in order to release viral particles; the cells die naturally [51, 52, 55]. This delays an inflammatory response, which in turn delays the adaptive immune response. Moreover, HPV locally infects the epithelium, which is an immune-privileged zone. The epithelium does not have much immune activity, which makes it difficult for immune responses to be triggered in the presence of antigen [51, 52, 55]. In these ways, HPV can impede the immune response, which limits the immune response against HPV. Even though HPV is effective at evading

the immune response, in most infections a response is triggered and often HPV is effectively cleared.

Once the adaptive immune response is triggered, primed T-cells and B-cells work together to clear the HPV infection. Killer T-cells primed against the HPV antigen detect the early proteins E2 and E6, expressed on the surface of the infected basal cells [10, 17, 21] and eliminate the infected cells. B-cells are primed to produce antibodies by coming into contact with L1 viral protein epitopes [51, 53, 54]. Antibodies are produced and interact with and deactivate the viral L1 protein on the outer capsid, which prevents further implantation of viral DNA into susceptible cells [51, 54]. Unfortunately, because HPV often elicits a weak immune response, antibody concentrations are often quite low in natural infections [24, 51, 54]. Some individuals do not even acquire detectable antibody levels after infection [55]. Because of these low antibody levels, and the high rates of eventual HPV clearance, it is believe that the cell-mediated response is primarily responsible for clearing HPV infections [51,52,54].

In order to combat the burden of HPV and HPV induced cancers, two vaccines have been developed. Gardasil^{TM} (Merck & Co.), a quadravalent vaccine, protects against four types of HPV: 16, 18, 6, and 11. The first two types are the most highly associated with cervical cancer, while the other two are most associated with genital warts. The second vaccine Cervarix^{TM} (GlaxoSmithKline) is bivalent and protects against HPV-16 and -18. These vaccines work in similar ways. Virus like particles (VLPs), which contain L1 capsid proteins but no viral DNA, are injected directly into muscle tissue. Because the vaccine is injected directly into the host, this puts the L1 protein into direct contact with the immune system, which activates strong humoral and cell-mediated immune responses, preventing later infection [51, 54]. Three doses

of the vaccine produce high antibody concentrations, which after follow-up remain high in the patients [31, 43]. These vaccines show high efficacy in preventing infection with these specific HPV types, thus lessening the burden of cervical cancer and other cancers caused by these HPV types.

2.2 The Model

Our model considers sites on the cervical epithelium. One site is essentially the resulting infectious portion of the epithelium that develops after infection of one basal cell unit. We organize the sites into different compartments based on infection status. The compartment denoted H refers to healthy sites, S refers to susceptible sites which have been uncovered by an abrasion, E refers to exposed sites which have been infected but are not yet producing virus, and I refers to sites that have reached the infectious stage of the viral cycle. These infectious sites produce and release virus particles V at rate f. These virus particles are naturally destroyed at a rate δ. Virus particles are able to infect susceptible sites at a rate βV. Healthy sites are abraded into susceptible sites at a rate χ and these susceptible sites recover back to healthy sites at a rate ρ. Once a site has become exposed with HPV, it matures into an infectious site at a rate σ. This completes the infection cycle of HPV.

We also introduce an immune response through killer T-cells Z, which will be triggered by the presence of exposed and infectious cells. Production of immune cells occurs at a constant rate ζ without the presence of an infectious agent. Propagation of these cells due to the presence of HPV infection will depend upon the current number of immune cells Z and the combined number of infected cells $(E + I)$. This process occurs at a rate of $\gamma(E+I)$. These T-cells can effectively clear infections from

cells in the E and I compartments at a rate of αZ and have a natural death rate of μ. The entire process of this model is illustrated in Figure 2.1, a compartmental diagram highlighting the infection and clearance processes through flows.

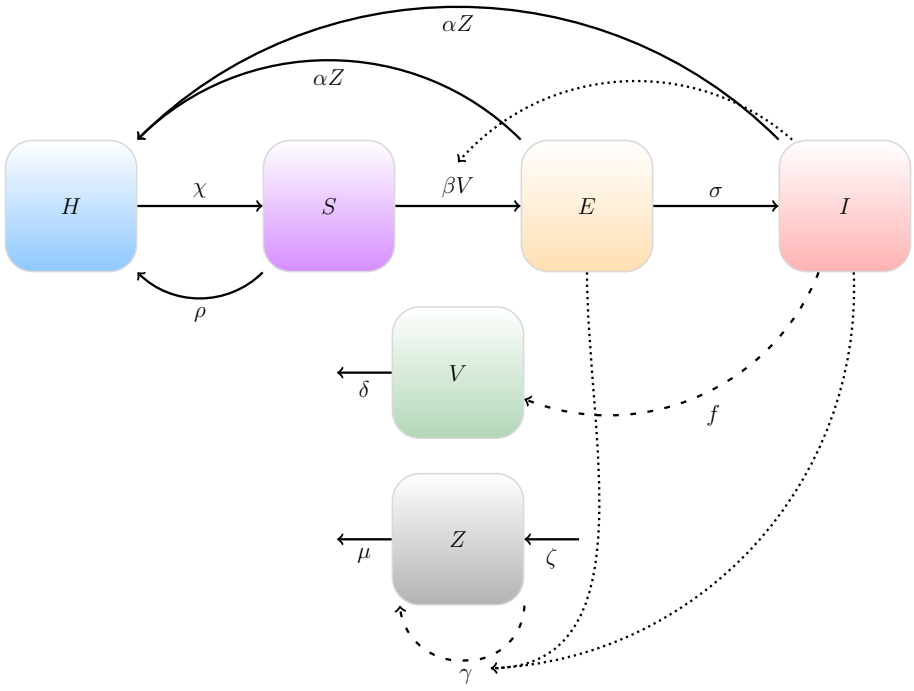

Figure 2.1: Flow diagram of the infectious process of HPV within the host.

This model is also expressed as a system of differential equations (1) below.

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\chi H + \rho S + \alpha Z\frac{(E+I)}{N}, \tag{1a}$$
$$\frac{\mathrm{d}S}{\mathrm{d}t} = \chi H - \rho S - \beta V\frac{S}{N}, \tag{1b}$$
$$\frac{\mathrm{d}E}{\mathrm{d}t} = \beta V\frac{S}{N} - \sigma E - \alpha Z\frac{E}{N}, \tag{1c}$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E - \alpha Z\frac{I}{N}, \tag{1d}$$
$$\frac{\mathrm{d}V}{\mathrm{d}t} = fI - \delta V, \tag{1e}$$
$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \zeta N + \gamma Z\frac{(E+I)}{N} - \mu Z, \tag{1f}$$

To simplify the analysis, we set $N = 1$ and consider H, S, E, and I to be the proportion of healthy, susceptible, exposed, and infectious sites, respectively.

Parameter Estimation

We begin by discussing the various biological parameters of the model and estimates of these parameters from the existing literature. There are a number of biological parameters in the system that require estimation from previous physiological and virological studies. Please see Table 2.1 below to review the various parameters that build up the model:

Firstly, consider the abrasion-recovery process. We assume a rate of abrasion $\chi = 0.015$. This is an estimate that we developed. We assume that in one occurrence of sexual intercourse, 10% of the basal cells become uncovered due to abrasion. If individuals have sexual intercourse at a rate of $1/7$ (≈ 0.15) times per day on average, the rate of abrasion is $\chi = 0.10(0.15) = 0.015\,\mathrm{day}^{-1}$. We estimate the rate of recovery ρ from the average cycle of the epithelium, approximately 30 days. Now, we only

Table 2.1: Estimates for the parameters used in the models

Host Parameters	**Description**	**Value**	**Source**
χ	Abrasion rate of the epithelium	0.015 day^{-1}	Est.
ρ	Recovery rate of the epithelium after abrasion	0.6 day^{-1}	Est.
ζ	Rate of baseline immune cell production	0.01 day^{-1}	Est. [42]
γ	Rate of immune cell production in the presence of HPV viral proteins	2day^{-1}	[4]
μ	Rate of immune cell death	0.5 day^{-1}	[4]
ε	Rate of memory cell proliferation	0.02 day^{-1}	[4]
K_M	Carrying capacity of memory cells	$0.01N$ cells	[4]
N	Site population size of the organ	10,000 sites	Const.
Virus Parameters	**Description**	**Value**	**Source**
f	Rate of virus production	600 $\frac{\text{copies}}{(\text{site}\cdot\text{day})}$	[51]
δ	Natural rate of viral destruction	0.138 day^{-1}	[39]
σ	Rate at which infectious sites can start producing virus	0.03 day^{-1}	Est.
Host-Virus Parameters	**Description**	**Value**	**Source**
α	Killing rate of infected sites due to the immune response	0.5 day^{-1}	[35]
β	Infection rate of susceptible sites by virus particles	0.003 day^{-1}	[8]

require that only layer of the epithelium is recovered to shield the basal layer from infection. One layer of cells of the epithelium is approximately 10 micrometres of the 180 micrometres of the whole epithelium. Thus the average time of recovery for one layer of the epithelium is approximately $(10/180)(30) = 1.67$ days, and thus $\rho = 0.60\,\text{day}^{-1}$.

We also consider the estimates of HPV virus parameters. HPV requires the complete epithelial cell cycle in order to produce and release viral particles. Therefore, we estimate that the average time from an exposed site to become infectious is $\frac{1}{\sigma} = 30$ days, and thus $\sigma = 0.03\,\text{day}^{-1}$. The rate of viral production can be determined from the amount of viral copies produced from the squamous layer per day. The burst size of a single infectious site is 1000 copies per cell [51], and we multiply this by the rate of layer recovery, ρ, thus the rate of viral production is $f = 1000(0.6) = 600$ copies per cell per day. The natural decay rate of of these viral particles is set to 0.138 day^{-1} [39]. These viral particles can then infect uncovered basal cells at a rate 0.003 day^{-1} [8].

There are not many studies examining specific immune parameters during an HPV infection. Thus we use studies from other viral studies to inform our parameter estimates. A study by De Boer et al. [4] examined proliferation and apoptosis rates of CD8+ T-cells in response to Lymphocytic Choriomeningitis viral infections. They found that T-cells lived on average for 2 days before apoptosis, *i.e.,* $\mu = 0.5\,\text{day}^{-1}$. The proliferation rates of CD8+ T-cells occurred about about three per day, that is $\gamma \approx 3\,\text{day}^{-1}$. However, proliferation rates can vary between individuals. Furthermore, HPV is not particularly immunogenic, so we can assume that $\gamma = 2\,\text{day}^{-1}$. When considering the rate at which primed T-cells are able to clear HPV, it is not completely clear. Because the epithelium is an immune privileged zone, we set the baseline immune response ζ to be quite low, $\zeta = 0.01$ day^{-1} [42].

We also consider memory cell dynamics in the immune system. The same study by De Boer et al. [4] found that the proliferation rate of memory cells was about $\varepsilon = 0.01\,\text{day}^{-1}$. Furthermore, the capacity of the memory cells recruited during an

infection was found to be about 5% of the peak population of T-cells [4].

2.3 Results

In this section, we solve for the equilibria of the model (1) and also derive the expression for $\mathcal{R}_0$. We also showcase the numerically solved solutions of the model and discuss the dynamics of the model for different parameter values. In sections sections 2.4 and 2.5 we introduce memory cells and an immune system delay process into the model and discuss how these effects alter the dynamics.

2.3.1 The Healthy Equilibrium, *HE*

We will first consider the case with no infection present, which we call the healthy equilibrium (HE). The equilibrium is expressed as $HE = (\bar{H}, \bar{S}, \bar{E}, \bar{I}, \bar{V}, \bar{Z}) = (\bar{H}, \bar{S}, 0, 0, 0, \bar{Z})$. Given this equilibrium condition, we can find the remaining equilibrium values for the healthy and susceptible patches and the immune response. This is done by setting the differential equations to zero, and solving for the remaining H, S, and Z terms in the steady states.

$$\overline{Z} = \frac{\zeta}{\mu}, \tag{2a}$$

$$\overline{H} = \frac{\rho}{\chi + \rho}, \tag{2b}$$

$$\overline{S} = \frac{\chi}{\chi + \rho}. \tag{2c}$$

These values provide the final expression for the healthy equilibrium HE:

$$HE = \left(\frac{\rho}{\rho+\chi}, \frac{\chi}{\rho+\chi}, 0, 0, 0, \frac{\zeta}{\mu}\right) \tag{3}$$

The stability of the healthy equilibrium is examined in more detail in Appendix A.2.

2.3.2 The Basic Reproduction Number, $\mathcal{R}_0$

An important value in epidemiology is the basic reproduction number $\mathcal{R}_0$. The value of $\mathcal{R}_0$ is the average number of new infections from one infected individual in a fully susceptible population at the beginning of an outbreak. This parameter has many important implications for the control of infectious disease. Intuitively, if a system has an $\mathcal{R}_0 < 1$, then fewer than one individual will be infected on average at the onset of the disease, which means the disease cannot spread effectively through the population, resulting in no epidemic. In the case where $\mathcal{R}_0 > 1$, the disease is able to infect more than one individual on average, and an epidemic can occur. The severity of the epidemic and the magnitude of the control efforts to combat the disease depend on the value of $\mathcal{R}_0$. In the context of our model, we are not concerned with individual humans but rather sites in the epithelium. A complete infectious cycle can be thought of as a single infectious site I producing viral particles, these viral particles successfully infecting a susceptible site, and that newly exposed site E surviving latency to become infectious again. Thus in the context of our model, $\mathcal{R}_0 < 1$ means the HPV infection cycle is not completed on average and an infection is unable to establish. For $\mathcal{R}_0 > 1$ the infection cycle is completed more than once on average, and thus the infection may be able to establish and persist. This $\mathcal{R}_0$ value signifies the within-host reproduction number and only informs the establishment of

an infection after exposure to HPV, but does not inform the spread of HPV between individuals at the population level.

We formulate an expression for $\mathcal{R}_0$ in more detail by closely examining the infection cycle of a singly infectious site at the beginning of the epidemic. Recall that the infectious site I is able to produce viral particles at a rate f, and it is cleared by the existing immune system at rate $\frac{\alpha\zeta}{\mu}$. Thus, the average number of viral particles produced before this site is cleared is equal to $\frac{f\mu}{\alpha\zeta}$. These virus particles are able to infect susceptible sites at a rate β and are cleared naturally at a rate δ. Because the proportion of the susceptible sites at the beginning of the epidemic is $\frac{\chi}{\chi+\rho}$, these virus particles infect an average of $\frac{\beta\chi}{\delta(\chi+\rho)}$ susceptible sites. These newly exposed sites become infectious at a rate σ, and are also cleared at a rate $\frac{\alpha\zeta}{\mu}$, thus the average number of exposed sites that survive latency is $\frac{\sigma\mu}{\sigma\mu+\alpha\zeta}$. This cycle is illustrated in more detail as a flow diagram in Figure 2.2.

Multiplying the terms, we arrive at the following expression for $\mathcal{R}_0$.

$$\mathcal{R}_0 = \left(\frac{f\mu}{\alpha\zeta}\right)\left(\frac{\beta\chi}{\delta(\chi+\rho)}\right)\left(\frac{\sigma\mu}{\sigma\mu+\alpha\zeta}\right) \tag{4a}$$

$$= \frac{f\beta\chi\sigma\mu^2}{\alpha\zeta\delta(\chi+\rho)(\sigma\mu+\alpha\zeta)} \tag{4b}$$

The derivation of this value is also confirmed using van den Driessche and Watmough's method of the next generation matrix [59] discussed in detail in Appendix A.1. Thus given the formulation of $\mathcal{R}_0$ the healthy equilibrium HE is stable when $\mathcal{R}_0 < 1$ and unstable otherwise. When applied to the parameters that were estimated and gathered from literature, we find that $\mathcal{R}_0 = 23.86$. It should be noted that this is

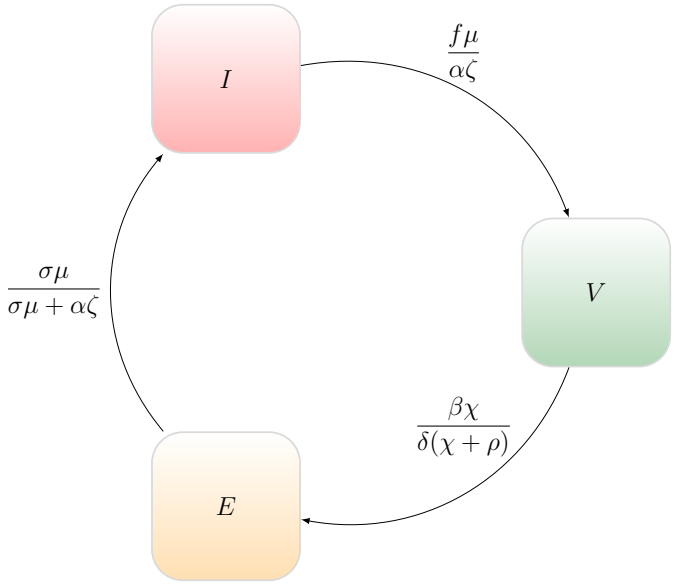

Figure 2.2: A flow diagram illustrating the infection cycle of HPV: a singly infectious site I produces virus particles to infect a susceptible site. This newly exposed site then survives latency to become infectious, completing the cycle. This process is used to formulate the expression for $\mathcal{R}_0$.

a hypothesized and imprecise estimate; however, it does provide a justification as to why HPV is able to infect so many individuals after being exposed to HPV. It is not an expression for the between-host reproduction number, and thus cannot inform how an infection will spread through the population.

2.3.3 The Diseased Equilibrium, *DE*

We now consider the case when the disease is effectively able to establish an infection in the host. We refer to this state as the diseased equilibrium (DE), *i.e.,* when $\overline{E}, \overline{I}, \overline{V} \neq 0$. Setting each of the differential equations above to equal zero, we solve

449 for the disease equilibrium using Maple.

$$\overline{H} = \frac{\overline{Z}^2\alpha\mu + \overline{Z}(\gamma\rho - \alpha\zeta - \mu\rho) + \rho\zeta}{(\chi+\rho)\gamma\overline{Z}} \tag{5}$$

$$\overline{S} = \frac{-\overline{Z}^2\alpha\mu + \overline{Z}(-\chi\mu + \chi\gamma + \alpha\zeta) + \chi\zeta}{(\chi+\rho)\gamma\overline{Z}} \tag{6}$$

$$\overline{E} = \frac{\sigma(\mu\overline{Z} - \zeta)}{\gamma(\alpha\overline{Z} + \sigma)} \tag{7}$$

$$\overline{I} = \frac{\sigma(\mu\overline{Z} - \zeta)}{\gamma\overline{Z}(\alpha\overline{Z} + \sigma)} \tag{8}$$

$$\overline{V} = \frac{\overline{I}f}{\delta} = \frac{f\sigma(\mu\overline{Z} - \zeta)f}{\delta\gamma\overline{Z}(\alpha\overline{Z} + \sigma)} \tag{9}$$

450 The solution for $\overline{Z}$ is a root of the cubic:

$$P(Z) = Z^3(\alpha^2\delta\gamma(\chi+\rho)) + Z^2(\alpha\sigma(\beta f\mu + (\chi+\rho)\delta\gamma)) + Z(\beta f\sigma(\chi(\mu-\gamma) - \alpha\zeta)) - \beta\chi f\sigma\zeta \tag{10}$$

451 This cubic is difficult to solve for explicitly, but we examine the shape of the poly-
452 nomial in order to learn more about the roots, and thus the equilibrium. In Ap-
453 pendix A.3, we show that there is only one positive root of $P(Z)$ when $\mathcal{R}_0 > 1$, which
454 shows that the diseased equilibrium DE is only biologically relevant when $\mathcal{R}_0 > 1$.

455 Finally, we visualize the healthy equilibrium and diseased equilibrium, by solving
456 the system of ODEs numerically in **R** for different values of $\mathcal{R}_0$. Considering Fig-
457 ure 2.3, we see that for when $\mathcal{R}_0 < 1$ the infection dies out, whereas when the value
458 of $\mathcal{R}_0 > 1$, the system reaches the diseased equilibrium.

459 We also examine how the values of the diseased equilibrium change as the param-
460 eters change. We plot the solutions for the value of the equilibrium as a function of
461 the parameter values. We scale the equilibrium value $\overline{V}$ by the maximum value of

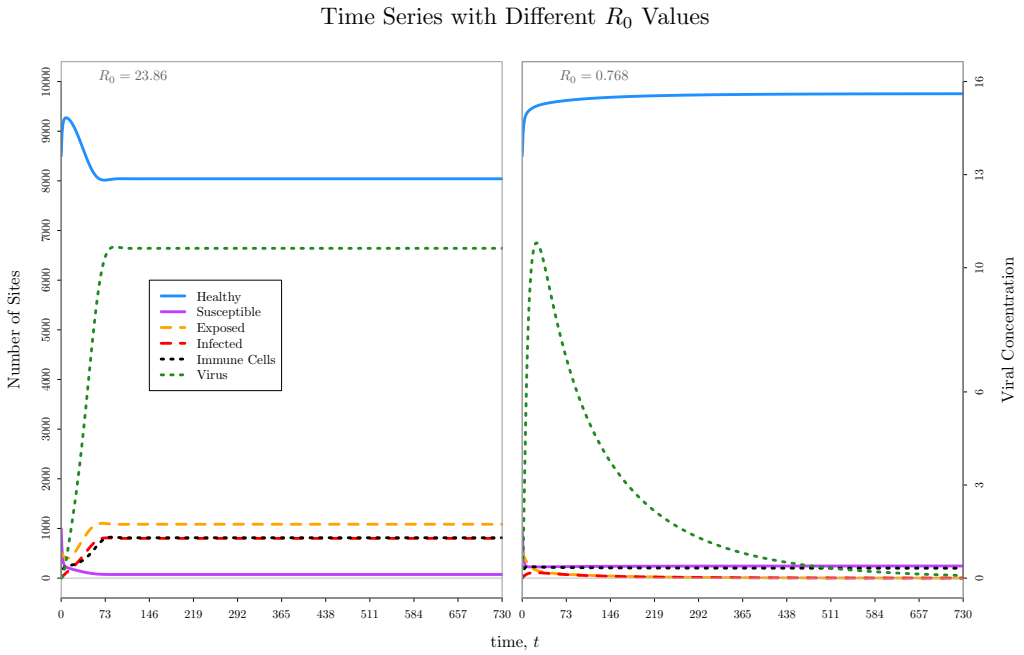

Figure 2.3: We consider the basic within host model for HPV for two different $\mathcal{R}_0$ values (varying δ and α in this case). We see that in the first panel for $\mathcal{R}_0 = 23.86$ the infection establishes and persists, reaching the DE. In the second panel, we have that $\mathcal{R}_0 = 0.768$, and the infection dies out and approaches the HE.

$\overline{V}$ explored in order to show the dynamics on the same plot, as viral load is much higher in magnitude than the proportion of infected cells. In these plots, the values of $\mathcal{R}_0$ are also shown as a function of the parameter in question. We first consider how the rate of killer T-cell propagation γ affects the system. The expression for $\mathcal{R}_0$ does not include this parameter, and thus $\mathcal{R}_0$ does not change if γ changes. This is observed because $\mathcal{R}_0$ is a measure of how a pathogen will spread in an infection-naive individual. Thus, the propagation rate due to presence of HPV will not have an effect on whether an infection will establish. It does have important implications for clearance, however. As γ increases, the the values $\overline{E}$ and $\overline{I}$ decrease, illustrated

in Figure 2.4. If these values are low enough, the infection may be cleared due to stochastic effects. This follows from the hypothesis that HPV is cleared predominantly by the cell-mediated immune response. Variability between these propagation rates may explain why some individuals are able to clear infections naturally while others sustain persistent infections.

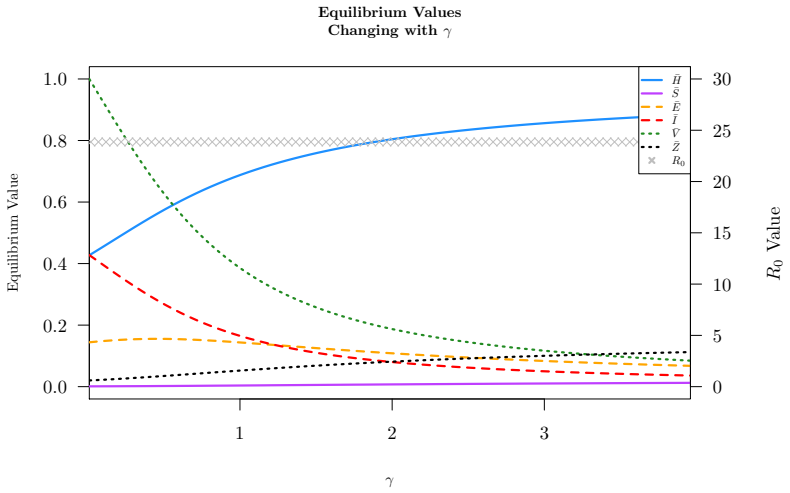

Figure 2.4: This plot shows the diseased equilibrium value as a function of γ. It can be seen that the values of $\mathcal{R}_0$ (gray x's) stay constant for all γ but the values for $\overline{E}, \overline{I}$, and $\overline{V}$ all decrease as γ increases.

We also explore how the rate of abrasion of the epithelium of the host organ χ affects the diseased equilibrium. Increasing this parameter increases the rate at which sites are uncovered and thus become susceptible to infection. This larger proportion of available susceptible sites increases the number of infections that occur $\overline{E}$ and $\overline{I}$. Interestingly, $\overline{I}$ and $\overline{V}$ increase and then decreases as χ increases. The viral production $\overline{V}$ and number of infectious sites $\overline{I}$ is maximized at $\chi = 0.0105$ under these particular parameter values. This increase in the number of exposed sites but

decrease in infectious sites and viral load is subtle. As the abrasion rate increases, more susceptible sites are produced through micro-abrasions in the epithelium. This increases the number of infections that occur. Because more sites are being infected, the immune response $\overline{Z}$ is increased. This increases the clearance of both exposed and infected sites. More exposed sites are cleared before they can transition to being infectious. This effect results in a net increase in the number of exposed sites but the reduction of infectious sites, and thus a reduction in viral load. It is already known that the number of sexual partners is associated with HPV infections. It would be worth examining if these same risky behaviours are also risk factors for persistent HPV infections. That is, the increase in sexual activity with different partners may increase the chance for an individual to contract an HPV infection, but they may also help sustain present infections as well.

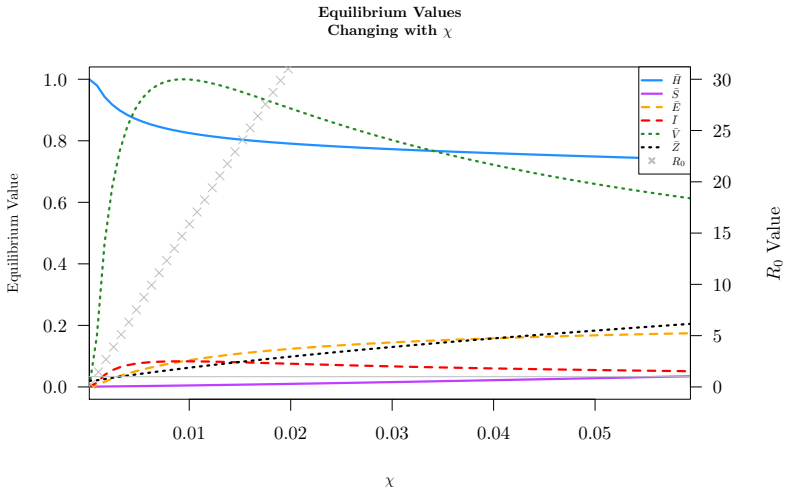

Figure 2.5: This plot illustrates the equilibrium value of the diseased equilibrium as a function of χ. We see that as χ increases, $\overline{E}$ increases due to more available susceptible sites to infect. However, $\overline{I}$ and $\overline{V}$ increase then decrease as χ increases, most likely due to an increase in immune activity clearing infectious I sites.

The other parameters affect the value of the diseased equilibria in predictable ways. As the rate of viral production f and rate of infection β increase, the equilibrium values $\overline{E}$, $\overline{I}$, and $\overline{V}$ also increase. Increasing the rate of clearance by the killer T-cells α has a large effect on reducing the viral load $\overline{V}$ and the present infection $\overline{E}$ and $\overline{I}$. If we increase the rate at which abraded sites recover ρ, then we see a decline in the infected site equilibria $\overline{E}, \overline{I}$, and $\overline{V}$. As δ increases we see a sharp decrease in $\overline{V}, \overline{E}$, and $\overline{I}$. This decrease is due to the elimination of virus particles preventing further infection. By increasing the rate at which exposed sites become infectious, σ, we increase the equilibrium value $\overline{I}$ and thus $\overline{V}$. By increasing the viral load, we see an increase in $\overline{E}$; however, for large values of σ exposed sites turn over to infectious so quickly that $\overline{E}$ begins to decline. Increasing the baseline level of immune cells increases the initial immune response, which has a negative effect on the infection equilibrium values, *i.e.,* $\overline{E}, \overline{I}$, and $\overline{V}$ decrease. Conversely, increasing the rate at which immune cells die μ, decreases the number of infected sites that can be cleared before the immune cells die. That is, $\overline{E}, \overline{I}$, and $\overline{V}$ increase as μ increases. The figures of the equilibrium values as functions of the parameters can be found in Appendix B.

2.4 Immune System with Memory Model

An important question in HPV research is why some individuals develop memory following an HPV infection and why many individuals do not. Here we explore the effects of memory cells within the context of our current HPV model. To do this, we include one more compartment into our model (M) referring to memory cells. We will also make some simplifying assumptions regarding memory cells to provide analytic simplicity. Firstly, memory cells are developed in the current presence of

killer T-cells Z at a rate ε and are limited by a carrying capacity for memory cells, K_M. Secondly, we assume that memory cells are long lasting and are not eliminated given the timescale of our model. If longitudinal studies or models are going to be considered examining the potential for HPV reinfection later in life, then a loss of memory cells should be explored. Thirdly, we assume that memory cells aid in the clearance of HPV similarly to how killer T-cells clear HPV.

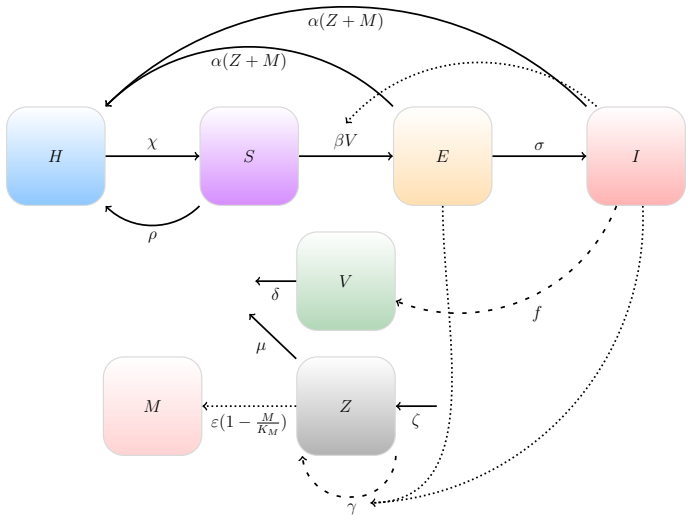

Figure 2.6: HPV Compartmental Model with Memory

524 Adding this compartment changes the system of differential equations slightly:

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\chi H + \rho S + \alpha(Z+M)\frac{(E+I)}{N}, \tag{11a}$$
$$\frac{\mathrm{d}S}{\mathrm{d}t} = \chi H - \rho S - \beta V\frac{S}{N}, \tag{11b}$$
$$\frac{\mathrm{d}E}{\mathrm{d}t} = \beta V\frac{S}{N} - \sigma E - \alpha(Z+M)\frac{E}{N}, \tag{11c}$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E - \alpha(Z+M)\frac{I}{N}, \tag{11d}$$
$$\frac{\mathrm{d}V}{\mathrm{d}t} = fI - \delta V, \tag{11e}$$
$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \zeta N + \gamma Z\frac{(E+I)}{N} - \mu Z, \tag{11f}$$
$$\frac{\mathrm{d}M}{\mathrm{d}t} = \varepsilon Z\left(1 - \frac{M}{K_M}\right), \tag{11g}$$
$$N = H + S + E + I. \tag{11h}$$

525 If we assume that there are no memory cells at the beginning of a naive infection,
526 $M_0 = 0$, then we can see that the $\mathcal{R}_0$ value for this system remains the same as before:

527

$$\mathcal{R}_0 = \frac{\beta f \chi \sigma \mu^2}{\delta(\alpha\zeta)(\chi+\rho)(\sigma\mu+\alpha\zeta)} \tag{12}$$

528 However, due to the proliferation of memory cells that are able to clear the infec-
529 tion, we must examine an expression for the effective reproductive number $\mathcal{R}_e$. As
530 the system equilibrates, the memory cells approach the value K_M, which can clear
531 the infection at rate α. Thus the expression for the effective reproduction number is:

532

$$\mathcal{R}_e = \frac{\beta f \chi \sigma \mu^2}{\delta(\alpha\zeta + \alpha K_M \mu)(\chi+\rho)(\sigma\mu+\alpha\zeta+\alpha K_M\mu)}. \tag{13}$$

533 In this sense, we may have that the infection is able to invade for $\mathcal{R}_0 > 1$, but if

$\mathcal{R}_e < 1$ then as time progresses, the production of memory cells will result in full clearance of the infection. If both $\mathcal{R}_0, \mathcal{R}_e > 1$, then the disease may persist. This is further illustrated in Figure 2.7.

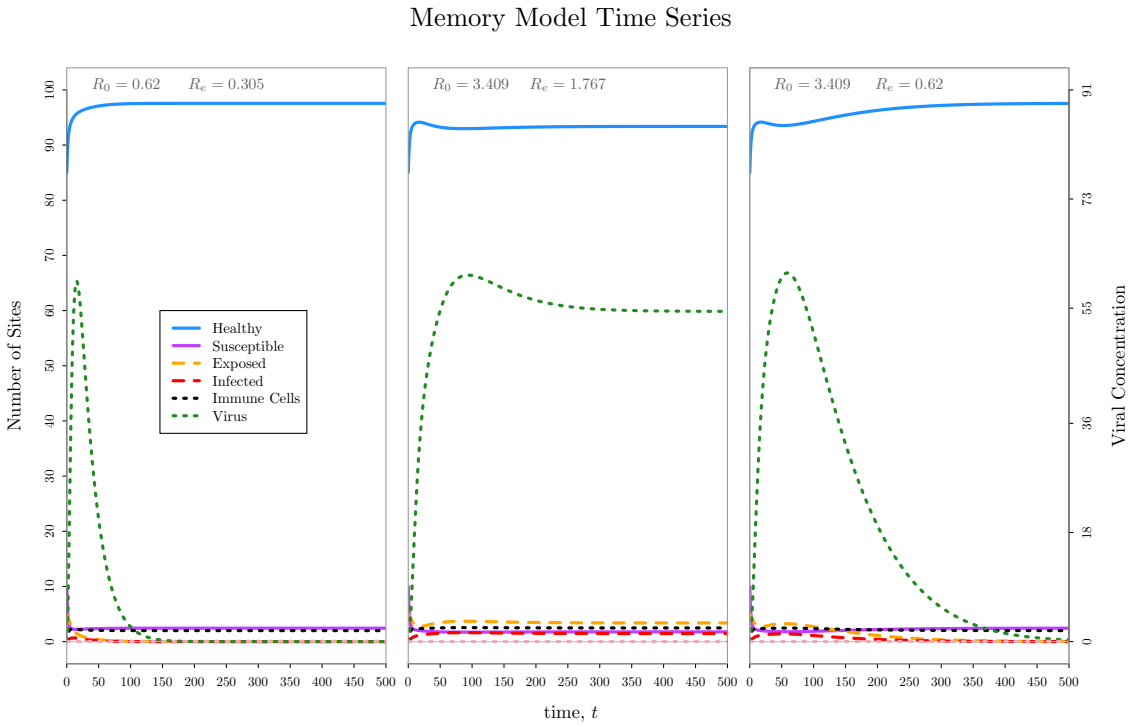

Figure 2.7: The following is a time series for the system with memory cells. In the first panel, we see that $\mathcal{R}_0 < 1$ so the infection cannot invade. In the second and third panels, we see that $\mathcal{R}_0 > 1$, so the infection can establish. However, in the third panel, we have changed the value of K_M such that $\mathcal{R}_e < 1$, and the infection is eventually cleared.

2.5 The Immune Response Delay Model

The human papillomavirus is a poor immunogen in that it does not elicit a strong immune response during infection. HPV is particularly good at evading the immune

system, such that many individuals do not even seroconvert after infection. That is, they don't produce long lasting antibodies.

Because of this, we introduce a time delay from the start of an infection until the immune response is able to "see" the pathogen and proliferate a specific immune response. We set up a similar model to the one before, but introduce a switch statement into the immune response differential equation:

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\chi H + \rho S + \alpha(Z+M)(E+I), \tag{14a}$$
$$\frac{\mathrm{d}S}{\mathrm{d}t} = \chi H - \rho S - \beta SV, \tag{14b}$$
$$\frac{\mathrm{d}E}{\mathrm{d}t} = \beta SV - \sigma E - \alpha(Z+M)E, \tag{14c}$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E - \alpha(Z+M)I, \tag{14d}$$
$$\frac{\mathrm{d}V}{\mathrm{d}t} = fI - \delta V, \tag{14e}$$
$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \zeta + s(t)\gamma Z(E+I) - \mu Z, \tag{14f}$$
$$\frac{\mathrm{d}M}{\mathrm{d}t} = s(t)\varepsilon Z\left(1 - \frac{M}{K_M}\right). \tag{14g}$$

In this case the switch statement keeps the immune response to the pathogen "turned off" until after some threshold time T_{start}, at which point the immune system detects HPV and is "turned on." We also do not see the production of memory cells primed against this pathogen until this system is "turned on." This function $s(t)$ is a step function with the following definition:

$$s(t) = \begin{cases} 0 & t < T_{start} \\ 1 & t \geq T_{start} \end{cases} \tag{15}$$

In this case, we have an innate immune response to the pathogen HPV when $t < T_{start}$ and we have an active immune response when $t \geq T_{start}$. We define T_{start} as the minimum time at which the cumulative infection increases above some immune detection threshold. That is, we set

$$T_{start} = \min(\tau), \text{ such that } C(\tau) > C_{thresh} \tag{16}$$

where $C(\tau) = E(\tau) + I(\tau)$ is the cumulative infection at time τ. This simulates a period where the host is naive to the antigen until after enough antigen is produced to trigger an adaptive immune response. To examine this system in more detail, we analyze the equilibria for each state.

2.5.1 Equilibria Analysis

Notice here, that in the case of the healthy equilibrium HE $(E = 0, I = 0)$, we have the exact same solution regardless of the switch value $s(t)$:

$$HE = \left(\frac{\rho}{\chi + \rho}, \frac{\chi}{\chi + \rho}, 0, 0, \frac{\zeta}{\mu}\right) \tag{17}$$

This leads to the same $\mathcal{R}_0$ values from the previous models regardless of the switch function value.

$$\mathcal{R}_0 = \frac{\beta f \chi \sigma \mu^2}{\delta \alpha \zeta (\chi + \rho)(\sigma \mu + \alpha \zeta)} \tag{18}$$

This is because at the beginning of an infection, regardless of the state of the switch parameter $s(t)$, we have the same infection free steady state.

While the healthy equilibrium has only one solution for both values of the switch

parameter ($s(t) = 0$ or $s(t) = 1$), there will be two different disease equilibria: one when the active immune system is turned off ($s(t) = 0$) and one when the active immune system is turned on ($s(t) = 1$). We will examine the solutions in both of these cases.

Active Immune System Off, $s(t) = 0$

By considering the case where $t < T_{start}$, this system of differential equations simplify to the following:

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\chi H + \rho S + \alpha Z(E + I), \tag{19a}$$
$$\frac{\mathrm{d}S}{\mathrm{d}t} = \chi H - \rho S - \beta SV, \tag{19b}$$
$$\frac{\mathrm{d}E}{\mathrm{d}t} = \beta SV - \sigma E - \alpha ZE, \tag{19c}$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E - \alpha ZI, \tag{19d}$$
$$\frac{\mathrm{d}V}{\mathrm{d}t} = fI - \delta V, \tag{19e}$$
$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \zeta - \mu Z, \tag{19f}$$
$$\frac{\mathrm{d}M}{\mathrm{d}t} = 0. \tag{19g}$$

Notice, starting the initial memory response $M_0 = 0$, means that $M(t) = 0$ for all time t. Here we can see that $\overline{Z} = \frac{\zeta}{\mu}$, and using this we can easily solve this system of

equations and obtain the following solution:

$$\overline{H} = \frac{\delta\alpha\zeta((\alpha\zeta+\sigma\mu)(\mu\rho-\alpha\zeta)+\frac{\beta f}{\delta}\sigma\mu^2)}{\beta f\sigma\mu^2(\alpha\zeta+\chi\mu)}, \tag{20a}$$

$$\overline{S} = \frac{\delta\alpha\zeta(\sigma\mu+\alpha\zeta)}{\beta f\sigma\mu^2}, \tag{20b}$$

$$\overline{E} = \frac{\delta\alpha\zeta(\frac{\beta f}{\delta}\sigma\mu^2\chi-\alpha\zeta(\sigma\mu+\alpha\zeta)(\chi+\rho))}{\beta f\sigma\mu(\alpha\zeta+\chi\mu)(\alpha\zeta+\sigma\mu)}, \tag{20c}$$

$$\overline{I} = \frac{\delta(\frac{\beta f}{\delta}\sigma\mu^2\chi-\alpha\zeta(\sigma\mu+\alpha\zeta)(\chi+\rho))}{\beta f(\alpha\zeta+\chi\mu)(\alpha\zeta+\sigma\mu)}, \tag{20d}$$

$$\overline{V} = \frac{f\overline{I}}{\delta} \tag{20e}$$

$$\overline{Z} = \frac{\zeta}{\mu} \tag{20f}$$

Notice here that, if $\mathcal{R}_0 < 1$, *i.e.,*

$$\mathcal{R}_0 < 1 \tag{21}$$

$$\frac{\beta f\chi\sigma\mu^2}{\delta\alpha\zeta(\chi+\rho)(\sigma\mu+\alpha\zeta)} < 1 \tag{22}$$

$$\frac{\beta f}{\delta}\chi\sigma\mu^2 < \alpha\zeta(\chi+\rho)(\sigma\mu+\alpha\zeta) \tag{23}$$

And from here we can see that the diseased equilibrium values $\overline{E}$ and $\overline{I}$ in (20) are negative when $\mathcal{R}_0 < 1$, outside of the biologically relevant domain. Thus, this equilibrium will not be obtained when $\mathcal{R}_0 < 1$ given a biologically relevant initial condition.

Active Immune System On, $s(t) = 1$

In this case, we have that the active immune response has identified the antigen, and the system recruiting new immune cells directly proportional to the current concentration of immune cells and the number of infected (E and I) sites. This system of differential equations is then the following:

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\chi H + \rho S + \alpha(Z+M)(E+I), \tag{24a}$$
$$\frac{\mathrm{d}S}{\mathrm{d}t} = \chi H - \rho S - \beta SV, \tag{24b}$$
$$\frac{\mathrm{d}E}{\mathrm{d}t} = \beta SV - \sigma E - \alpha(Z+M)E, \tag{24c}$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E - \alpha(Z+M)I, \tag{24d}$$
$$\frac{\mathrm{d}V}{\mathrm{d}t} = fI - \delta V, \tag{24e}$$
$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \zeta + \gamma Z(E+I) - \mu Z, \tag{24f}$$
$$\frac{\mathrm{d}M}{\mathrm{d}t} = \varepsilon Z\left(1 - \frac{M}{K_M}\right). \tag{24g}$$

In this situation we obtain the same Disease Equilibrium as the previous models, which have previously been analyzed. We will conclude some details in regards to how the system switches from no active immune response to the active immune response. We have already discussed that the basic reproduction number $\mathcal{R}_0$ is the same for both $s(t) = 0$ and $s(t) = 1$, which means that in this situation, HPV would not be able to invade either system when $\mathcal{R}_0 < 1$.

We simulate this deterministic system by using the ordinary differential equation solver package `deSolve` in **R**. We set up the system of differential equations as discussed in system (14). We implemented the switch statement initially by setting the

T_{start} to be the end time of the simulation, to allow for the naive system to begin. We updated the T_{start} variable by keeping track of the cumulative infection $C(t)$ and updated $T_{start} = \tau$ to be the first time $C(\tau) > C_{thresh}$ using `event` and `root` options implemented in the `ode()` function in the `deSolve` package. The results of the model along with parameter values are shown in Figure 2.8 in the middle panel.

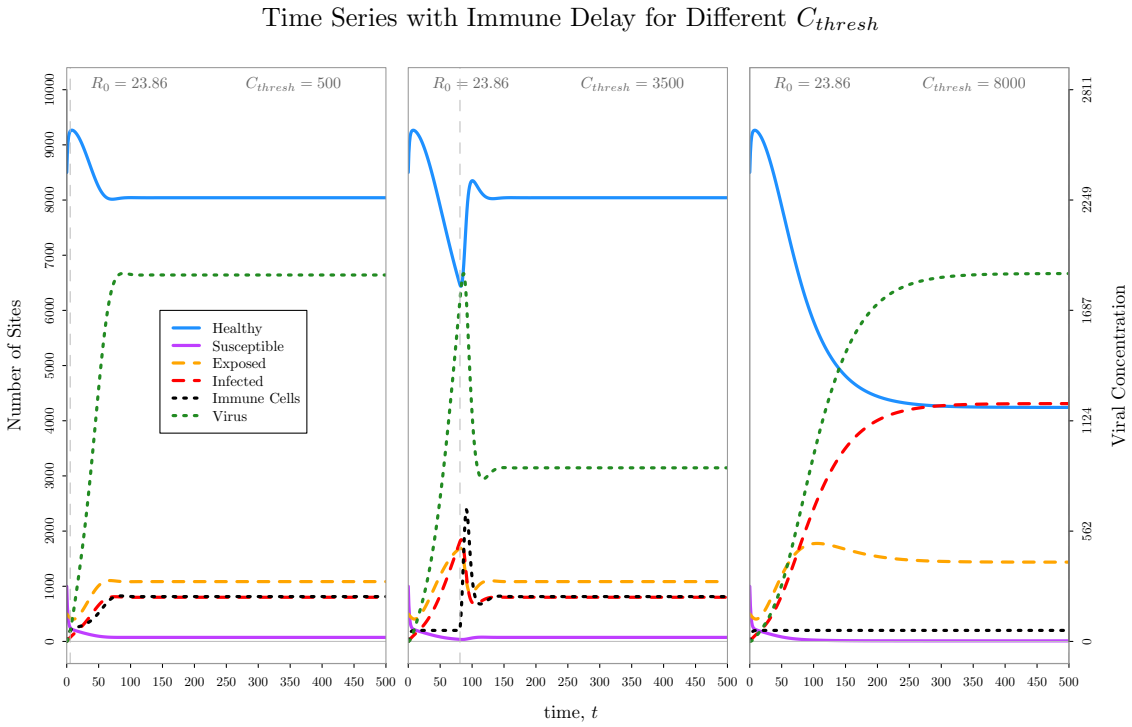

Figure 2.8: The left panel illustrates the base case when the immune response is triggered immediately. The middle panel showcases the delayed immune response with $C_{thresh} = 3500$, after which it settles onto the base equilibrium. The right panel illustrates the case when the $C_{thresh} = 8000$ is never met, and thus the system settles on the inactive immune response equilibrium.

Given this updating method, it may be possible that the the cumulative infection $C(t) < C_{thresh}$ for all time t. In this situation, the host never develops an immune response specific to the antigen and the HPV infection is able to persist if established.

This is illustrated in the time series in Figure 2.8 in the right panel.

2.6 Discussion

Within host modelling is particularly helpful for understanding and disentangling many of the complex host-pathogen interactions that take place during an infection with HPV. In particular, these models can help us understand some of the current questions surrounding HPV infection, such as why certain individuals develop persistent infections, how clearance is possible without eliciting a strong immune response or even any antibody response, and what factors may result in latent infections with HPV. Importantly, these within host models provide a basis by which to further examine these questions. Further research may be done through observational and physiological studies and by using empirical evidence alongside models better tailored to answer specific questions.

In this study, we developed various within host models that account for number of different immune response scenarios. Importantly, we formulate an expression for the within-host basic reproduction number $\mathcal{R}_0$. Using this value, we are able to better understand why an infection may or may not be able to establish within a host. We also explicitly solve for the equilibria of our models, and use these values to inform the likelihood of clearance. We can see that an infection can be reduced small enough such that the pathogen may be cleared due to some stochastic event. This can explain the variable clearance times of many individuals and may further explain persistent infections in about 10% of the population. What we show is the clearance is possible, and often likely even when the deterministic, continuous system converges to the Disease Equilibrium.

Furthermore, because we were able to analytically solve for the basic reproduction number and equilibria from our models, we are able to examine what parameters may potentially affect disease dynamics in a predictable way. In particular, we showed that the generation rate of T-cells in the presence of the pathogen (γ) has no effect on the ability of the pathogen to establish in our model. That is, the γ term does not appear in the expression for $\mathcal{R}_0$. This makes intuitive sense in that the immune response is triggered by the presence of the pathogen, and at the beginning of the infection there isn't any established. However, this rate γ does have important implications for the value of the diseased equilibrium. As γ increases, both $\overline{E}$ and $\overline{I}$ decrease. It may be possible for the system to decrease these values below one individual infection site, in which case full clearance is established. However, in most cases, the immune response reduces the value of the diseased equilibrium such that clearance may occur through stochastic events. Therefore, it may be a likely case that persistent infections are due to weak CTL propagation in the host.

We also consider how the parameter χ affects the value of $\mathcal{R}_0$. For values of $\chi \ll \rho$, then $\mathcal{R}_0$ increases linearly as χ increases. Because abrasions in the epithelium of the cervix are caused during sexual intercourse and ρ is related to the natural cell replication cycle, it is sensible to assume that $\chi \ll \rho$. In this way, an increase in χ can increase the chances of an infection establishing in the host and may decrease the probability of clearance.

We also consider the case for the establishment of immune memory. Immune memory is important as it may be able to help clear current infections by reducing the equilibrium values of the exposed and infectious sites such that they are cleared more easily. As well, as the infection progresses and memory immune cells are established,

the effective reproduction $\mathcal{R}_e$ number decreases. If this value decreases below one, then the infection will be cleared and a future infection will not establish, granted the memory cells remain for that duration of time. We do not consider a loss of immune cells in these models. Important questions regarding reinfection with certain HPV types later in life and the potential for latent infections are currently being examined, and mathematical models developed to help elucidate these problems should include factors that consider the loss of immune memory.

Lastly, we consider the case where we introduce a delay in the activation of an immune response. In particular, we introduce a switch function that is triggered when the cumulative infection $(E + I)$ reaches a particular threshold C_{thresh}. When the system is off $(E + I < C_{three})$, we have no active immune response and the pathogen can spread in the host with no clearance apart from that by the already present innate immune response $(\frac{\zeta}{\mu})$. If the cumulative infection reaches the threshold, then the immune system is activated, and the system settles on the normal diseased equilibrium.

This chapter provides solid groundwork for modelling the within-host spread of HPV. We consider the biological mechanisms used to inform the construction of our model. We examine the equilibria of the model, and discuss how clearance is possible even when the infection is able to establish. Ultimately, this work aims to provide the ground-work by which to explore within-host HPV dynamics in future studies.

Chapter 3

A Spatial Simulation Model

3.1 Introduction

The human papillomavirus (HPV) is a pathogen that locally infects the bottom-most basal cells of the epithelium. These sites are only able to become infected after the top-most layer of the epithelium has been abraded, exposing the bottom-most layer basal cells. Virus particles then infect these susceptible sites and begin the infection cycle. Virus particles are developed within the host sites as the cells replicate and move up the epithelium. As the cells move closer to the top of the epithelium, the cells flatten and form the upper squamous layer. Virus particles are released into the surrounding environment as the cells undergo natural cell death and desquamate.

In this section, we consider a localized infection structure. We assume that virus particles are only able to infect susceptible sites in a restricted neighbourhood. This differs from other within host models [34, 35] of HPV, which consider homogenous mixing. We run simulations with the local neighbourhood and compare it to an analogous simulation with a global neighbourhood to examine how the local structure

may affect the dynamics of the infection.

We establish a base set of parameters that are derived from literature values and estimates. We run these base parameters for both the global neighbourhood and local neighbourhood to compare the differences in spatiality. We also compare the global neighbourhood results with the deterministic model results to confirm event rates. We show that the locality plays an important role on the establishment or clearance of an infection for the same parameter values. This suggests that when developing within host models to examine viral kinetic parameters, spatiality may be an important factor when fitting parameters to data, particularly when comparing it to population level data.

3.2 Methods

We develop a stochastic, spatial model and implement it as an adaptation of the Gillespie Algorithm [23]. We initialize the system as an organ of N sites arranged as a grid on a torus. Each site i has a neighbourhood of sites which it is connected to. The size of the neighbourhood at site i is $\mathcal{N}(i)$. In this chapter, we focus on two different neighbourhood structures: global and local. The global neighbourhood structure of a site includes all other sites with equal weighting. In this way, our spatial simulation mimics a homogenous system. The local neighbourhood only contains the four closest sites to the focal site—those North, South, East, and West of the site. The local structure is illustrated in Figure 3.1. This spatial model is an extension of the homogenous, deterministic model discussed in Chapter 2. The sites may have one of four different states: healthy, susceptible, exposed, and infectious. A healthy site i becomes susceptible after an abrasion event. This susceptible site may become healthy

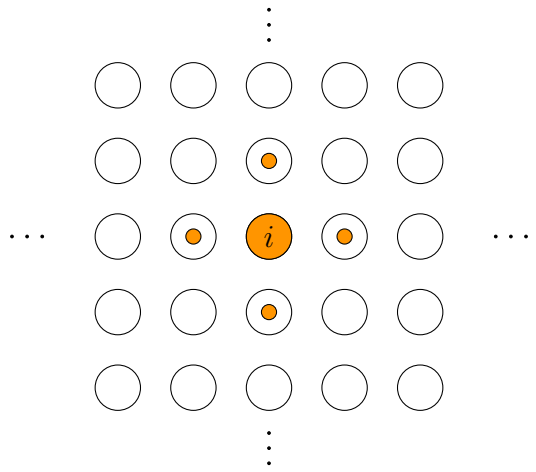

Figure 3.1: Local neighbourhood illustration. The focal site i is connected to the sites above, below, to the left, and to the right of the focal site. All these sites are connected as such, and arranged as a lattice on a torus.

again after recovery, or it may become infected by a viral particle. After infection, it becomes exposed, where it then becomes infectious following a maturation event. These infectious sites are then able to produce virus particles, which can infect other sites. Both infectious sites and exposed sites can be cleared by the immune system Z, after which they become healthy again. To simplify our spatial model, we make the assumption of fast viral dynamics. We solve for V using a pseudo-equilibrium (Equation (1)), and use this to calculate rates of infection events.

$$\overline{V} = \frac{fI}{\delta} \tag{1}$$

To ensure that this substitution maintains the dynamics of the model, we compare the numerical solutions of the base model with the fast dynamics model in Figure 3.2. We see that the dynamics are very similar when we make the pseudo-equilibrium substitution into the deterministic model, and as such we apply this method when

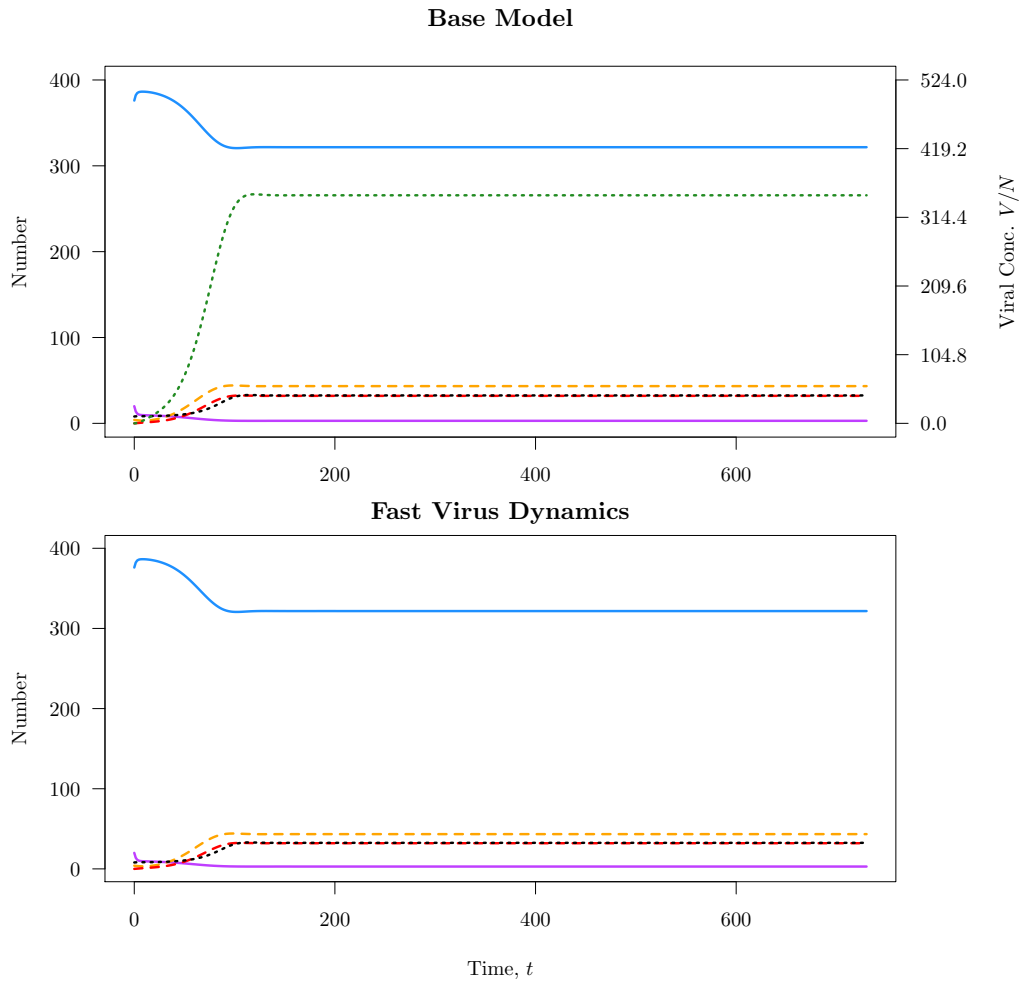

Figure 3.2: These plots compare the dynamics of the basic deterministic model with that of the fast-virus model. It can be seen that the dynamics of the sites are very similar, and maintain the overall dynamics.

developing the spatial models.

The Gillespie Algorithm developed for this particular model is discussed below. We begin by calculating the rates of the events. The rate of event j occurring on site i is represented as $c_j(i)$. We then calculate the rate of event j occurring by summing up over all the individual rates $c_j(i)$, $c_j = \sum_i c_j(i)$. The total rate of events is the

summation of the total rates over all events j, $c = \sum_j c_j$.

The time to next event Δt is sampled from an exponential distribution with parameter c, $\Delta t \sim \text{Exp}(c)$. Event j is selected with probability $\frac{c_j}{c}$, and site i is selected with probability $\frac{c_j(i)}{c_j}$. The site i is updated accordingly (see Table 3.1), time is updated $t = t + \Delta t$, and all the rates are also updated accordingly. The rates are derived in section 3.2.1 and are also listed in Table 3.1. The simulation terminates when $t > t_{\max}$. This algorithm is summarized in Algorithm 1.

Algorithm 1 Gillespie Algorithm

1: Initialize the population and set time, $t = 0$.
2: Calculate rate constants c_j for each event j by summing up the individual rate constants $c_j(i)$ over all sites i.
3: Sample the length of the time step from an exponential distribution $\Delta t \sim \text{Exp}(c)$, with parameter $c = \sum_j c_j$.
4: Select the event which occurs with probability c_j/c.
5: Execute the chosen event on site i chosen randomly with probability $c_j(i)/c_j$ and update the time, $t = t + \Delta t$, and the states of the sites. Go to step 2 and repeat until either $t > t_{\max}$ or the predetermined maximum number of iterations is completed.

3.2.1 Calculation of the Event Rates

We discuss the calculation of all the event rates and the subsequent probabilities of each event. An abrasion event occurs at a healthy site i at a rate of χ. Summing this up over all the healthy sites, we obtain the total rate of an abrasion c_{a} over all the sites in the organ:

$$c_{\text{a}} = \sum_{i=1}^{H} \chi = \chi H. \tag{2}$$

Similarly, the rate of a recovery event on a susceptible site i is ρ, and thus the rate of a recovery event over all the sites is

$$c_{\mathrm{r}} = \sum_{i=1}^{S} \rho = \rho S. \tag{3}$$

We consider an infection event occurring when a infectious site i is able to infect a susceptible site j in its neighbourhood, $\mathcal{N}_{\mathrm{S}}(i)$. We normalize by dividing by the total number of neighbours of site i, $\mathcal{N}(i)$.

$$c_{\mathrm{inf}}(i) = \beta V_i = \frac{\beta f}{\delta} \frac{\mathcal{N}_{\mathrm{S}}(i)}{\mathcal{N}(i)} \tag{4}$$

Summing this up over all infectious sites, we obtain the total rate of an infection event occurring

$$c_{\mathrm{inf}} = \sum_{i=1}^{S} c_{\mathrm{inf}}(i) = \sum_{i=1}^{I} \frac{\beta f}{\delta} \frac{\mathcal{N}_{\mathrm{S}}(i)}{\mathcal{N}(i)}. \tag{5}$$

Instead of keeping track of all possible Infectious-Susceptible pairs, which can be computationally expensive, instead we sum of the rates of infection over all infectious sites:

$$c_{\mathrm{inf}} = \sum_{i=1}^{I} \frac{\beta f}{\delta} = \frac{\beta f I}{\delta}, \tag{6}$$

and then randomly select an infectious site i and a neighbouring site j. If the neighbouring site j is susceptible, then the infection event is carried out and the states updated. Otherwise, it does not occur, and the states are kept the same.

The rate of a maturation event occurring at an exposed site i is σ, and thus the

752 total rate of a maturation event occurring over all sites is

$$c_{\text{mat}} = \sum_{i=1}^{E} \sigma = \sigma E. \tag{7}$$

753 Lastly, a clearance event may occur at either an exposed or infectious site i with rate
754 αZ_i, where Z_i is the number of immune cells at site i. We can calculate the total rate
755 of clearance of a site with state X by summing up over all X sites

$$c_{\text{cl,X}} = \sum_{i=1}^{X} \alpha Z_i, \tag{8}$$

756 where X is the number of sites of state X, which is either exposed or infectious sites
757 in this case. The events and their corresponding rates are summarized in Table 3.1.

Table 3.1: Event Rates

Event	Action	Individual Rate	Total Rate
Abrasion (a)	$H \to S$	$c_{\text{a}}(i) = \chi$	$c_{\text{a}} = \chi H$
Recovery (r)	$S \to H$	$c_{\text{r}}(i) = \rho$	$c_{\text{r}} = \rho S$
Infection (inf)	$S \to E$	$c_{\text{inf}}(i) = \frac{\beta f}{\delta}\frac{\mathcal{N}_{\text{I}}(i)}{\mathcal{N}(i)}$	$c_{\text{inf}} = \frac{\beta f I}{\delta}$
Maturation (mat)	$E \to I$	$c_{\text{mat}}(i) = \sigma$	$c_{\text{mat}} = \sigma E$
Clearance (cl,E)	$E \to H$	$c_{\text{cl,E}}(i) = \alpha Z_i$	$c_{\text{cl,E}} = \sum_{i=1}^{E} \alpha Z_i$
Clearance (cl,I)	$I \to H$	$c_{\text{cl,I}}(i) = \alpha Z_i$	$c_{\text{cl,I}} = \sum_{i=1}^{I} \alpha Z_i$

758 We keep track of the immune cells Z deterministically. We update the number of
759 immune cells at site i using Euler's Method

$$Z_i(t + \Delta t) = Z_i(t) + \frac{\mathrm{d}Z_i(t)}{\mathrm{d}t}\Delta t, \tag{9}$$

where $\frac{\mathrm{d}Z_i}{\mathrm{d}t}$ is calculated based on the neighbourhood of the site i

$$\frac{\mathrm{d}Z_i(t)}{\mathrm{d}t} = \zeta + \gamma Z_i(t)\frac{\mathcal{N}_{\mathrm{E,I}}(i)}{\mathcal{N}(i)} - \mu Z_i, \tag{10}$$

and where $\mathcal{N}_{\mathrm{E,I}}(i) = \mathcal{N}_{\mathrm{E}}(i) + \mathcal{N}_{\mathrm{I}}(i)$ is the number of neighbours surrounding i with state E or I. Furthermore, the step size Δt is the same time step sampled from the exponential distribution: $\Delta t \sim \mathrm{Exp}(c)$. From the initial system, the expected value of $\Delta t \approx 0.027$ days. This value is small enough such that the Euler's Method solution will not diverge from the true solution. This is further confirmed pictorially when comparing the stochastic results and the deterministic results (compare Figure 3.3 and Figure 3.4).

When considering the global neighbourhood *i.e.,* $\mathcal{N}(i) = N$ for all sites i, we are able to make some simplifying assumptions. Firstly, that the number of infectious sites in a site i's neighbourhood is just the total number of infectious sites, $\mathcal{N}_{\mathrm{I}}(i) = I$. This reduces the rate of an infection event to be

$$c_{\mathrm{inf}} = \sum_{i=1}^{S} \frac{\beta f}{\delta}\frac{I}{N} = \frac{\beta f S I}{\delta N}. \tag{11}$$

In a global neighbourhood, the immune cells are distributed evenly amongst sites, $Z_i = \frac{Z}{N}$. Thus the total rate of a clearance event occurring on exposed or infectious sites together is

$$c_{\mathrm{cl,E\,and\,I}} = \sum_{i=1}^{E+I} \alpha Z_i = \alpha\frac{Z(E+I)}{N}. \tag{12}$$

Thus we can just consider the total amount of immune cells in the global neighbourhood scenario. We update the immune cells in the same way as (10) but over all the

sites

$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \sum_{i=1}^{N} \frac{\mathrm{d}Z_i}{\mathrm{d}t}, \tag{13a}$$

$$= \sum_{i}^{N} \zeta + \gamma Z_i \frac{\mathcal{N}_{\mathrm{E,I}}(i)}{\mathcal{N}(i)} - \mu Z_i, \tag{13b}$$

$$= \zeta N + \sum_{i=1}^{N} \gamma Z_i \frac{\mathcal{N}_{\mathrm{E,I}}(i)}{\mathcal{N}(i)} - \mu Z, \tag{13c}$$

$$= \zeta N + \gamma \frac{Z(E+I)}{N} - \mu Z. \tag{13d}$$

Considering these global rates, we expect that the global neighbourhood simulations are qualitatively similar to the homogenous, deterministic model discussed in Chapter 2.

After running the simulations for different parameter values, we calculate the proportion of infections that lead to clearance within two years, the mean time to clearance, and the 95% quantile ranges (QR) of time to clearance.

3.3 Results

We run the system for 100 realizations, stopping when $t > 730$ days, or two years. We set the base parameter values to the those found in Table 3.2.

Table 3.2: Parameter values used in spatial simulations

Parameter	Description	Value (day^{-1})	Source
α	Rate of clearance of infected sites by the immune system	0.5	[35]
β	Infection rate of a virus particle on a susceptible site	0.003	[8]
f	Rate of virus produced by infectious site	$600\frac{\text{copies}}{\text{cell}\cdot\text{day}}$	[51]
δ	Rate of natural viral decay	0.138	[39]
χ	Abrasion rate of the epithelium	0.015	Est.
ρ	Recovery rate of the epithelium	0.6	Est.
σ	Maturation rate of newly exposed sites	0.03	Est.
γ	Rate of T-cell propagation in presence of pathogen	2.0	[4]
ζ	Base T-cell repose rate	0.01	Est. [42]
μ	Rate of T-cell death	0.5	[4]
N	Number of sites	400 sites	Const.

We also consider various other parameter values. In the global neighbourhood, we consider clearance parameters by setting the parameters to those found in Table 3.2 but changing the rate of clearance due to primed T-cells to $\alpha = 2.0\ \text{day}^{-1}$and changing the rate of primed T-cell propagation to $\gamma = 14\ \text{day}^{-1}$.

We also run these base parameters (Table 3.2) for the local simulation, but also include a case where the local infection is able to establish. The establishment parameters are the same as those set in the base simulation but changing the rate of primed T-cell propagation to $\gamma = 0.2\ \text{day}^{-1}$. These results are explained and visualized below.

3.3.1 Global Neighbourhood

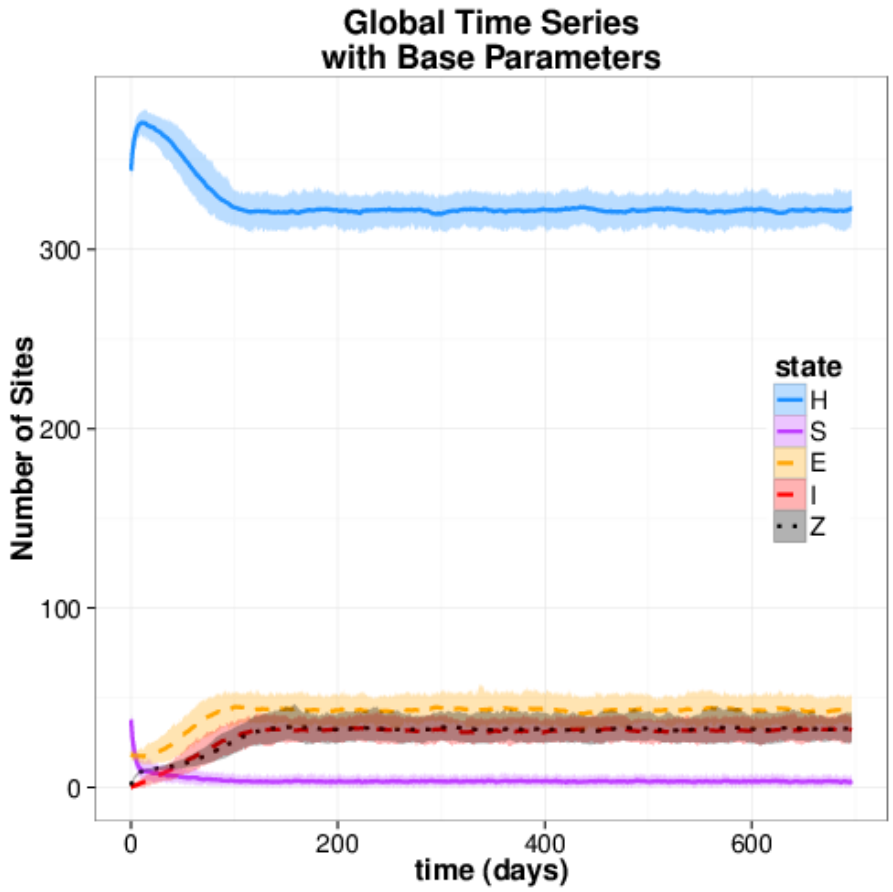

Figure 3.3: Stochastic simulation results for the global neighbourhood model with base parameters. The bands include the 90% inter-quantile range for the different realizations; the line is the mean of all realizations.

Comparing the stochastic results in Figure 3.3 to the deterministic model results in Figure 3.4, we see that the global simulation matches well with the deterministic solution. This suggests that the rates have been properly initialized for the spatial model.

Under these parameter values, we see no clearance of the pathogen. To examine

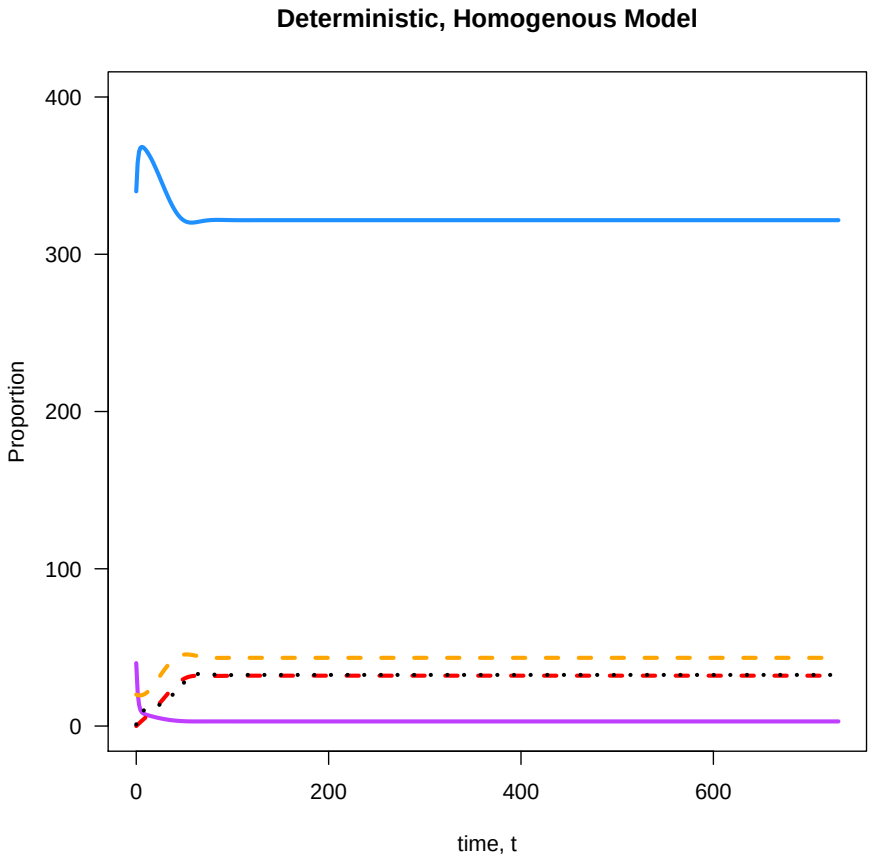

Figure 3.4: Deterministic simulation results with base parameters.

the potential for stochastic clearance of the pathogen, we vary the host immune parameters γ (the rate of T-cell propagation) and α (the rate of clearance by the primed T-cells). We set $\alpha = 2$ day^{-1}and $\gamma = 14$ day^{-1}. The time series can be found in Figure 3.5. Under these parameter values, clearance occurred 78% of the time, and the mean time to clearance was 260.8 and 95% QR (73.7, 541.7) days, while the median was 250 days. The distribution of clearance times can be found in Figure 3.7. We define clearance as the lack of infectious and exposed sites in the system, and we define time to clearance as the first instance after initial infection when both exposed

and infectious site are zero.

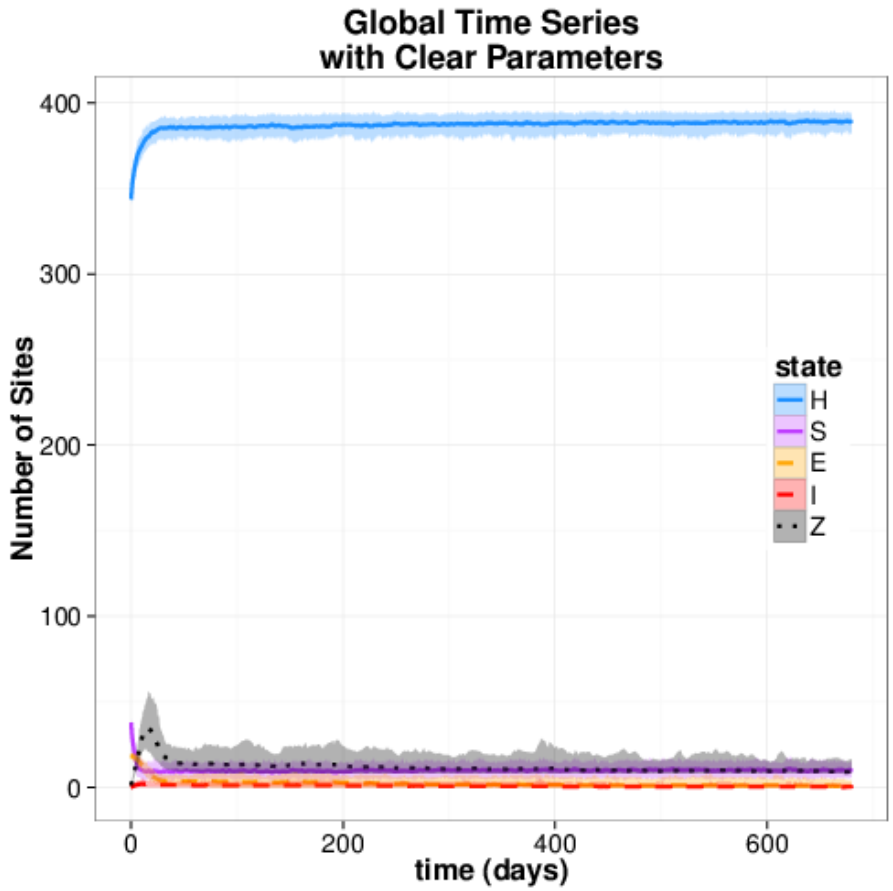

Figure 3.5: Stochastic simulation results for the global neighbourhood model with clearance parameters. We set $\alpha = 2$ day^{-1}and $\gamma = 14$ day^{-1}and keep the remaining parameter values the same. The bands include the 90% inter-quantile range for the different realizations; the line is the mean of all realizations. This set of parameters illustrates the stochastic clearance events.

We also compare these clearance results to the deterministic model time series (Figure 3.6). While it may seem that the system dies out, it actually has a basic reproduction value of $\mathcal{R}_0 = 3.41$. It's that the diseased equilibrium values are so low that they are cleared by stochastic effects in the spatial model.

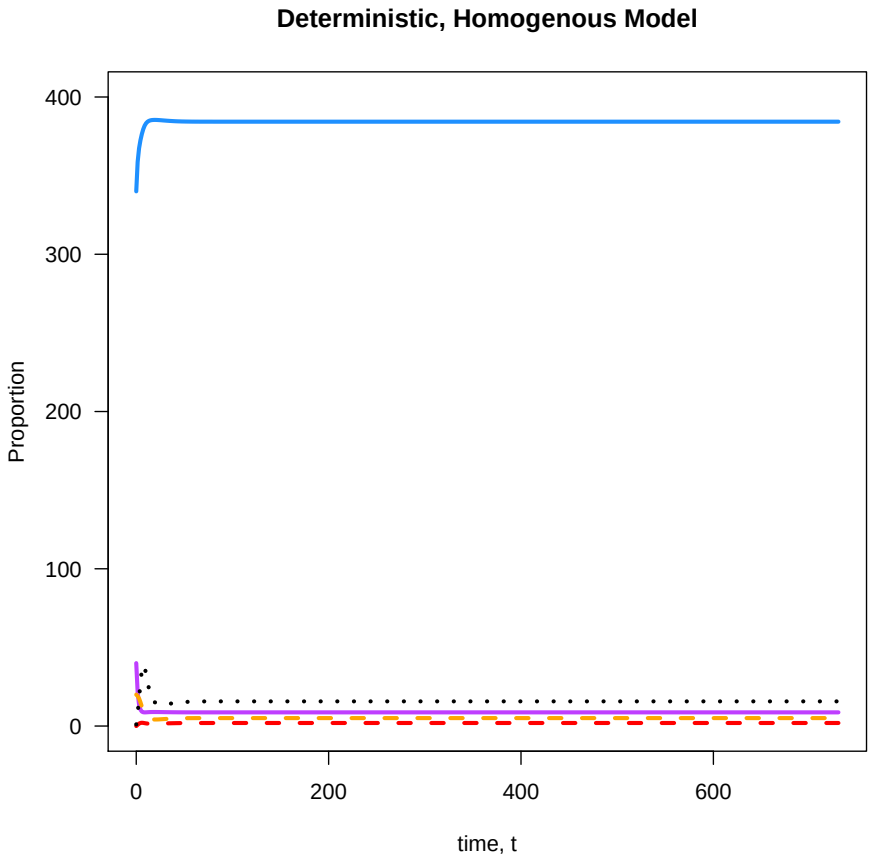

Figure 3.6: Deterministic simulation results with clearance parameters: $\alpha = 2.0$ day^{-1}and $\gamma = 14$ day^{-1}.

Cross Sections in Time of the Global Neighbourhood Simulation

We can see one realization for this at different snapshots in time (Figure 3.8). This illustrates how the infection occurs on the host organ. In the global neighbourhood, the infection is able to spread from one part of the organ to any other part of the organ as long as there is a a susceptible site for infection. The realization simulated in Figure 3.8 uses the parameter values in Table 3.2, and we see that the infection is sustained. We also observe similar snapshots in time for the scenario with clearance

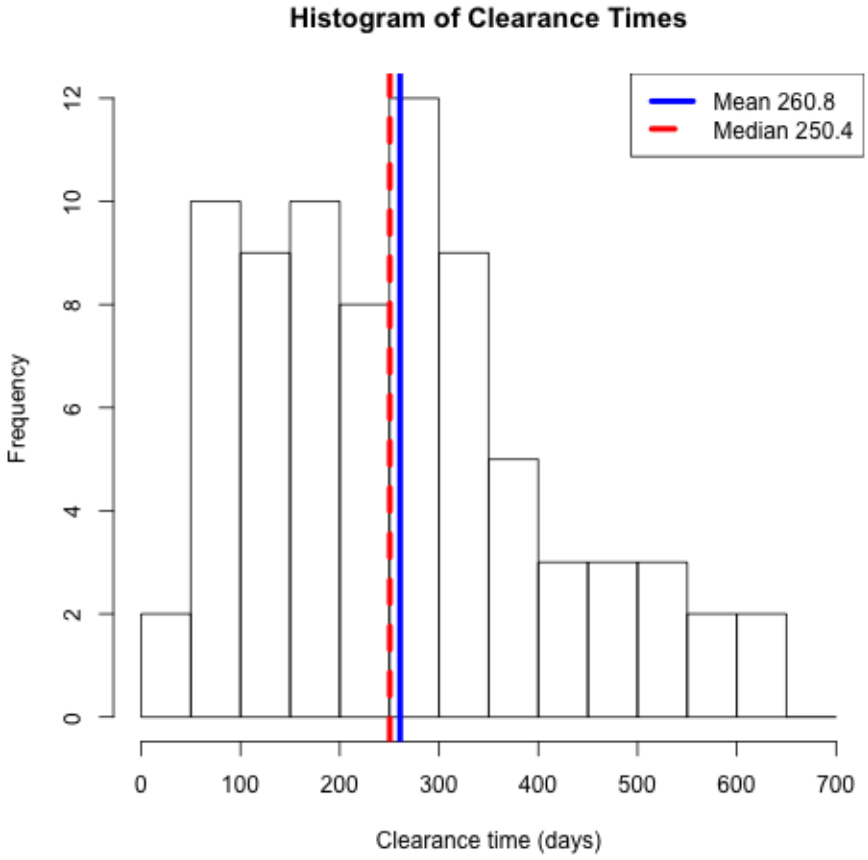

Figure 3.7: Histogram of the distribution of clearance times when setting $\alpha = 2$ day^{-1}and $\gamma = 14$ day^{-1}for the global neighbourhood structure. The mean is 260.8 days and the median 250 days.

parameters. Here it is seen that the infection dies out particularly quickly in this realization. The speed at which the infection is cleared using the clearance parameters can be illustrated through snapshots of the site states at different points in time in Figure 3.9.

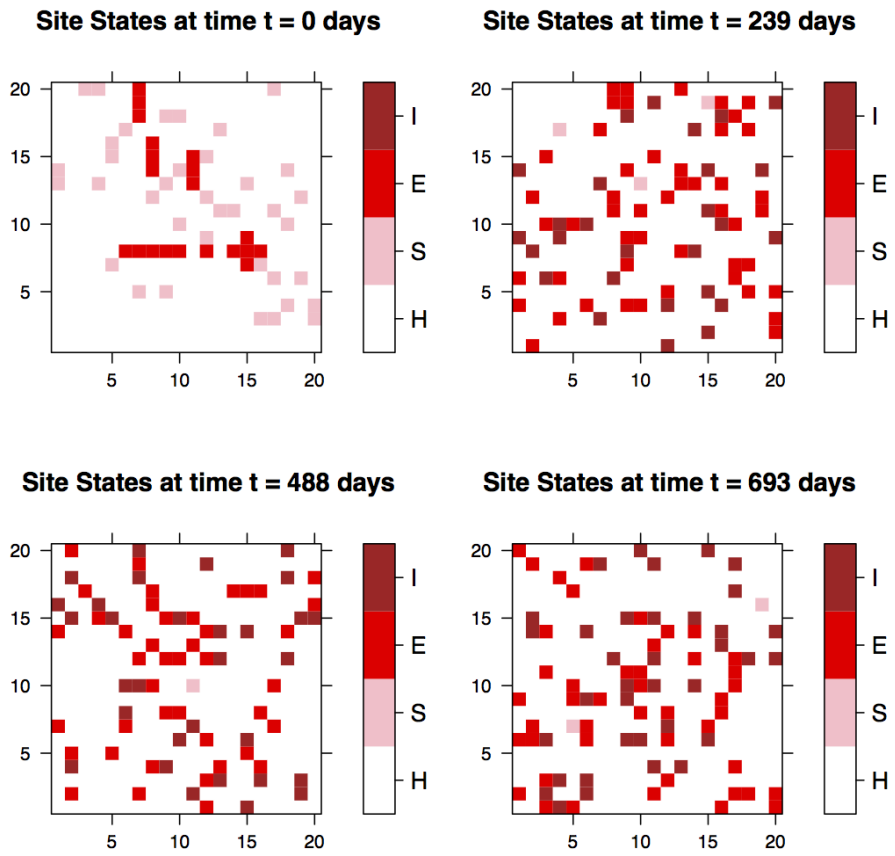

Figure 3.8: This figure provides a snapshot of the organ at different times for one realization. This provides a visualization of what happens during the infection. Because this is a global neighbourhood structure, it can be seen that the infection spreads across the organ and not in a connected fashion.

3.3.2 Local Neighbourhood

We run the simulation for the local neighbourhood under a number of different parameter values. For the same parameter values of the global simulation (see Table 3.2), the local simulation results in clearance in every realization. This can be seen in the time series in Figure 3.10.

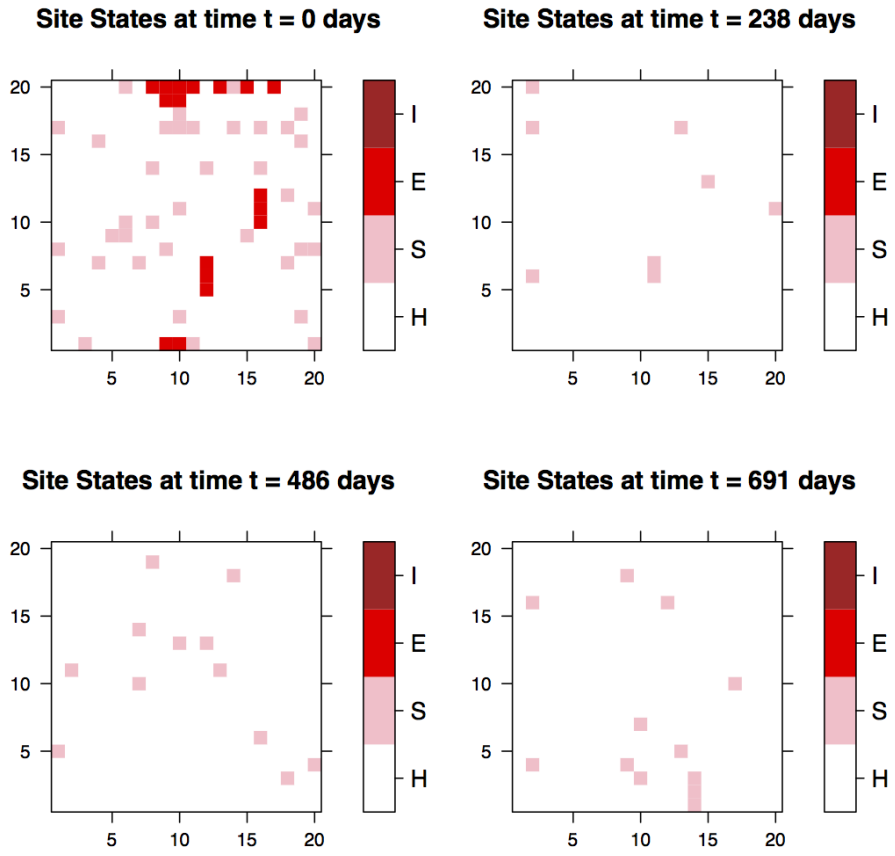

Figure 3.9: This figure provides a snapshot of the organ at different times for one realization. This provides a visualization of what happens during the infection. While this is a global infection scenario, under the clearance parameters, the system clears very quickly.

When we run the local simulation with the base parameters for 100 realizations, we see that each realization results in clearance. The mean time for clearance was found to be 5.0 with a 95% QR (4.53, 6.10) days, and the median time to clearance was 4.89 days. The distribution for clearance times can be found in Figure 3.11. This shows that with the same parameters, the local neighbourhood restricts the infection process so much that it dies out before it can even establish.

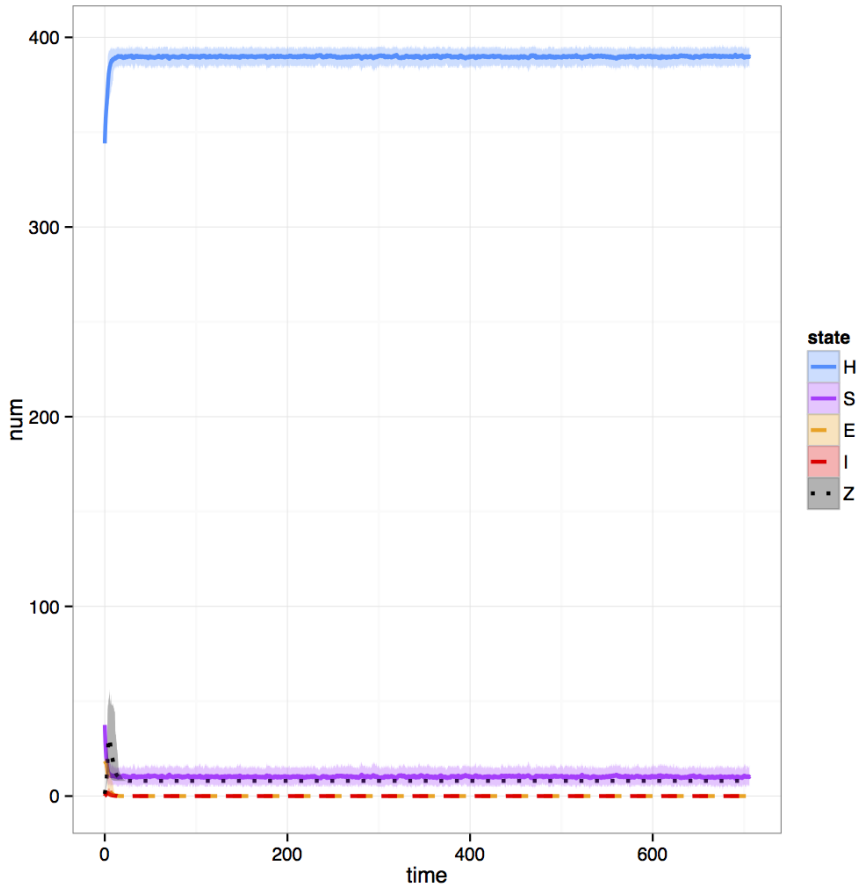

Figure 3.10: Stochastic simulation results for the local neighbourhood model with base parameters. The bands include the 90% inter-quantile range for the different realizations, the line is the mean of all realizations. We ran this for 100 realizations. Under local infection, we see that clearance is much more common than in the global neighbourhood.

We can also visualize this local infection process by examining the states at different times during a realization Figure 3.12. In this figure, the infection quickly dies out, and there is no infection left.

We can also examine what may happen if an infection were able to establish in this local neighbourhood. We set $\gamma = 0.2$ day^{-1}in order to examine what may

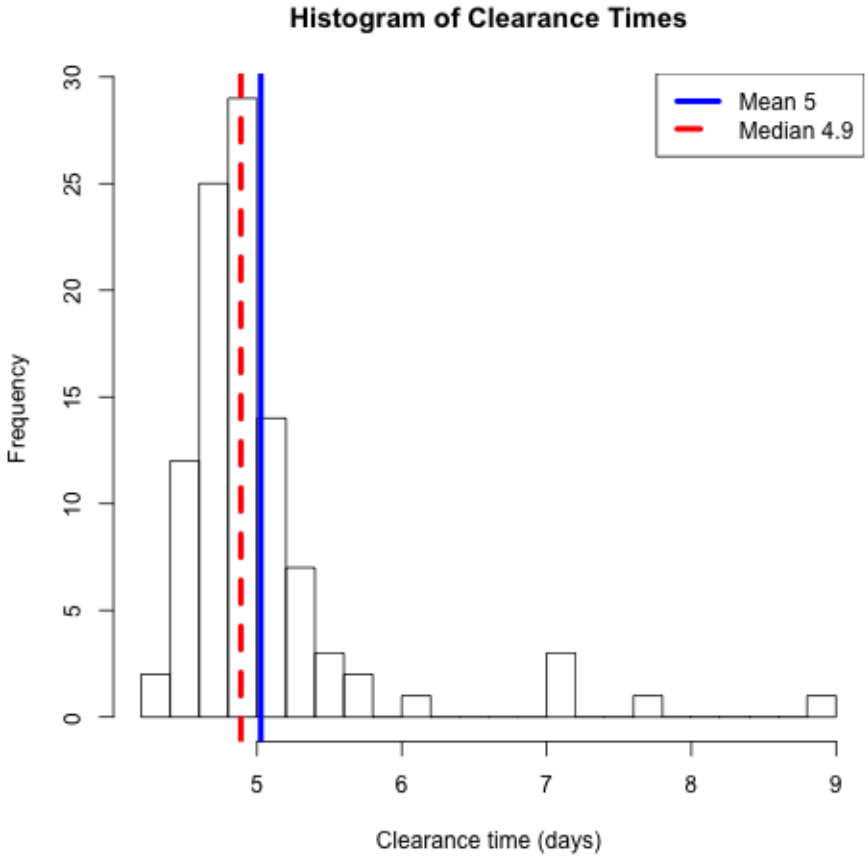

Figure 3.11: Histogram for clearance time given the local neighbourhood simulation using the base parameter values. The mean value is 5.02 days and the median 4.89 days.

842 happen when an infection is able to establish. In this scenario, the propagation rate
843 of the primed T-cells is so low that the infection will be able to spread to surrounding
844 susceptible sites before it can be cleared by the present immune response. We set these
845 parameter values to illustrate the local infection process. We display the aggregated
846 time series in Figure 3.13. In Figure 3.14, we take snapshots of the infection status of
847 the sites at different points in time. We see here, that the infection is sustained, but

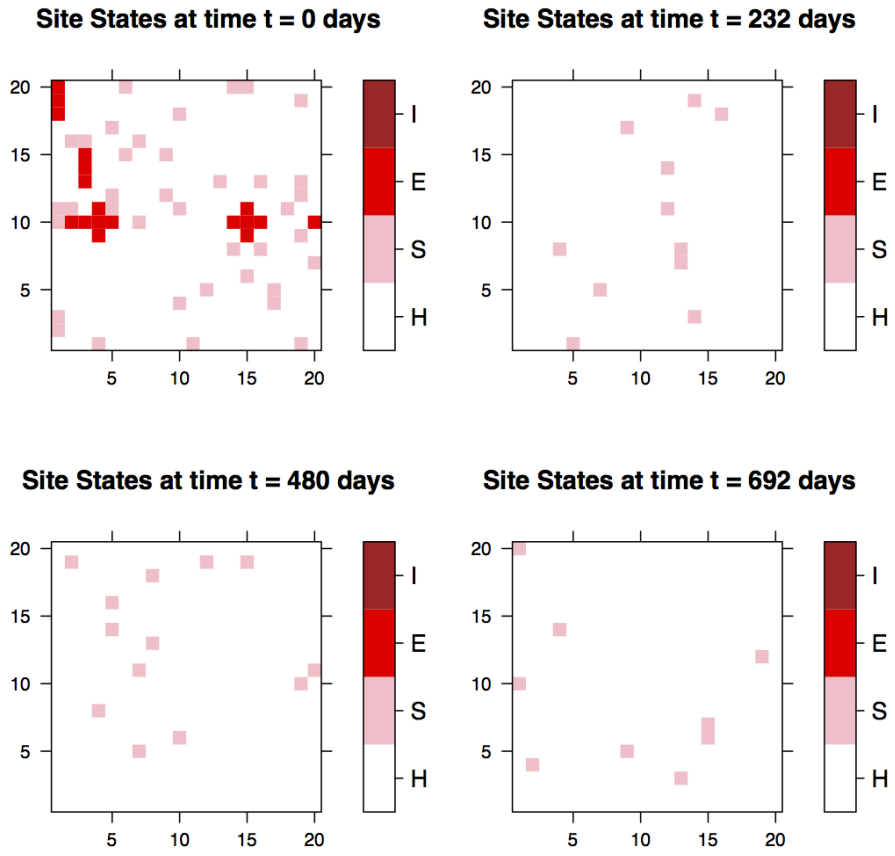

Figure 3.12: Here is a snapshot of a local neighbourhood realization for different times. Here we can see that the infection quickly dies out.

also spreads locally from the initial point of infection. For this set of parameters, we run the simulation for one year, $t = 365$ days, and for 50 realizations. The infection established 100% of the time and is cleared 0% of the time after one year of infection. Because this scenario does not confer any clearance after one year, these parameters are likely to not be correct as the results are not consistent with the 80-90% clearance rate observed in population data.

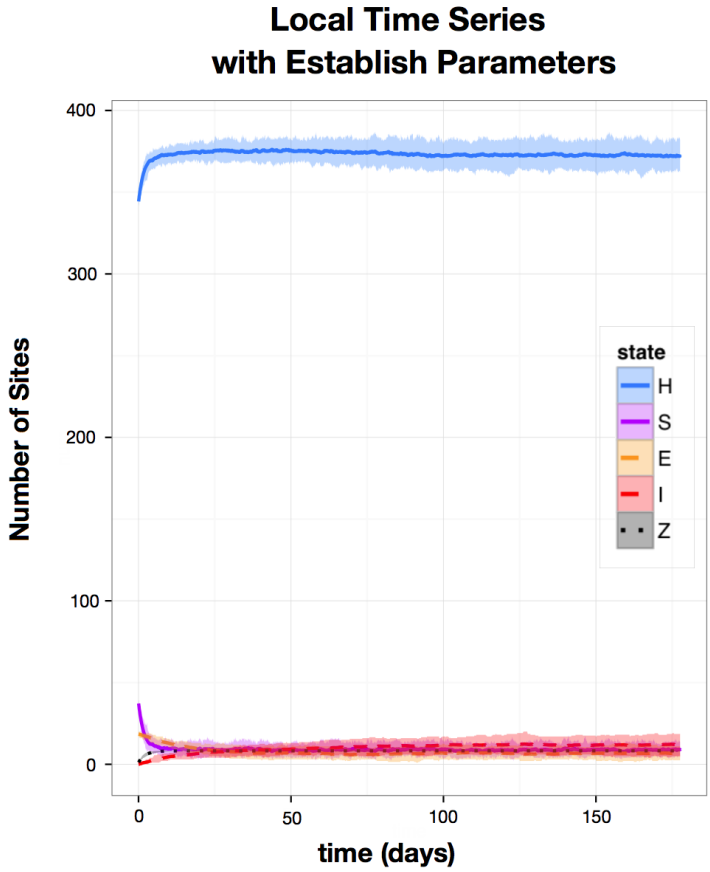

Figure 3.13: Here is the time series for the the local neighbourhood simulation using the establish parameters. These parameters are the same as the base parameters, but setting $\gamma = 0.2\text{day}^{-1}$. We see here that the infection is able to establish.

3.4 Discussion

We adapted the base within-host model discussed in Chapter 2 to a stochastic spatial model. We considered two different neighbourhoods, a global and a local neighbourhood. The global neighbourhood mimics a homogeneously mixed population and was used to compare the event rates to the base deterministic model. These models strongly agreed within one another and had similar dynamics, suggesting that

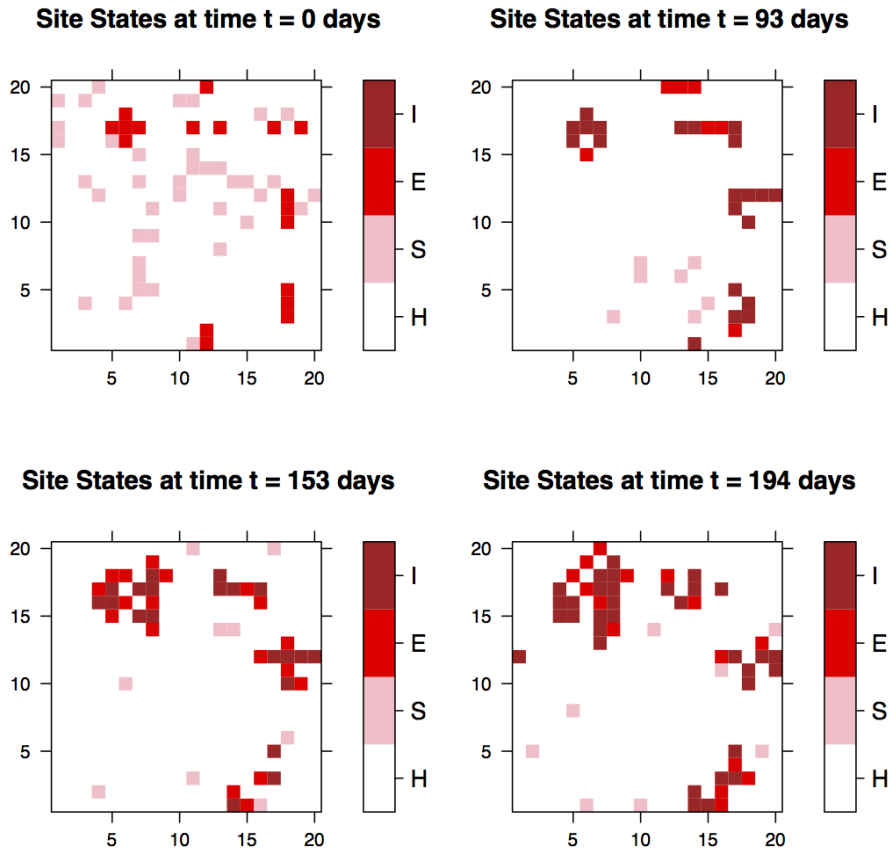

Figure 3.14: Here is a snapshot of a local neighbourhood realization for different times using the establish parameters, $\gamma = 0.2\text{day}^{-1}$. Here the infection does not die out right away, and we can see how it spreads in a local fashion.

the event rates were properly calibrated. Using the base model parameters, it was shown that clearance occurred 0% of the time, which is very unlikely given that most infections are cleared about 80-90% of the time [2, 33, 38, 57]. This suggests that the cell-mediated immune response is greater than in the parameters defined in base model (Table 3.2). By changing the immune parameters, rate of clearance due to primed T-cells $\alpha = 2$ day^{-1}and rate of primed T-cell propagation $\gamma = 14$ day^{-1}, we

were able to obtain realistic clearance values. Under these parameters we observed a clearance rate of 78%, which is similar to the population clearance rates [2,33,38,57]. Furthermore, the mean time to clearance was determined to 260.8 and 95% QR (73.7, 541.7) days, median 250 days, or about 9 months. This may suggest that the immune parameters are greater than previously thought.

Comparing the global and local neighbourhood simulation results provides some interesting outcomes. The local neighbourhood provided an extreme case of locality, where infectious sites were only able to infect susceptible sites directly in contact with the infectious site. The local structure significantly reduces the ability for the infection to establish before being cleared by the immune system when compared to the global neighbourhood structure. For the base parameters, the global model resulted in persistent infections for each realization, while the local model with the same parameters resulted in complete clearance for all realizations. In fact, using these same parameter values, we saw 100% clearance with a mean clearance time of 5.02 days, median 4.89 days. This is much shorter than those infections discussed in literature [57] Therefore, locality can play an important role whether an HPV infection will establish or die out in the host.

We also explore the case when the infection is able to establish in the case of the local neighbourhood structure. In this scenario, we set $\gamma = 0.2$ day^{-1}, and there was 100% establishment with 0% clearance for 50 realizations after one year. Because of the large difference in results between the global neighbourhood and local neighbourhood, it is unlikely that HPV in fact spreads in either of these extremes completely.

It is unlikely that HPV truly acts in either of these extreme scenarios, but rather

most likely has some locality but with a larger neighbourhood than solely those in direct contact with the focal site. When developing within-host mathematical models for parameter fitting to data, a homogenous mixing model may over or under estimate parameter values. For example, in our global neighbourhood, we were able to obtain approximately 89% clearance by using the the base parameter values but changing $\gamma = 14$ day^{-1}and $\alpha = 2$ day^{-1}. In the other extreme, using the base parameters in the local neighbourhood simulation, we obtain 100% clearance. This suggests that the establishment or clearance of an infection is significantly affected by the spatiality of the infection process. Therefore, these points should be considered when building within-host models, especially when using these models to fit parameters to data.

Chapter 4

Multistrain HPV Models

4.1 Introduction

The human papillomavirus has been shown to be responsible for almost all cases of cervical cancer, and is also highly associated with a number of other cancers such as anal cancer and oropharyngeal cancers. There are over 100 different types of HPV, over forty of which infect the anogenital tract. HPV types are differentiated from one another by the genetic sequence of the L1 capsid protein of the virus [19, 25]. HPV types are also categorized into low-risk and high-risk types based on their association with the development of cancerous and pre-cancerous lesions. Furthermore, HPV types are categorized into various genera. We focus on the alpha (α) genus as these types infect the mucosal regions. HPV types are further divided into different species within the α-genus based on phylogentic differences.

Two important types of HPV are known to be the cause of the majority of cervical cancer cases: HPV-16 and HPV-18. HPV-16, a member of the α-9 species, is associated with 50% of cervical cancer cases, and HPV-18 (α-7) is associated with

20% [25]. To combat the burden of these HPV types, two different vaccines have been developed to protect against these two high-risk HPV types. Cervarix^{TM} is a bivalent vaccine that protects against these two high-risk types. Gardasil^{TM}, on the other hand, is a quadravalent vaccine and protects against these two high-risk types alongside two low-risk types—HPV-6 and -11—which are highly associated with genital warts. Both of these vaccines have shown significant levels of efficacy and are thought to be successful in reducing the number of high-risk HPV infections and subsequently the number of cervical cancer cases.

While these vaccines show strong efficacy in protection against these two HPV types, there are other high-risk types that these vaccines do not confer protection against. This has led researchers to explore the potential for "type replacement" [16, 28,35,41,56]. This is an ecological phenomenon wherein the protection against certain types of a pathogen increases the niche space for other pathogen types, potentially increasing the prevalence of these non-vaccine pathogen types. If type replacement were to occur alongside the vaccine program, then there may be an increase in high-risk, non-vaccine HPV types, which could limit the estimated protective effects of the HPV vaccine. Because these HPV vaccines are relatively new to the public, there is limited longitudinal data to support or refute the potential of HPV type replacement. There has been one study so far that illustrated a possible case of type replacement in young women [29]. It found that there was an increased prevalence of high-risk, non-vaccine HPV in young girls who were vaccinated but not in those girls who were not vaccinated. While this may provide some limited support for type replacement, more careful studies should be conducted to examine the possibility of type replacement in other populations.

In order to estimate the potential for HPV type replacement before it occurs in the population, researchers examined these questions using mathematical models. Past models have examined the potential for HPV type replacement at the population level. Major findings of these models show that type replacement will only occur if HPV types are competitive, in that they will fight for space or resources in the host [16, 28]. If types are independent or facilitative, then type replacement will not occur. In fact, if two HPV types are facilitative, then a vaccination effort may have additional benefits in reducing the prevalence of the facilitative, non-vaccine type alongside the vaccine type. It should be noted that these scenarios are contingent upon the ecological interactions between the two HPV types within the host.

It is not yet fully understood how different HPV types interact within the host. Population level studies show that infections with multiple HPV types are quite common. The Centers for Disease Control and Prevention report that 5 to 30% of individuals infected with mucosal HPV are also infected with multiple types [25]. Because of this high rate of multiple type infections, some researchers believe that HPV types interact independently or synergistically within the host [13, 56]. This high rate of multiple infections may not be due to within host facilitation. Rather they may be caused by an individual's behaviour increasing the likelihood of being multiply infected with various HPV types. It is known that risk factors for HPV infection are predominantly based on sexual behaviours such as lifetime and recent sex partners, and these factors rather than type facilitation may be driving multiple HPV infections in some individuals [6, 57]. While a consensus for how HPV types compete or do not compete is still not clear, there are some forms of competition that are known.

It is known that there are some competitive interactions between HPV types, in

particular through cross-reaction in the immune system. Williams et al. [61] examined the T-cell cross-reactivity response between HPV 11 and other types. Furthermore, Scherpenisse et al. [46] showed cross-reactivity in the antibody response between HPV types within the same species (α-9 and α-7 species). These types of competition are referred to as apparent competition, as the presence of one pathogen can drive the elimination of the other through a cross-reactive immune response.

Along with cross-reactivity, there is also evidence of resource competition in HPV coinfections. Because HPV types have a large portion of shared DNA, it is likely that they require similar cellular functions in order to produce viral particles. MacLaughlin et al. [32] examined the effects of coinfection on the virus production in an *in vivo* study. They found that cells coinfected with HPV-18 and another HPV type resulted in lower viral counts in the other HPV types compared to singly infected cells. In all cases, HPV-18 dominated cell resources, which may explain the success of HPV-18 in developing persistent infections in women. Xi et al. [62] also compared viral loads in coinfected patients with viral loads in singly infected patients. They found that HPV-16 and HPV-18 viral loads were significantly lower in coinfections with HPV types from the same α-9 and α-7 species, respectively. They discuss the possibility of resource competition as the cause of the viral load reduction. However, they also attribute the magnitude of the decrease in part to cross-reactivity in the immune response.

In this chapter we explore the potential for coexistence between two competing types of HPV. We introduce two competing HPV types within our model, each

which elicits a separate but cross-reactive immune response. These HPV types compete directly through space competition and indirectly through immune system cross-reactivity (apparent competition). We show that for certain levels of cross-reactivity, coexistence is possible. Restricted to only two years, the unofficial threshold for persistent infections, we see that effective coexistence is possible for even larger levels of cross-reactivity than in the long term scenario. We do not add to the debate about whether type replacement is occurring or if it surely will or will not occur. We solely examine the coexistence of multiple HPV types in the host even in the presence of competition. This suggests that competition cannot be ruled out solely because of high rates of HPV coinfection in hosts.

4.2 Methods

We develop a mathematical, multi-strain model for the spread of HPV within the host. The multi-strain model considers two different HPV strains infecting the surface of the epithelium. Our model considers apparent competition through cross-reactivity in the immune response. We report these results in Section 4.3

4.2.1 The Model

We consider a similar compartmental model to System (1) where sites in the cervix epithelium are categorized by infection status: healthy (H), susceptible (S), exposed (E), and infectious (I). In this model, we consider two pathogens, and thus sites may be exposed or infectious with one of two types (E_1, I_1, E_2, or I_2). We distinguish between type-specific virus particles. That is, sites infectious with type one (I_1) will

produce V_1 virus particles and those infectious with type two (I_2) will produce V_2 virus particles. We also consider type specific immune responses, that is we have T-cells that are primed against one type (Z_1) and T-cells primed against the other type (Z_2). We discuss various forms of competition that may play a role in HPV dynamics.

A within-host, multistrain model has been examined previously by Murall et al. [35]. Specifically, they review space competition, resource competition, and apparent competition but also type facilitation and independence. Murall et al. [35] consider resource competition as the acquisition and use of cellular processes in the development of viral particles in coinfected sites. While an *in vitro* study [32] has shown that different HPV strains are able to coinfect the same site, this same phenomenon has not yet been examined *in vivo*. Furthermore, in almost all cases in the study, one HPV type (specifically HPV 18) dominated cell functions over the other coinfected type. In fact, viral production of the weaker coinfected type was significantly less in coinfected sites than in singly infected sites. This reduction was not recorded for the dominant type, however.

Murall et al. consider the scenario of independence between HPV types, which they define as no interaction at all. That is, coinfection is allowed unhindered, and viral particles are able to produce viral particles of both types at the same rate as solely infected patches. This means the coinfected sites are able to produce virus past the normal capacity of the cells, which is biologically improbable. This implies that independence is unlikely both between sites and within sites. The dominance of HPV-18 viral production in coinfected sites [32] even suggests that neutrality may be unlikely within the same host site. However, more work should be done to rule out

neutrality. Furthermore, care must be taken for those wishing to examine the effects of coinfection in patch models such as this; asymmetry through the super-infection process may be unintentionally introduced into the model. This is highlighted in more detail in Appendix C.

Furthermore, our model induces the necessity of abrasions for an infection event to occur. To model the potential for coinfection, it would be required to keep track of the abraded and non-abraded exposed and infectious sites for each strain, which complicates a multiple HPV type super-infection model significantly. Because of this, and with the unlikeliness of neutral interactions within the host site, we do not include this form of competition into our model. We introduce two forms of competition through space competition—HPV types are not able to super-infect the same site—and apparent competition—HPV types may elicit and be cleared by an immune response primed by a different type. The mechanisms for apparent competition are highlighted in more detail below (Section 4.2.1 Apparent Competition).

Apparent Competition

We introduce apparent competition through cross-reactivity of similar HPV types. The factor for cross-reactivity q varies between 0 and 1. When $q = 0$, that means that the two types are not cross-reactive at all, and completely independent immune responses are established. When $q = 1$, the types are 100% cross-reactive, and each immune response can identify and clear both of the HPV types. There are two ways we introduce cross-reactivity into the model: through propagation of T-cells primed

against one strain by coming into contact with the other strain:

$$\frac{\mathrm{d}Z_1}{\mathrm{d}t} = \zeta + \frac{\gamma_1 Z_1(E_1 + I_1)}{N} + q\frac{\gamma_1 Z_1(E_2 + I_2)}{N} - \mu Z_1 \tag{1a}$$
$$\frac{\mathrm{d}Z_2}{\mathrm{d}t} = \zeta + \frac{\gamma_2 Z_2(E_2 + I_2)}{N} + q\frac{\gamma_2 Z_2(E_1 + I_1)}{N} - \mu Z_2 \tag{1b}$$

As well, we consider cross clearance of one type of pathogen by T-cells primed against the other strain:

$$\frac{\mathrm{d}E_1}{\mathrm{d}t} = \beta_1 S V_1 - \sigma E_1 - \frac{\alpha_1 Z_1 E_1}{N} - q\frac{\alpha_2 Z_2 E_1}{N}, \tag{2a}$$
$$\frac{\mathrm{d}I_1}{\mathrm{d}t} = \sigma E_1 - \frac{\alpha_1 Z_1 I_1}{N} - q\frac{\alpha_2 Z_2 I_1}{N}, \tag{2b}$$
$$\frac{\mathrm{d}E_2}{\mathrm{d}t} = \beta_2 S V_2 - \sigma E_2 - \frac{\alpha_2 Z_2 E_2}{N} - q\frac{\alpha_1 Z_1 E_2}{N}, \tag{2c}$$
$$\frac{\mathrm{d}I_2}{\mathrm{d}t} = \sigma E_2 - \frac{\alpha_2 Z_2 I_2}{N} - q\frac{\alpha_1 Z_1 I_2}{N}. \tag{2d}$$

Using these competition interactions, we introduce our full multistrain model for two HPV types. The system of differential equations defining our model can be found in System (3), and a flow diagram visually explaining the model can be found in Figure 4.1.

$$\begin{aligned}\frac{\mathrm{d}H}{\mathrm{d}t} = & -\chi H + \rho + \frac{(\alpha Z_1 + q\alpha Z_2)(E_1 + I_1)}{N} \\ & + \frac{(\alpha Z_2 + q\alpha Z_1)(E_2 + I_2)}{N} \end{aligned} \tag{3a}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \chi H - \rho S - \beta_1 \frac{SV_1}{N} - \beta_2 \frac{SV_2}{N}, \tag{3b}$$

$$\frac{\mathrm{d}E_1}{\mathrm{d}t} = \beta_1 \frac{SV_1}{N} - \sigma E_1 - \alpha_1 \frac{Z_1 E_1}{N} - q\alpha_2 \frac{Z_2 E_1}{N}, \tag{3c}$$

$$\frac{\mathrm{d}E_2}{\mathrm{d}t} = \beta_2 \frac{SV_2}{N} - \sigma E_2 - \alpha_2 \frac{Z_2 E_2}{N} - q\alpha_1 \frac{Z_1 E_2}{N}, \tag{3d}$$

$$\frac{\mathrm{d}I_1}{\mathrm{d}t} = \sigma I_1 - \alpha_1 \frac{Z_1 I_1}{N} - q\alpha_2 \frac{Z_2 I_{12}}{N} \tag{3e}$$

$$\frac{\mathrm{d}I_2}{\mathrm{d}t} = \sigma I_2 - \alpha_2 \frac{Z_2 I_2}{N} - q\alpha_1 \frac{Z_1 I_{12}}{N} \tag{3f}$$

$$\frac{\mathrm{d}V_1}{\mathrm{d}t} = f_1 I_1 - \delta_1 V_1 \tag{3g}$$

$$\frac{\mathrm{d}V_2}{\mathrm{d}t} = f_2 I_2 - \delta_2 V_2 \tag{3h}$$

$$\frac{\mathrm{d}Z_1}{\mathrm{d}t} = \zeta N + \frac{\gamma_1 Z_1 (E_1 + I_1)}{N} + q\frac{\gamma_1 Z_1 (E_2 + I_2)}{N} - \mu Z_1 \tag{3i}$$

$$\frac{\mathrm{d}Z_2}{\mathrm{d}t} = \zeta N + \frac{\gamma_2 Z_2 (E_2 + I_2))}{N} + q\frac{\gamma_2 Z_2 (E_1 + I_1)}{N} - \mu Z_2 \tag{3j}$$

We numerically solve this system of differential equations using the `ode()` function in **R**. We solve for the equilibria by running the system until $t = 10,000$ days, trying to reach equilibrium. We find the values for the infection equilibria $\overline{E}$ and $\overline{I}$ in both strains for different values of cross-reactivity, q. We also consider the same system, but taking the values of $\overline{E}$ and $\overline{I}$ after two years, $t = 730$ days. The definition of persistent infection is not yet well defined. However, it is commonly thought to be an infection that lasts longer than two years. These results can be seen in Figure 4.3.

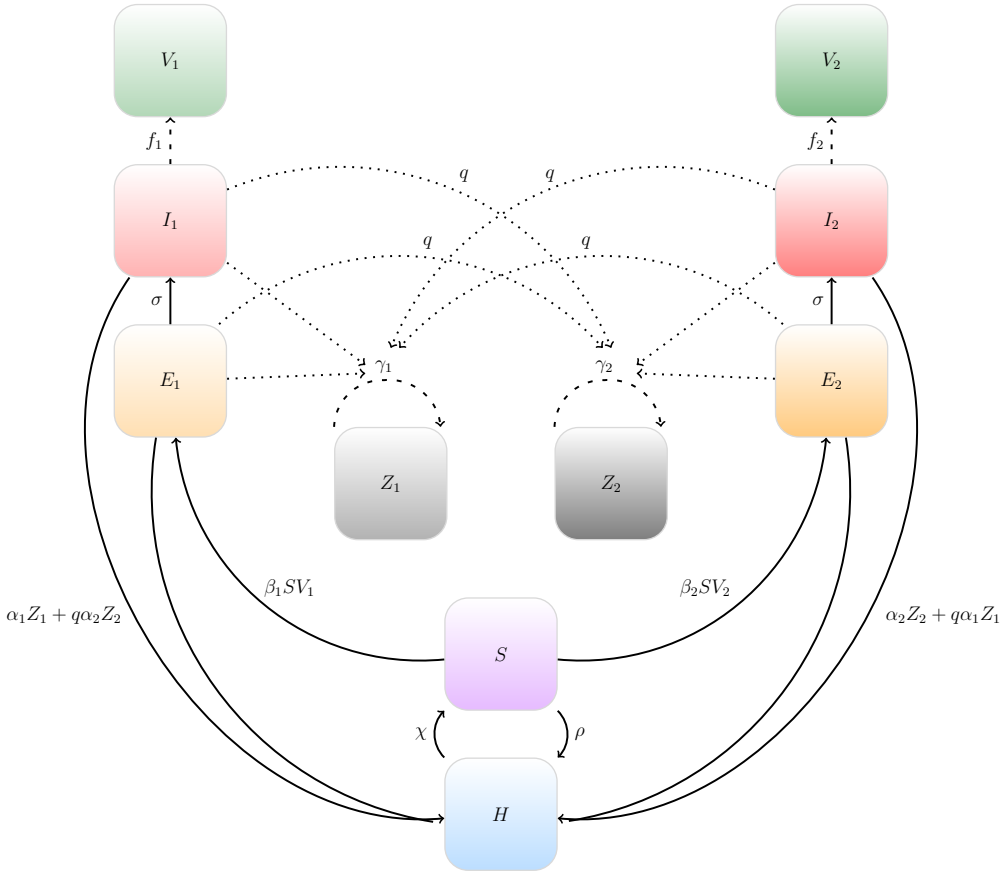

Figure 4.1: A flow diagram illustrating the within-host multi-strain HPV model. The cross-reactive term q can be seen in how the exposed and infectious sites affect the propagation of strain-primed T-cells and also in the clearance of exposed and infectious sites.

4.3 Results

We analyze this system by examining how the system evolves for different values of cross-reactivity, $q \in [0, 1]$. We define the elimination of one strain when the number of exposed or infectious sites goes below one site. If the equilibrium value is below one, then on average it will result in clearance of that strain in an actual host. Looking when the number of exposed and infectious sites of a particular strain dips below one

(horizontal gray line, Figure 4.2), we see that coexistence of the two HPV strains is present for values of about $q < 0.70$. For values $q > 0.70$ we see that the second strain (the "weaker" strain) dies out. This shows that if the HPV types illicit distinct but cross-reactive immune responses, then coexistence is possible even when space competition is present and the two strains are not able to infect the same sites. This is illustrated in Figure 4.2.

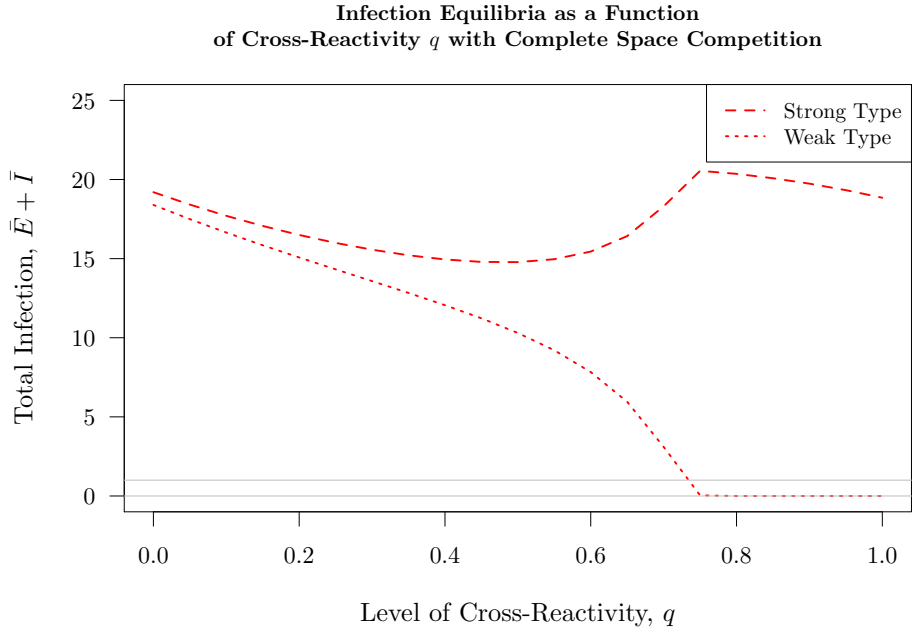

Figure 4.2: This plot shows the values of the disease equilibrium, $\overline{E}$ and $\overline{I}$, as a function of the cross-reactivity factor q in the case with complete space competition $x = 0$. It can be seen that we have coexistence for both strain 1 and 2 for values $q < 0.70$. Large values of q eliminates strain 2, the "weaker" strain, from the system.

When we examine the curves in more detail, we can observe some interesting patterns. Consider first the curves for strain 2 (dotted) in Figure 4.2. We see that as the level of cross-reactivity increases, both the curves for the exposed sites and

infectious sites decrease. This is because as cross-reactivity is increased, the presence of strain 1 increases the cross-reactive immune response against strain 2 enough to deplete its numbers quite rapidly. This pattern continues until about $q > 0.75$ where the strain 2 is depleted completely (flat line at $y = 0$). We induce "clearance" before this, however. When strain 2 cross the gray line $y = 1$, then the strain will be eliminated from the system necessarily as there are fewer than 1 infected individuals in this case.

Now, examining the curves for strain 1 (dashed) we see some more complex dynamics. Similarly, as we increase the level of cross-reactivity from $q = 0$, we see a reduction in the number of exposed (orange) and infectious (red) sites of strain 1. This is again due to the increase in cross-reactive immune responses due to the presence of strain 2. However, for larger q values, approximately when $0.6 < q < 0.75$ then the equilibrium values for the exposed and infectious sites of strain 1 increase. This is because the numbers of strain 2 infected sites are so low, that strain 1 is able to recover from the cross-protective effects. We also see a reduction in both the exposed and infectious sites of strain 1 when $q > 0.75$, which occurs when strain 2 has been eliminated from the system. This is because strain 1 is cross-reacting with the base immune response induced against strain 2, even when there is not any pathogen of this type present in the system.

Examining the equilibria of $\overline{E}$ and $\overline{I}$ after two years, $t = 365(2) = 730$ days (Figure 4.3), it can be seen here that coexistence is possible after two years for larger values of q. In fact, we see that coexistence is possible for all values $q < 0.8$. However, the equilibrium value in this case are very low, and clearance due to stochastic effects becomes quite likely. The dynamics of these curves are similar to those found in the

case when we let the system go to equilibrium (Figure 4.2), but are not as striking because the system has not completely settled to the equilibrium value.

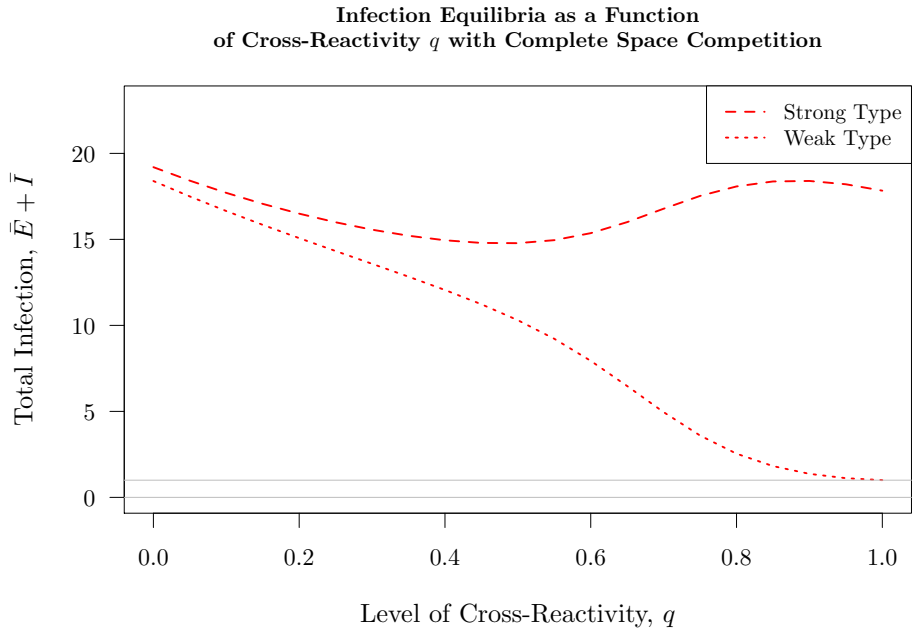

Figure 4.3: This plot shows the values of $\overline{E}$ and $\overline{I}$ after two years as a function of the cross-reactivity factor q in the case with complete space competition $x = 0$.

Importantly here, we can see that coexistence is possible for different values of $q < 0.70$. This shows that competition between types within the host still confers coexistence between the two strains. Thus, the high rates of multi-type infections do not show that the HPV types are not competing. Previous mathematical models suggest that competition is a requirement for type replacement [16], and current epidemiological data have been used to suggest that the high rates of coinfection with multiple types at the population level do not support competitive interactions between HPV [13, 56]. However, we have shown that even when considering space competition and cross-reactivity, coexistence is possible within the host. That is,

coexistence between strains may still possible and the high rates of multiple strain infections are not enough to rule of competition between HPV types. More work should be done to consider the effects of within-host dynamics on the population dynamics. Furthermore, HPV type replacement cannot necessarily be ruled out.

4.4 Discussion

We developed a two-type model for HPV that takes into account separate immune responses, but which are linked through a cross-reactivity term. This is a form of apparent competition that is very common in infectious diseases. We also induce further competition through space competition. Sites can only be infected by one HPV type. While there is *in vitro* evidence of type coexistence within the same infected host cell [32], we do not incorporate this scenario into our model. Including this would significantly complicate the model. Because we account for an abrasion process in our infection cycle, we would be required to keep track of all abraded and non-abraded infectious, exposed, and uninfected sites. Thus we decided to focus on only space competition and apparent competition in our model.

We find that coexistence between the two HPV types is possible when the the level of cross reactivity is below 70% cross-reactive. Furthermore, when restricting the time to only two years, which is the unofficial threshold for persistent infections, we see that coexistence is possible when the types are less than 80% cross reactive. This considers strains of two different strengths. In all cases where coexistence is not possible, the weaker strain dies out.

This shows that coexistence of HPV types is possible in the same host even though the types may be competitive. This contradicts some theories that HPV types are

not competitive with one another because of the high rates of coinfection in the same patients. While the potential for type replacement has been determined to be reliant on competition between HPV types, we do not make any claims that HPV type replacement will or will not occur from this model. To do something like that, researchers should consider linking within-host models with population level models in some way. While HPV types may be competitive within the same host, how these effects translate to population level effects is still not completely understood.

Chapter 5

Conclusion

This thesis establishes a mathematical model with which to examine the spread of HPV within the host. We reviewed the biological mechanisms of infection and clearance by the immune system to inform the construction of our model. Using this model, we solved for the healthy equilibrium HE, when no HPV is present, and the diseased equilibrium DE, when HPV is able to establish in the host. We found an expression for the basic reproduction number $\mathcal{R}_0$, which provides a threshold for whether an infection will be able to establish in the host and potentially persist. We also examined some alterations to the model by including memory cells and a delay in the immune response. By including memory cells, an infection was able to establish if $\mathcal{R}_0 > 1$ but could be completely cleared if the effective reproduction number dipped below 1, $\mathcal{R}_e < 1$. The delay in the immune response was found to not have much change in the dynamics of the disease unless the immune response is never triggered. We also examined how the parameter values effect the value of the diseased equilibrium. We found that while γ, the rate of immune cell propagation, did not have an effect on whether the infection will establish, it did have an important effect on

the value of the diseased equilibrium. If γ is large enough, the diseased equilibrium value decreased enough for clearance to occur through stochastic effects, which could explain the 80-90% rate of clearance of infections after two years.

We also adapted this mathematical model into a stochastic spatial model. We introduced two different neighbourhood structures (global and local) to examine how spatiality may affect the spread, establishment, and clearance of an infection. We found that for a set of base parameters estimated from literature, the global neighbourhood model resulted in 100% establishment with 0% clearance. On the other hand, running the same parameter values on the local neighbourhood model resulted in no establishment and 100% clearance with a mean time to clearance of 5.12 days. Neither of these cases seemed to encapsulate the actual rate of clearance of infections (80-90%), so it is unlikely that a typical HPV infection spreads completely locally or completely globally. We also showed that the locality of the infection process in within-host models has an important impact on the establishment or clearance of an infection. This has important implications when developing within-host models to fit parameters to data. Depending on how you define the spatiality of the model, parameter values may be under- or over-estimated.

Lastly, we examined the potential for coexistence between two competing HPV strains. We explored two forms of competition between the HPV types: space competition, where HPV types compete for infection sites, and apparent competition through immune cross-reactivity. We also discussed the potential for resource competition, the competition for cellular processes and resources in cells coinfected with multiple types. However, we found that this form of competition is somewhat unlikely for HPV, and importantly difficult to implement in our model without inducing some

form of unintended bias. This bias is discussed in more detail in Appendix C. We found that when the level of cross-reactivity was below $q < 0.70$, or rather the types were 70% cross-reactive, then coexistence is possible. This suggests that researchers are unable to use the high rates of multiple type infections as evidence that HPV types are not competitive. While we did not add to the debate for the potential of type replacement, we showed that because competition between types in the host cannot be ruled out, type replacement should also not be ruled out as a potential outcome of vaccination.

Ultimately, this thesis provided a basis with which to examine the spread of HPV within the host using mathematical models. These models discussed in this thesis can be adapted to help answer specific open questions. When developing complex mathematical models to help understand complicated viral dynamics, it is important to start with a base model and add complexities where necessary. This thesis provided groundwork with which to develop more specific within-host HPV models.

Appendices

Appendix A

Deterministic Analysis

A.1 Derivation of $\mathcal{R}_0$

We confirm our biological interpretation of the basic reproduction number by solving for $\mathcal{R}_0$ using the next generation matrix method developed by van den Driessche and Watmough [59]. Consider the set of disease free states X_s,

$$X_s = \{x \geq 0 \,|\, x_i = 0, i = 1, ..., m\}, \tag{A.1.1}$$

where m is the number of compartments that refer to infected agents. Without loss of generality, order the compartments such that the infected compartments come first. For our model, we have

$$X_s = \{E_s, I_s, V_s, H_s, S_s, Z_s\} = \{0, 0, 0, H_s, S_s, Z_s\} \tag{A.1.2}$$

Because we are concerned with whether the infection will be able to effectively reproduce, we will examine the next generation matrix proposed by van den Driessche and Watmough of the infectious units. In our model, the infectious units are the exposed and infected classes E and I, and also the virus V.

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \beta SV - \sigma E - \alpha ZE, \tag{A.1.3a}$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E - \alpha ZI, \tag{A.1.3b}$$
$$\frac{\mathrm{d}V}{\mathrm{d}t} = fI - \delta V. \tag{A.1.3c}$$

We can partition these differential equations into two different parts. That is, $\mathcal{F} - \mathcal{V}$, where $\mathcal{F}$ refers to *new* agents in each class, and $\mathcal{V}$ represents all other movement between the classes, such as elimination or movement from one class to another. Thus we have that

$$\mathcal{F} = \begin{pmatrix} \beta SI \\ 0 \\ fI \end{pmatrix}; \ \mathcal{V} = \begin{pmatrix} \sigma E + \alpha ZE \\ -\sigma E + \alpha ZI \\ \delta V \end{pmatrix} \tag{A.1.4}$$

We can consider the Jacobian of the system, which can be simplified to the matrices $\mathbf{F} = \frac{\partial \mathcal{F}}{\partial x}$ and $\mathbf{V} = \frac{\partial \mathcal{V}}{\partial x}$.

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & \beta S \\ 0 & 0 & 0 \\ 0 & f & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \frac{\beta\chi}{\chi+\rho} \\ 0 & 0 & 0 \\ 0 & f & 0 \end{pmatrix} \tag{A.1.5a}$$

$$\mathbf{V} = \begin{pmatrix} \sigma + \alpha Z & 0 & 0 \\ -\sigma & \alpha Z & 0 \\ 0 & 0 & \delta \end{pmatrix} = \begin{pmatrix} \sigma + \frac{\alpha\zeta}{\mu} & 0 & 0 \\ -\sigma & \frac{\alpha\zeta}{\mu} & 0 \\ 0 & 0 & \delta \end{pmatrix} \tag{A.1.5b}$$

The next generation matrix is defined as $\mathbf{FV}^{-1}$. van den Driessche and Watmough define $\mathcal{R}_0$ as the spectral radius of the next generation matrix $\rho(\mathbf{FV}^{-1})$.

$$\mathbf{FV}^{-1} = \begin{pmatrix} 0 & 0 & \frac{\beta\chi}{\chi+\rho} \\ 0 & 0 & 0 \\ 0 & f & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\frac{\alpha\zeta}{\mu}+\sigma} & 0 & 0 \\ \frac{\mu\sigma}{\alpha\zeta(\frac{\alpha\zeta}{\mu}+\sigma)} & \frac{\mu}{\alpha\zeta} & 0 \\ 0 & 0 & \frac{1}{\delta} \end{pmatrix} \tag{A.1.6a}$$

$$= \begin{pmatrix} 0 & 0 & \frac{\beta\chi}{\chi+\rho} \\ 0 & 0 & 0 \\ \frac{f\mu\sigma}{\alpha\zeta(\frac{\alpha\zeta}{\mu}+\sigma)} & \frac{f\mu}{\alpha\zeta} & 0 \end{pmatrix} \tag{A.1.6b}$$

The eigenvalues of this system are the following.

$$\lambda_0 = 0 \tag{A.1.7a}$$

$$\lambda_\pm = \pm\mu\sqrt{\frac{\beta\chi\sigma}{\delta\alpha\zeta(\rho+\chi)(\alpha\zeta+\mu\sigma)}} \tag{A.1.7b}$$

Therefore, the spectral radius is $\bar{\rho} = \mu\sqrt{\frac{\beta\chi\sigma}{\delta\alpha\zeta(\rho+\chi)(\alpha\zeta+\mu\sigma)}}$. However, recall that our infectious cycle requires going through two steps before we consider the full generation. Firstly, viral particles are produced and infect another site, then this site must survive latency. In that require we require two generations, and we square the spectral radius. This formalized the final expression for the basic reproduction number:

$$\mathcal{R}_0 = \frac{f\beta\chi\sigma\mu^2}{\delta\alpha\zeta(\chi+\rho)(\sigma\mu+\alpha\zeta)} \tag{A.1.8}$$

This confirms our biological derivation of the basic reproduction number.

A.2 Linearization of the Healthy Equilibrium, *HE*

To examine the stability of this system, we linearize the system at the *HE* and examine the eigenvalues. To simplify our system, we will replace $H = 1 - S - E - I$, which removes the variable H from the system,

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \chi(1-S-E-I) - \rho S - \beta V\frac{S}{N}, \tag{A.2.9a}$$
$$\frac{\mathrm{d}E}{\mathrm{d}t} = \beta V\frac{S}{N} - \sigma E - \alpha Z\frac{E}{N}, \tag{A.2.9b}$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E - \alpha Z\frac{I}{N}, \tag{A.2.9c}$$
$$\frac{\mathrm{d}V}{\mathrm{d}t} = fI - \delta V, \tag{A.2.9d}$$
$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \zeta N + \gamma Z\frac{(E+I)}{N} - \mu Z. \tag{A.2.9e}$$

We will now consider the Jacobian of this simplified system.

$$DF_{\bar{x}} = \begin{pmatrix} -\chi-\rho-\beta V & -\chi & -\chi & -\beta S & 0 \\ \beta V & \sigma-\alpha Z & 0 & \beta S & -\alpha E \\ 0 & \sigma & -\alpha Z & 0 & -\alpha I \\ 0 & 0 & f & -\delta & 0 \\ 0 & \gamma Z & \gamma Z & 0 & \gamma(E+I)-\mu \end{pmatrix} \tag{A.2.10a}$$

We then evaluate this Jacobian at the healthy equilibrium HE.

$$DF_{HE} = \begin{pmatrix} -\chi-\rho & -\chi & -\chi & \frac{-\beta\chi}{\chi+\rho} & 0 \\ 0 & \sigma-\frac{\alpha\zeta}{\mu} & 0 & \frac{\beta\chi}{\chi+\rho} & 0 \\ 0 & \sigma & \frac{-\alpha\zeta}{\mu} & 0 & 0 \\ 0 & 0 & f & -\delta & 0 \\ 0 & \frac{\gamma\zeta}{\mu} & \frac{\gamma\zeta}{\mu} & 0 & -\mu \end{pmatrix} \tag{A.2.11}$$

To simplify the system, we rearrange the matrix accordingly to break this system into blocks. The top, left block considers the healthy and susceptible compartments (S) of our system; the middle block considers the immune response (Z); and the bottom, right block considers the infection states (E, I, V)

$$DF_{HE} = \left(\begin{array}{c|c|ccc} -\chi-\rho & -\chi & -\chi & \frac{-\beta\chi}{\chi+\rho} & 0 \\ \hline 0 & -\mu & \frac{\gamma\zeta}{\mu} & \frac{\gamma\zeta}{\mu} & 0 \\ \hline 0 & 0 & \sigma-\frac{\alpha\zeta}{\mu} & 0 & \frac{\beta\chi}{\chi+\rho} \\ 0 & 0 & \sigma & \frac{-\alpha\zeta}{\mu} & 0 \\ 0 & 0 & 0 & f & -\delta \end{array}\right) \tag{A.2.12}$$

This is an upper triangular matrix, so we can examine the eigenvalues of the blocks down the main diagonal. Examining the first submatrix, we have eigenvalue $\lambda_1^1 = -\chi - \rho$. The eigenvalue of the second submatrix is clearly $\lambda_1^2 = -\mu$. We can see that these parts of the system are stable as the eigenvalues corresponding to these blocks are both negative. To consider the eigenvalues of the third submatrix, we examine roots of the characteristic equation.

$$0 = (\delta - \lambda)\left[\left(-\sigma - \frac{\alpha\zeta}{\mu} - \lambda\right)\left(-\frac{\alpha\zeta}{\mu} - \lambda\right)\right] + \frac{\beta\chi f\sigma}{(\rho+\chi)} \tag{A.2.13a}$$

$$0 = \lambda^3 + \lambda^2\left(\frac{\sigma\mu + 2\alpha\zeta}{\mu} + \delta\right) + \lambda\left[\left(\frac{\sigma\mu+\alpha\zeta}{\mu}\right)\left(\frac{\alpha\zeta}{\mu}\right) + \left(\frac{\delta(\sigma\mu+2\alpha\zeta)}{\mu}\right)\right] + \frac{\delta\alpha\zeta(\sigma\mu+\alpha\zeta)}{\mu^2} - \frac{\beta\chi f\sigma}{(\rho+\chi)} \tag{A.2.13b}$$

According to the Routh-Hurwitz criteria for characteristic polynomials of degree three

$$p(x) = a_3x^3 + a_2x^2 + a_1x^1 + a_0, \tag{A.2.14}$$

in order for the system to be stable, we require that the coefficients of the polynomial satisfy the following conditions, $a_i > 0$, for all $i \in \{0, 1, 2, 3\}$ and that $a_2a_1 > a_3a_0$ [27, 44]. We can see that the first three coefficients are clearly positive. Therefore, the

first condition for this to be stable requires that $a_0 > 0$

$$\frac{\delta\alpha\zeta(\sigma\mu + \alpha\zeta)}{\mu^2} - \frac{\beta\chi f\sigma}{(\rho+\chi)} > 0 \tag{A.2.15a}$$

$$\frac{\delta\alpha\zeta(\sigma\mu + \alpha\zeta)}{\mu^2} > \frac{\beta\chi f\sigma}{(\rho+\chi)} \tag{A.2.15b}$$

$$1 > \frac{\beta\chi f\sigma\mu^2}{\delta\alpha\zeta(\sigma\mu + \alpha\zeta)(\rho+\chi)} \tag{A.2.15c}$$

$$1 > \mathcal{R}_0 \tag{A.2.15d}$$

Thus we see that the healthy equilibrium is clearly unstable (as the first condition fails) when $\mathcal{R}_0 > 1$. We must confirm the second condition $a_1a_2 > a_0a_3$ for $\mathcal{R}_0 < 1$. We will first simplify and expand the right hand side of the inequality.

$$a_0a_3 = \frac{\delta\alpha(\sigma\mu + \alpha\zeta)}{\mu^2} - \frac{\beta f\chi\sigma}{(\rho+\chi)} \tag{A.2.16a}$$

$$= \frac{\delta\alpha(\sigma\mu + \alpha\zeta)(\rho+\chi) - \beta f\chi\sigma\mu^2}{\mu^2(\rho+\chi)} \tag{A.2.16b}$$

$$= \frac{\delta\alpha(\sigma\mu + \alpha\zeta)(\rho+\chi)(1-\mathcal{R}_0)}{\mu^2(\rho+\chi)} \tag{A.2.16c}$$

$$= \frac{\delta\alpha(\sigma\mu + \alpha\zeta)(1-\mathcal{R}_0}{\mu^2} \tag{A.2.16d}$$

$$= \frac{(1-\mathcal{R}_0)(\alpha\zeta)^2\delta}{\mu^2} + \frac{(1-\mathcal{R}_0)(\alpha\zeta)\delta\sigma}{\mu} \tag{A.2.16e}$$

Similarly, we will expand the left hand term a_1a_2 and then show that $a_1a_2 > a_0a_3$.

$$a_1a_2 = \left(\frac{(\sigma+\mu)(\alpha\zeta)}{\mu^2} + \frac{\delta(\sigma\mu+2\alpha\zeta)}{\mu}\right)\left(\frac{\sigma\mu+2\alpha\zeta}{\mu}+\delta\right) \tag{A.2.17a}$$

$$= \frac{2(\alpha\zeta)^3}{\mu^3} + \frac{5(\alpha\zeta)^2\delta}{\mu^2} + \frac{3(\alpha\zeta)^2\sigma}{\mu^2} + \frac{5(\alpha\zeta)\delta\sigma}{\mu} + \frac{(\alpha\zeta)\sigma^2}{\mu} + \frac{2(\alpha\zeta)\delta^2}{\mu} + \delta\sigma^2 + \delta^2\sigma \tag{A.2.17b}$$

$$> \frac{(1-\mathcal{R}_0)(\alpha\zeta)^2\delta}{\mu^2} + \frac{(1-\mathcal{R}_0)(\alpha\zeta\delta)\sigma}{\mu} \tag{A.2.17c}$$

$$= a_0a_3 \tag{A.2.17d}$$

as $(1-\mathcal{R}_0) < 5$. Finally, we have shown using the Routh-Hurwitz criterion that all the eigenvalues of the third submatrix are negative if and only if $\mathcal{R}_0 < 1$. In conclusion, we have shown that the healthy equilibrium $HE = \left(\frac{\rho}{\chi+\rho}, \frac{\chi}{\chi+\rho}, 0, 0, 0, \frac{\zeta}{\mu}\right)$ is stable for when $\mathcal{R}_0 < 1$ and unstable when $\mathcal{R}_0 > 1$.

A.3 Stability of the Disease Equilibrium, *DE*

First, recall that the solution to the disease equilibrium is

$$\overline{H} = \frac{\overline{Z}^2\alpha\mu + \overline{Z}(\gamma\rho - \alpha\zeta - \mu\rho) + \rho\zeta}{(\chi+\rho)\gamma\overline{Z}} \tag{A.3.18a}$$

$$\overline{S} = \frac{-\overline{Z}^2\alpha\mu + \overline{Z}(-\chi\mu + \chi\gamma + \alpha\zeta) + \chi\zeta}{(\chi+\rho)\gamma\overline{Z}} \tag{A.3.18b}$$

$$\overline{E} = \frac{\sigma(\mu\overline{Z} - \zeta)}{\gamma(\alpha\overline{Z} + \sigma)} \tag{A.3.18c}$$

$$\overline{I} = \frac{\sigma(\mu\overline{Z} - \zeta)}{\gamma\overline{Z}(\alpha\overline{Z} + \sigma)} \tag{A.3.18d}$$

$$\overline{V} = \frac{\overline{I}f}{\delta} = \frac{f\sigma(\mu\overline{Z} - \zeta)f}{\delta\gamma\overline{Z}(\alpha\overline{Z} + \sigma)} \tag{A.3.18e}$$

$$\tag{A.3.18f}$$

and the solution $\overline{Z}$ is the root to the polynomial

$$P(Z) = Z^3(\alpha^2\delta\gamma(\chi+\rho)) + Z^2(\alpha\sigma(\beta f\mu + (\chi+\rho)\delta\gamma)) + Z(\beta f\sigma(\chi(\mu-\gamma) - \alpha\zeta)) - \beta\chi f\sigma\zeta. \tag{A.3.19}$$

We analyze the nature of the polynomial to show that the diseased equilibrium is positive only when $\mathcal{R}_0 > 1$.

Lemma A.3.1. *The polynomial $P(z)$ has exactly one positive real root.*

Proof. We will first start by showing that the polynomial (A.3.19) has at least one positive real root. The polynomial $P(Z)$ is of the form:

$$P(Z) = aZ^3 + bZ^2 + cZ + d,\ Z \in \mathbb{R}, \tag{A.3.20}$$

where the coefficients are equal to:

$$a = \alpha^2\delta\gamma(\chi + \rho), \tag{A.3.21a}$$

$$b = \alpha\sigma(\beta f\mu + (\chi + \rho)\delta\gamma), \tag{A.3.21b}$$

$$c = \beta f\sigma(\chi(\mu - \gamma) - \alpha\zeta), \tag{A.3.21c}$$

$$d = -\beta\chi f\sigma\zeta, \tag{A.3.21d}$$

To show that we have *exactly* one positive root we will apply Descartes' Rule of Signs [12]. The theorem states that the number of positive real roots of a polynomial (ordered in decreasing order of the degree) is equal to either the number of times or less than that by some even number. For example, if a polynomial has coefficients which change sign three times, the number of positive real roots is either 3 or 1. Considering another example, if another polynomial has coefficients change sign four times, the number of positive real roots is either 4, 2, or 0. Examining the coefficients of the polynomial $P(Z)$, we see that $a, b > 0$, $d < 0$, and the coefficient c may be negative or positive depending on the values of the parameters. However, regardless of the value of c, we have only one change in the sign of the coefficients.

$$P(Z) = aZ^3 + \underbrace{bZ^2 + \overbrace{cZ}^{\text{switch, } c>0}}_{\text{switch, } c<0} + d \tag{A.3.22}$$

If $c < 0$, then we see a switch in the sign of the coefficients between b and c only. If $c > 0$, we see a switch in the sign of the coefficients between c and d only. As well, if $c = 0$, then we see a switch in sign between coefficients b and d only. Because we only have one sign change, we must have only one positive real root. □

Theorem A.3.1. *The diseased equilibrium DE is negative, and thus biologically irrelevant, for* $\mathcal{R}_0 < 1$.

Proof. In Lemma A.3.1 we showed that there was exactly one positive root for the polynomial $P(Z)$. This means that there is exactly one viable solution for the immune system at the diseased equilibrium, $\overline{Z}$, otherwise $\overline{Z} < 0$ and our solution would be biologically irrelevant. In order to show that the diseased equilibrium is not biologically relevant for $\mathcal{R}_0 < 1$, we must show that at least one of the equations in the system (A.3.18) is negative when $\mathcal{R}_0$. Looking at the system, we see that equations $\overline{E}, \overline{I}, \overline{V}$ are only negative when $(\mu\overline{Z} - \zeta) < 0$, that is $\overline{Z} < \dfrac{\zeta}{\mu}$, which is enough to show that the *DE* is biologically irrelevant. If $\overline{Z} > \dfrac{\zeta}{\mu}$, then we have that all of the solutions to the *DE* (A.3.18) are positive. In order to do this, we will actually show that $\overline{Z} < \dfrac{\zeta}{\mu}$ if and only if $\mathcal{R}_0 < 1$, which will be the result that we want.

We examine the sign of $P(\frac{\zeta}{\mu})$:

$$0 < P(\tfrac{\zeta}{\mu}) = \frac{\zeta^3\alpha^2\delta\gamma\chi}{\mu^3} + \frac{\zeta^3\alpha^2\delta\gamma\rho}{\mu^3} + \frac{\zeta^2\alpha\sigma\delta\gamma\chi}{\mu^2} + \frac{\zeta^2\alpha\sigma\delta\gamma\rho}{\mu^2} - \frac{\beta\zeta f\sigma\chi\gamma}{\mu} \tag{A.3.23a}$$

$$0 < \frac{\zeta\gamma}{\mu^3}\left(\zeta^2\alpha^2\delta\chi + \zeta^2\alpha^2\delta\rho + \zeta\alpha\sigma\delta\chi\mu + \zeta\alpha\sigma\delta\rho\mu - \beta f\sigma\chi\gamma\mu^2\right) \tag{A.3.23b}$$

$$0 < \zeta^2\alpha^2\delta\chi + \zeta^2\alpha^2\delta\rho + \zeta\alpha\sigma\delta\chi\mu + \zeta\alpha\sigma\delta\rho\mu - \beta f\sigma\chi\gamma\mu^2 \tag{A.3.23c}$$

$$\beta f\sigma\chi\gamma\mu^2 < \zeta^2\alpha^2\delta\chi + \zeta^2\alpha^2\delta\rho + \zeta\alpha\sigma\delta\chi\mu + \zeta\alpha\sigma\delta\rho\mu \tag{A.3.23d}$$

$$\beta f\sigma\chi\gamma\mu^2 < \delta(\alpha\zeta)(\alpha\zeta(\chi+\rho) + \sigma\mu(\chi+\rho)) \tag{A.3.23e}$$

$$\beta f\sigma\chi\gamma\mu^2 < \delta(\alpha\zeta)(\chi+\rho)(\alpha\zeta+\sigma\mu) \tag{A.3.23f}$$

$$\frac{\beta f\sigma\chi\gamma\mu^2}{\delta(\alpha\zeta)(\chi+\rho)(\alpha\zeta+\sigma\mu)} < 1 \tag{A.3.23g}$$

$$\mathcal{R}_0 < 1 \tag{A.3.23h}$$

Thus we have shown that $P\left(\frac{\zeta}{\mu}\right) > 0 \Leftrightarrow \mathcal{R}_0 < 1$. By continuity of $P(z)$ and because there is only one positive real root, this means that there exists a $\overline{Z} \in \left(0, \frac{\zeta}{\mu}\right)$ such that $P(\overline{Z}) = 0$, *i.e.,* the equilibrium $\overline{Z} < \frac{\zeta}{\mu}$ when $\mathcal{R}_0 < 1$ and $\overline{Z} > \frac{\zeta}{\mu}$ when $\mathcal{R}_0 > 1$ for the *DE*. Therefore, looking back to the *DE*, we can see that $\overline{E}, \overline{I}$, and $\overline{V}$ are all less than zero for when $\mathcal{R}_0 < 1$.

□

Appendix B

Examining Parameters in Deterministic Model

In this appendix we examine how the different parameter values affect the diseased equilibrium DE values. We plot the diseased equilibrium values for different values the various parameters, obtaining the equilibrium values after running the system for 2000 time units.

Firstly we examine the effects that α, the rate of clearance by the primed T-cells, has on the DE. We see that as α increases, the number of infected sites $\overline{E}, \overline{I}$ and the viral load $\overline{V}$ decrease because more of these infected sites are cleared by the immune response. Similarly, the immune response $\overline{Z}$ also decreases because fewer T-cells are required to clear the infection. This is illustrated in Figure B.0.1

When examining the effects that the rate of infection by virus β has on the DE, we see that $\overline{E}$ increases as β increases, because more sites are becoming exposed to infection. The number of infectious sites $\overline{I}$ and subsequently the viral load $\overline{V}$ increases initially, but then decreases slightly. This is because as the number of exposed sites

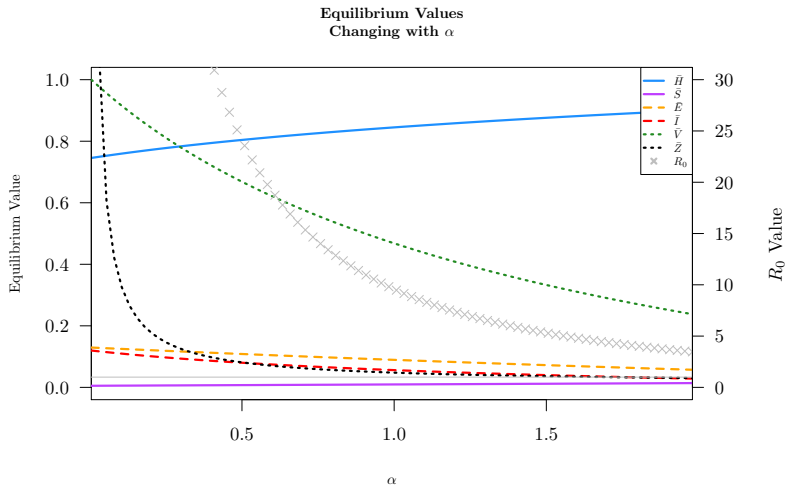

Figure B.0.1: The diseased equilibrium values as a function of α, the rate of clearance by T-cells. As α increases, it is seen that the infection equilibrium values decrease, while the number of healthy sites increases.

increases, so does the immune response $\overline{Z}$ and more exposed sites can be cleared before they mature to becoming infectious. This is illustrated in Figure B.0.2.

Examining the effects of the abrasion rate χ on the DE, we see that more sites can become infected after abrasion, increasing $\overline{E}$ as χ increases. A similar trend, the initial increase but then subsequent decrease in infectious sites $\overline{I}$ and virus $\overline{V}$, to the effects of β is observed here. However, it is more apparent. As the number of $\overline{E}$ sites increase due to more susceptible sites, the immune response $\overline{Z}$ also increases, which clears exposed sites before they can become infectious. This is illustrated in Figure B.0.3.

We also examine the effects of ρ the rate of recovery on the DE. We see that as ρ increases the number of available susceptible sites decrease, which also decreases the number of infected sites $\overline{E}, \overline{I}$ and the viral load $\overline{V}$. Subsequently, the lack of infection

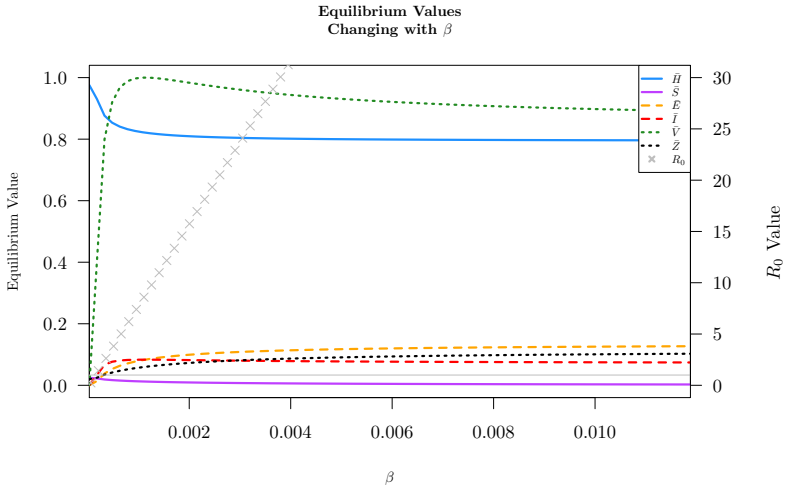

Figure B.0.2: The diseased equilibrium as a function of β, the rate of infection by a virus particle. As β increases then the infection equilibrium values increase along with the immune response, but then $\overline{V}$ and $\overline{I}$ decrease. This is because as more exposed sites are present, the immune response is increased, and these exposed sites can be cleared before maturing into infectious sites, depleting them for high values of β.

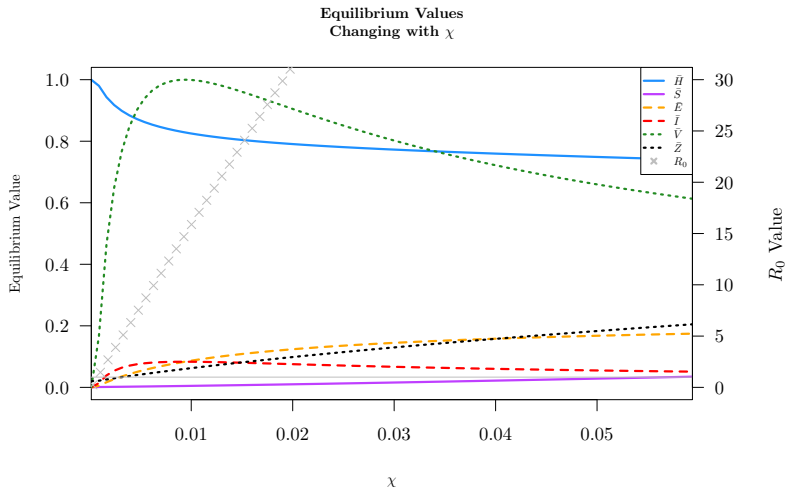

Figure B.0.3: This plot illustrates the equilibrium values of the diseased equilibrium as a function of χ. We see that as χ increases, $\overline{E}$ increases due to more available susceptible sites to infect. However, $\overline{I}$ and $\overline{V}$ increase then decrease as χ increases, most likely due to an increase in immune activity clearing infected I sites. The left axis is the proportion of sites and the right axis is the value of $\mathcal{R}_0$.

reduces the immune response $\overline{Z}$. This is illustrated in Figure B.0.4

We also consider how the rate of viral production f affects the solution DE. As f increases, the viral load increases $\overline{V}$, which results in more infection $\overline{E}, \overline{I}$ and subsequently more immune response to the higher levels of infection $\overline{Z}$. These findings are recorded in Figure B.0.5

Conversely, we examine how the rate of viral decay δ affects the DE. As δ increases, there are fewer virus particles $\overline{V}$, and thus less infection $\overline{E}, \overline{I}$, which results in a smaller immune response $\overline{Z}$. This is illustrated in Figure B.0.6.

The parameter σ, the rate of maturation of exposed sites into infectious sites, has some pretty interesting effects on the DE solution. As σ increases, more exposed sites transition to infected sites more quickly. This increases, $\overline{I}$ and $\overline{V}$, which results in

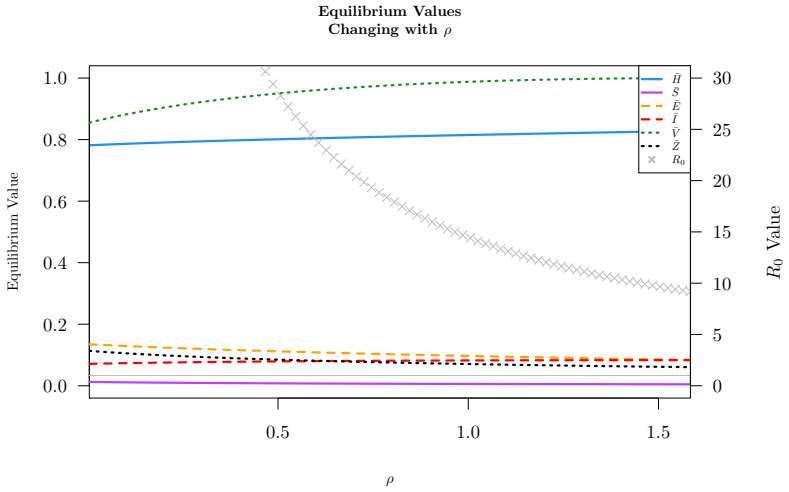

Figure B.0.4: The diseased equilibrium values as a function of ρ, the rate of cell recovery. As ρ increases the number of susceptible sites decreases, which subsequently decreases the number of infected sites. Healthy sites increase.

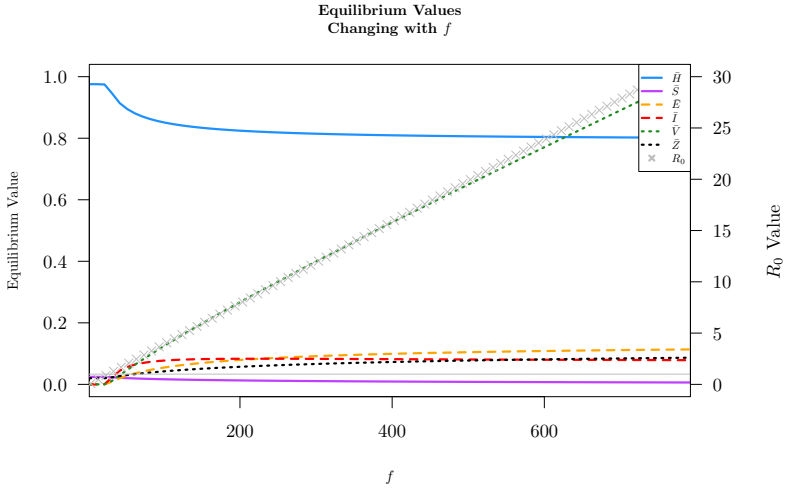

Figure B.0.5: The diseased equilibrium values as a function of f, the rate of viral production. As f increases, then the number of infected sites increases, and the healthy sites decrease. In response to more infected sites, the immune response also increases.

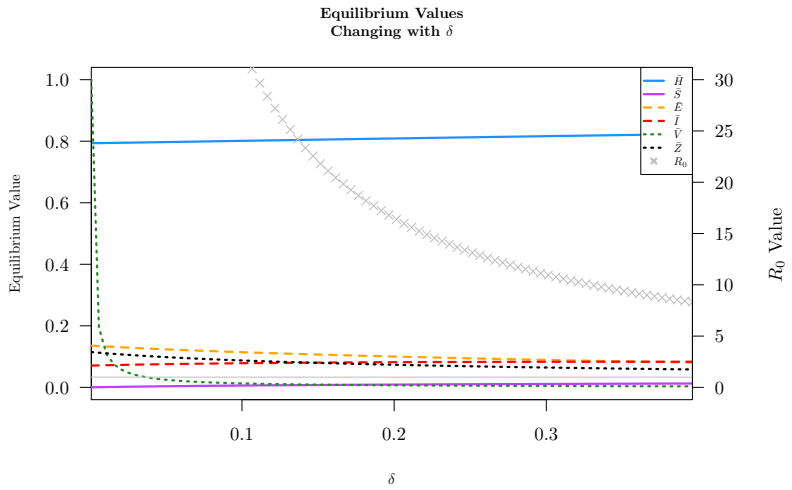

Figure B.0.6: The diseased equilibrium values as a function of δ, the rate of natural viral particle death. As δ increases the number of viral particles decreases significantly, which reduces the number of infected sites and subsequently the immune response.

more infections initially. However, as σ increases larger, then we see fewer exposed sites $\overline{E}$ because they are either transitioning to become infectious or are cleared by the overall increased immune response $\overline{Z}$. This is illustrated in Figure B.0.7.

While γ doesn't have an effect on the value of $\mathcal{R}_0$, it does have an important role on the value of the DE. As γ, the rate of propagation of primed T-cells, increases, there are more T-cells present in the system $\overline{Z}$, which decreases the overall infection $\overline{E}, \overline{I}$ and the viral load $\overline{V}$. Illustrated in Figure B.0.8.

The base immune activity rate ζ also has an overall affect on the diseased equilibrium value DE. As ζ increases, so does the overall immune response $\overline{Z}$, which decreases the overall infection $\overline{E}, \overline{I}$, and $\overline{V}$. As seen in Figure B.0.9.

Lastly, we explore the effects of the T-cell death rate μ on the diseased equilibrium value DE. As μ increases, the effective immune response $\overline{Z}$ decreases, which allows

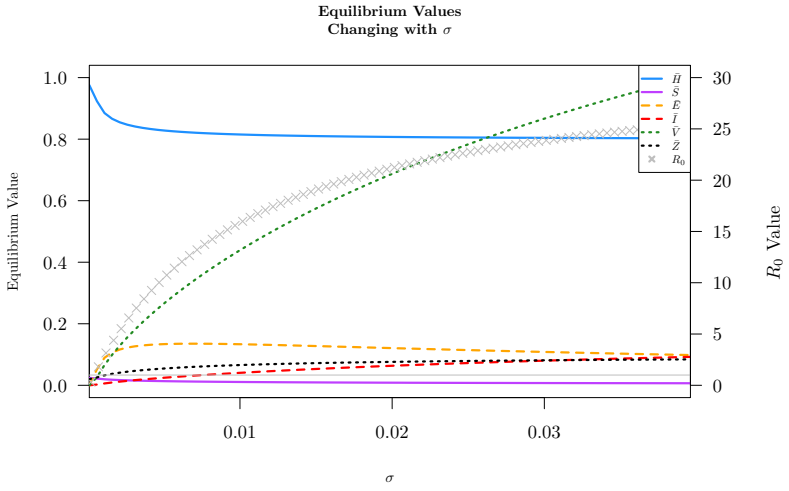

Figure B.0.7: The diseased equilibrium values as a function of σ, the rate of maturation of exposed sites to infectious. As σ increases more exposed sites are able to mature to infectious, which increases the viral load in the system, thus increasing the overall number of infected sites. The immune response is also increased due to the increase in infected sites.

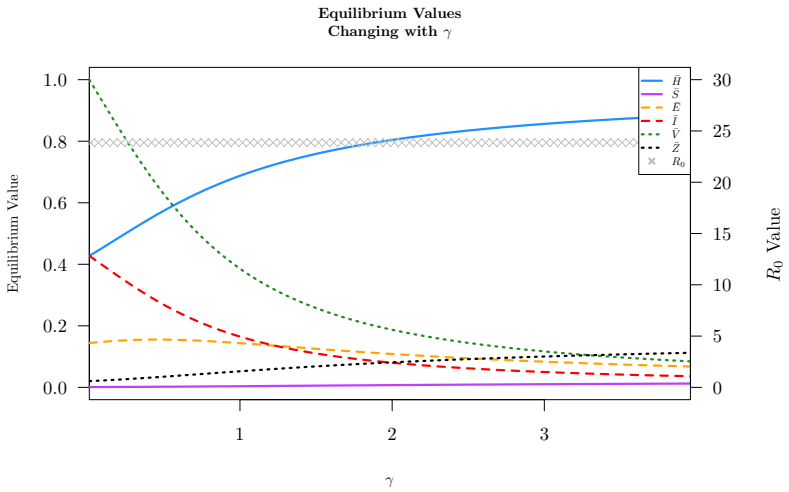

Figure B.0.8: This plot shows the diseased equilibrium values as a function of γ. It can be seen that the values of $\mathcal{R}_0$ (gray x's) stay constant for all γ but the values for $\overline{E}, \overline{I}$, and $\overline{V}$ all decrease as γ increases. The left axis is the proportion of sites and the right axis is the value of $\mathcal{R}_0$.

for the infection to establish and increase $\overline{E}, \overline{I}$, resulting in a higher viral load $\overline{V}$. This is observed in Figure B.0.10.

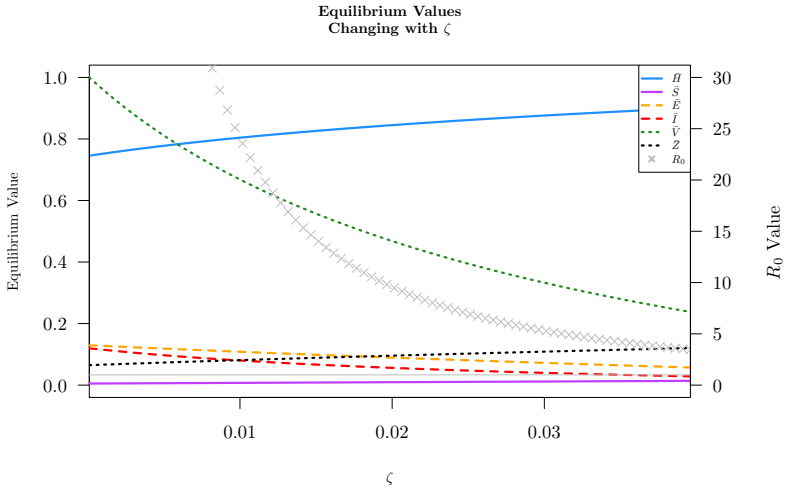

Figure B.0.9: The diseased equilibrium value as a function of ζ, the base immune response not in the presence of infection. As ζ increases, then the base number immune response increases, which decreases the number of infected sites, and the healthy sites thus increase.

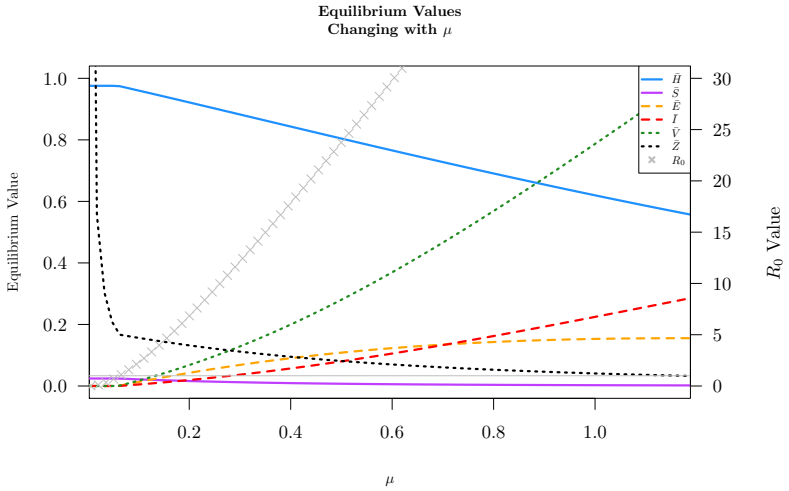

Figure B.0.10: The diseased equilibrium value as a function of μ, the natural death rate of immune cells. As μ increases immune cells die more often, thus the immune response decreases, which increases the number of infected sites. The number of healthy sites decrease.

Appendix C

Considering Super-Infections in a Patch Model

The effects of super-infection has been an interesting question for virologists, epidemiologists, and mathematical modellers for some time. In particular, some pathogens establish in a host and spread to other parts of the host via the release of viral particles. This is similar to plant-seed dispersal, and patch models have been used to model these viral processes [35]. However, when combining super-infection in a patch model can introduce asymmetry and bias to coinfected patches. This as been explored previously [1, 58], and here we outline this asymmetry and propose two techniques to avoid this asymmetry.

C.1 The Asymmetrical Case

First we will discuss the asymmetrical case that may be unintentionally implemented by modellers. We consider patches which are uninfected x, patches infected with only

one of each strain y_1 and y_2, and coinfected patches z. Coinfection occurs when one singly infected strain is infected with a different strain. Coinfection with both strains does not occur simultaneously. All infected strains are cleared at the same rate (we set it to 1 for simplicity), which become healthy after clearance. This is illustrated in the following flow diagram:

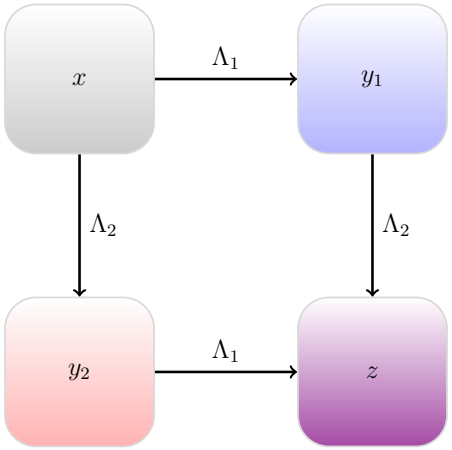

Figure C.1.1: Flow diagram for the simple patch model with super-infection.

This flow diagram is also expressed as a system of differential equations in System (C.1.1) below:

$$\frac{\mathrm{d}y_1}{\mathrm{d}t} = \Lambda_1 x - y_1 - \Lambda_2 y_1, \tag{C.1.1a}$$
$$\frac{\mathrm{d}y_2}{\mathrm{d}t} = \Lambda_2 x - y_2 - \Lambda_1 y_2, \tag{C.1.1b}$$
$$\frac{\mathrm{d}z}{\mathrm{d}t} = \Lambda_1 y_2 + \Lambda_2 y_1 - z, \tag{C.1.1c}$$
$$1 = x + y_1 + y_2 + z, \tag{C.1.1d}$$

where $\Lambda_1 = \beta_1(y_1 + p_1 z)$ and $\Lambda_2 = \beta_2(y_2 + p_2 z)$ are the forces of infection, and p_1 and p_2 are the proportions of cell resources allocated to each of strains 1 and 2, respectively,

in a coinfected patch.

We consider the boundary equilibria where only one strain y_1 is present and the other strain and the coinfected patches are not, $(\bar{x}, \bar{y_1}, \bar{y_2}, \bar{z}) = (\bar{x}, \bar{y_1}, 0, 0)$. We solve for $\bar{x}$ and $\bar{y_1}$.

$$\bar{x} = \frac{1}{\beta_1} \tag{C.1.2}$$

$$\bar{y_1} = \frac{\beta_1 - 1}{\beta_1} \tag{C.1.3}$$

Notice here, that for a valid equilibrium, we require that $\beta_1 > 1$, and thus we also impose that $\beta_2 > 1$. We also solve for the forces of infection at equilibria:

$$\overline{\Lambda_1} = \beta_1 - 1 \tag{C.1.4}$$

$$\overline{\Lambda_2} = 0 \tag{C.1.5}$$

Now consider the average number of new strain two or coinfected patches given a single patch infected with strain two, y_2. This is the invasion reproduction number, $\overline{\mathcal{R}}$. To determine this quantity, we must consider how new sites become infected with strain two. Firstly, a completely susceptible site $\bar{x}$ can become infected by either y_2 or a coinfected site that was first infected with y_1. Note that in the later case, we must have that the site initially infected with y_1 survives long enough to be coinfected. The number of susceptible x patches for infection is $\bar{x}$.

$$\bar{x} \to y_2 : \frac{\beta_2}{\beta_1(1 + \overline{\Lambda_1})} + \frac{p_2\beta_2\overline{\Lambda_1}}{\beta_1(1 + \overline{\Lambda_1})} = \frac{\beta_2}{\beta_1}\left(\frac{1 + p_2\overline{\Lambda_1}}{1 + \overline{\Lambda_1}}\right) \tag{C.1.6}$$

Similarly, we may have that a patch singly infected with strain 1 may become coinfected either by y_1 or by a coinfected site z. Recall that the coinfected sites z come from the new strain y_2 first being infected by y_1. The number of susceptible y_1 patches for infection is $\bar{y_1}$.

$$\bar{y_1} \to z : \frac{\beta_2(\beta_1 - 1)}{\beta_1(1 + \overline{\Lambda_1})} + \frac{p_2\beta_2(\beta_1 - 1)\overline{\Lambda_1}}{\beta_1(1 + \overline{\Lambda_1})} = \frac{\beta_2}{\beta_1}\left(\frac{(\beta_1 - 1) + (\beta_1 - 1)p_2\overline{\Lambda_1}}{1 + \overline{\Lambda_1}}\right) \tag{C.1.7}$$

We can sum these two cases and obtain an expression for $\overline{\mathcal{R}}$.

$$\begin{aligned}
\overline{\mathcal{R}} &= \frac{\beta_2}{\beta_1}\left(\frac{1 + p_2\overline{\Lambda_1} + (\beta_1 - 1) + (\beta_1 - 1)p_2\overline{\Lambda_1}}{1 + \overline{\Lambda_1}}\right) && \text{(C.1.8)}\\
&= \frac{\beta_2}{\beta_1}\left(\frac{\beta_1 + p_2\beta_1\overline{\Lambda_1}}{\beta_1}\right) && \text{(C.1.9)}\\
&= \frac{\beta_2}{\beta_1}(1 + p_2\overline{\Lambda_1}) && \text{(C.1.10)}\\
&= \frac{\beta_2}{\beta_1}(1 + p_2(\beta_1 - 1)) && \text{(C.1.11)}\\
& && \text{(C.1.12)}
\end{aligned}$$

Here we see that $1+p_2(\beta_1-1) > 1$, and thus $\overline{\mathcal{R}} > 1$ even when $\beta_1 > \beta_2$. This conflicts with the Competitive Exclusion Principle, in that two organisms cannot occupy the same niche, and the stronger one will win out. In this case, we may have that $\beta_1 > \beta_2$, but strain one can still be invaded by coinfected strains. This is an example of the asymmetry that can be caused when introducing superinfection into patch models.

Setting the value $p_1 = p_2 = 0.5$ and the transmission coefficients $\beta_1 = 3$ and $\beta_2 = 1.45$, we obtain the following solution curve solved using R. This is illustrated in figure C.1.2, where we see coexistence of patches infected with type y_1 and those which are coinfected z. Keeping the p_1, p_2 values the same ($p_1 = p_2 = 0.5$) and

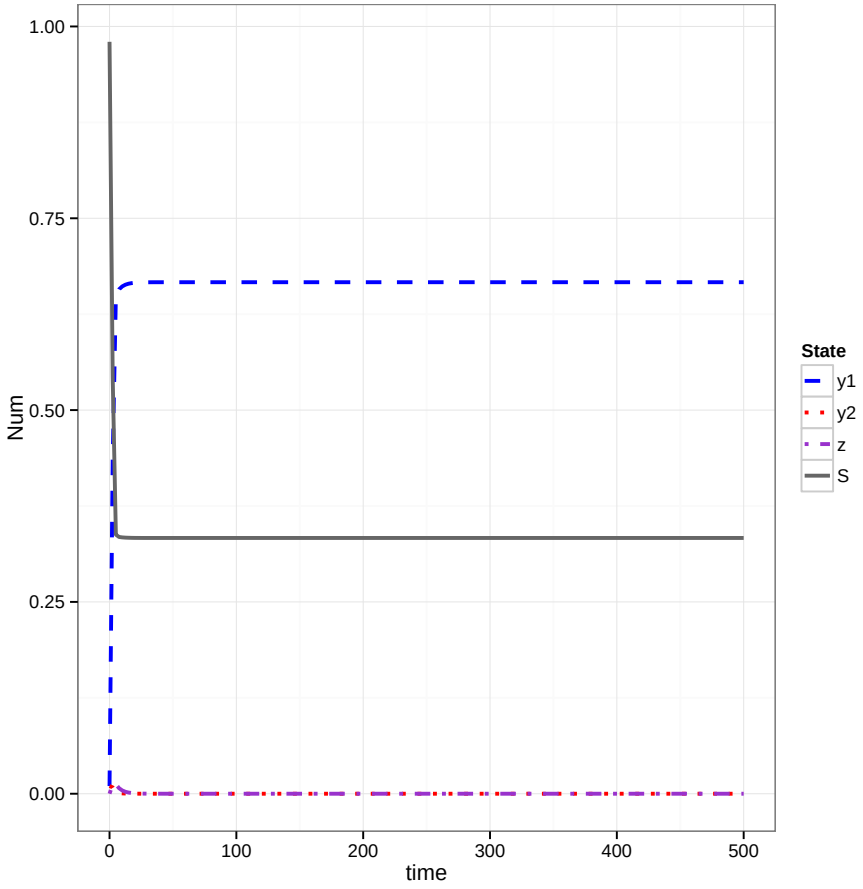

Figure C.1.2: This shows the case where the second pathogen, either as single infections y_2 or as super-infections z, is unable to invade. This occurs when $\overline{\mathcal{R}} < 1$. In this case, $\overline{\mathcal{R}} = 0.9\bar{6}$

changing $\beta_2 = 2.5$ we show that coexistence is possible, countering the Competitive Exclusion Principle. The system was numerically solved and the results can be shown in Figure C.1.3

The asymmetry introduced here is caused because of the exclusion of subsequent super-infection after super-infection by opposing types has occurred. Consider two patches, each initially infected with strain 1. The first patch, can be infected by

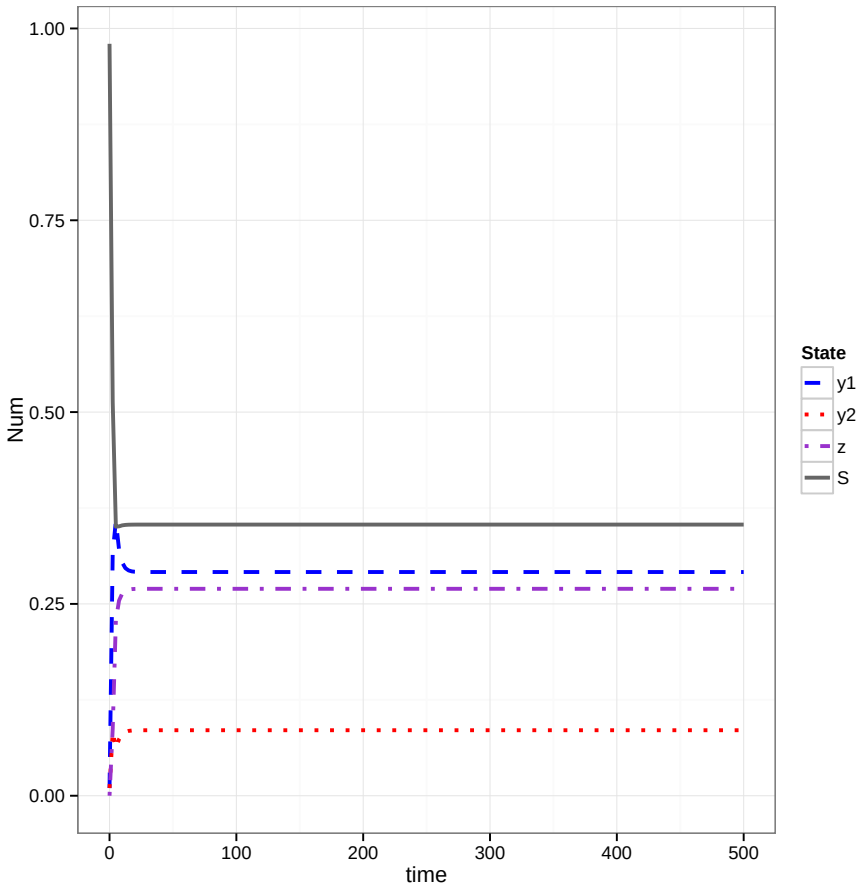

Figure C.1.3: This plot exemplifies the asymmetry that is introduced into the system. Even though the two strains occupy the same niche, and that strain 1 is stronger than strain 2, coexistence is possible because of the advantage of super-infected sites, z. Here $\bar{\mathcal{R}} = 1.6\bar{6}$.

strain 2, and that patch is now immune to subsequent super-infection. Now consider the second patch infected with strain 1, if it were come into contact with a virus of strain 1, we would just consider it to be singly infected with strain 1 and thus a super-infection event would not have occurred. However, it can still be infected by strain 2. Therefore, in this way, patches may come into contact with a different number of virus. In particular patches that come into contact with opposing types

become immune to subsequent super-infection, while patches may come into contact with viruses of the same strain before being super-infected by an opposing strain. In this way, a site which is coinfected by two different strains becomes advantageous over a site that is infected by the same strains. This is the asymmetry that was introduced into the previous model and is illustrated as a diagram in Figure C.1.4. There are a number of ways to include super-infection that does not introduce this form of asymmetry, and it is discussed further in Section C.2

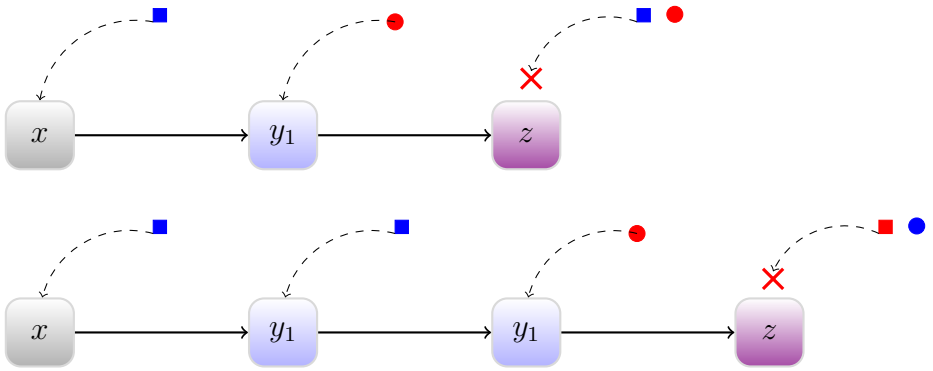

Figure C.1.4: This illustrates the asymmetry that can form in patch models. If we do not account for super-infection and subsequent recalcitrance of patches super-infected with the same strain, we can introduce asymmetry. Consider the top case, where a susceptible patch x comes into successful contact with virus of strain 1 (red square), and becomes infected with strain 1. It then comes into contact with a virus of strain 2 (blue circle), and becomes super-infected with two different strains. After this, it becomes recalcitrant to subsequent super-infection by any strain. In the bottom case, after initial infection with strain 1, the patch comes into contact with a virus of strain 1 again. However, because it does not become recalcitrant to super-infection after this second encounter, the patch can later be super-infected by a virus of strain 2. In this way, the bottom patch was essentially infected three times while the top one only twice. This confers a form of immunity to patches super-infected by two different types, which imposes an artificial advantage.

C.2 Symmetrical Super-Infection Scenarios

We have just highlighted how the initial implementation of super-infection can result in asymmetry in the system. Below we discuss some ways to implement super-infection that does not lead to they asymmetrical effect. We run simulations for each of the different scenarios and compare them to one another.

1. One may allow super-infection of 1 by 1 (and 2 by 2), and make these super-infected sites recalcitrant, that is unable to be subsequently coinfected. We call this the Same Strain Super-Infection Model.

2. The other method (that may work) is to allow for a patch to have k number of super-infections, and subsequent ones replace earlier ones. In the case of two strains, an obvious choice is $k = 2$. So if a strain is infected with type 1, then type 2, then type 2 again, the type 1 is replaced with the newest incoming strain 2 and this patch is overall producing virus of strain 2. We refer to this as the First-in-First-out Super-Infection Model.

C.2.1 Same Strain Super-Infection Model

We start by examining what happens when we allow same strain super-infection and impose immunity to subsequent super-infected patches. We can do this by including two new classes, z_1 and z_2 which are patches that have been super-infected by the same strains, 1 and 2 respectively. Here we have that all z_i for $i = 1, 2$ compartments are immune to subsequent super-infection. This system can be visualized as a flow diagram in Figure C.2.5.

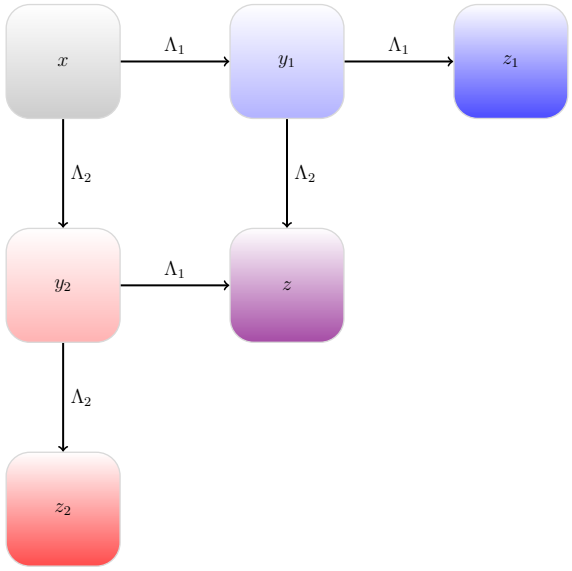

Figure C.2.5: Flow diagram for the super-infection patch model that considers same strain super-infection and subsequent immunity to further super-infection.

It is written as a system of differential equations (C.2.13).

$$\frac{\mathrm{d}y_1}{\mathrm{d}t} = \Lambda_1 x - \Lambda_1 y_1 - \Lambda_2 y_1 - y_1, \tag{C.2.13a}$$
$$\frac{\mathrm{d}y_2}{\mathrm{d}t} = \Lambda_2 x - \Lambda_2 y_2 - \Lambda_1 y_2 - y_2, \tag{C.2.13b}$$
$$\frac{\mathrm{d}z_1}{\mathrm{d}t} = \Lambda_1 y_1 - z_1, \tag{C.2.13c}$$
$$\frac{\mathrm{d}z_2}{\mathrm{d}t} = \Lambda_2 y_2 - z_2, \tag{C.2.13d}$$
$$\frac{\mathrm{d}z}{\mathrm{d}t} = \Lambda_1 y_2 + \Lambda_2 y_1 - z, \tag{C.2.13e}$$
$$1 = x + y_1 + y_2 + z_1 + z_2 + z, \tag{C.2.13f}$$

where $\Lambda_1 = \beta_1(y_1 + p_1 z + z_1)$ and $\Lambda_2 = \beta_2(y_2 + p_2 z + z_2)$ are the forces of transmission for strains 1 and 2, respectively. We assume that a patch that has been "super-infected" by the same strain produces the same amount of virus as a singly infected

patch.

We can visualize the Competitive Exclusion Principle if we have $\beta_2 < \beta_1$ (Figure C.2.6) or if $p_2 < p_1$ (Figure C.2.7), in which case the second strain is weaker, and is eventually eradicated from the system. We see coexistence in this scenario only when both $\beta_1 = \beta_2$ and $p_1 = p_2$.

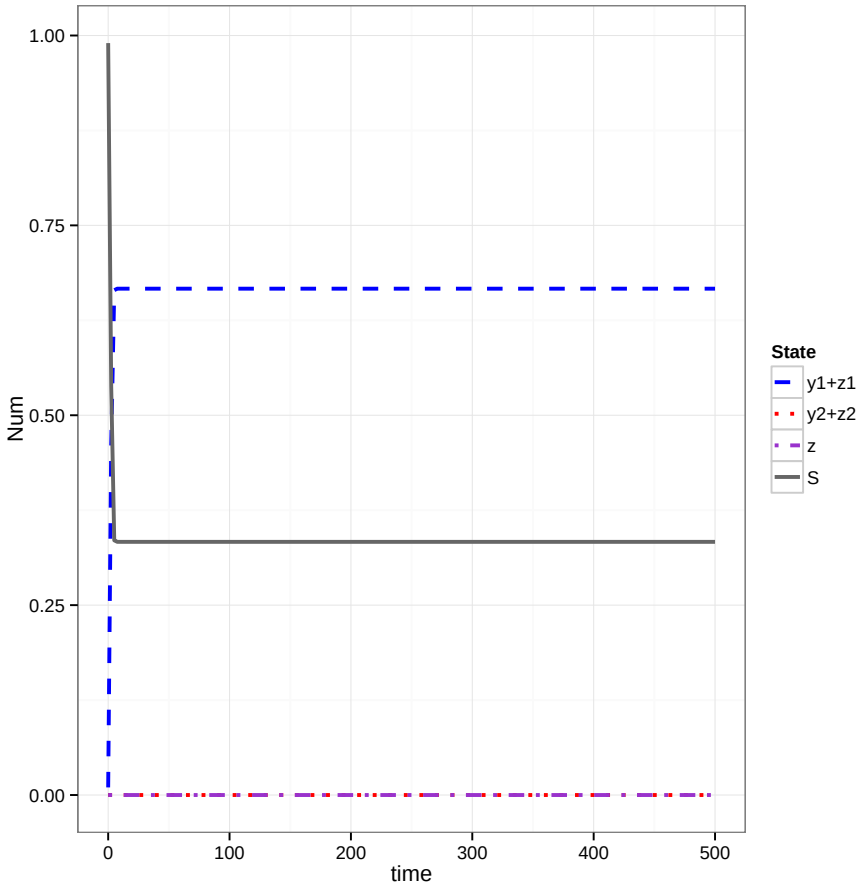

Figure C.2.6: Here we see when we allow for super-infection of the same type, followed by immunity to subsequent super-infection, we obtain symmetry in the system. We plot $y_1 + z_1$ (as they are producing the same amount of virus, blue:dashed), $y_2 + z_2$ (red:dotted), and z (purple:dash-dot). When $3 = \beta_1 > \beta_2 = 2.5$, we see that strain 1 dominates and strain 2 and the super-infected patches die out.

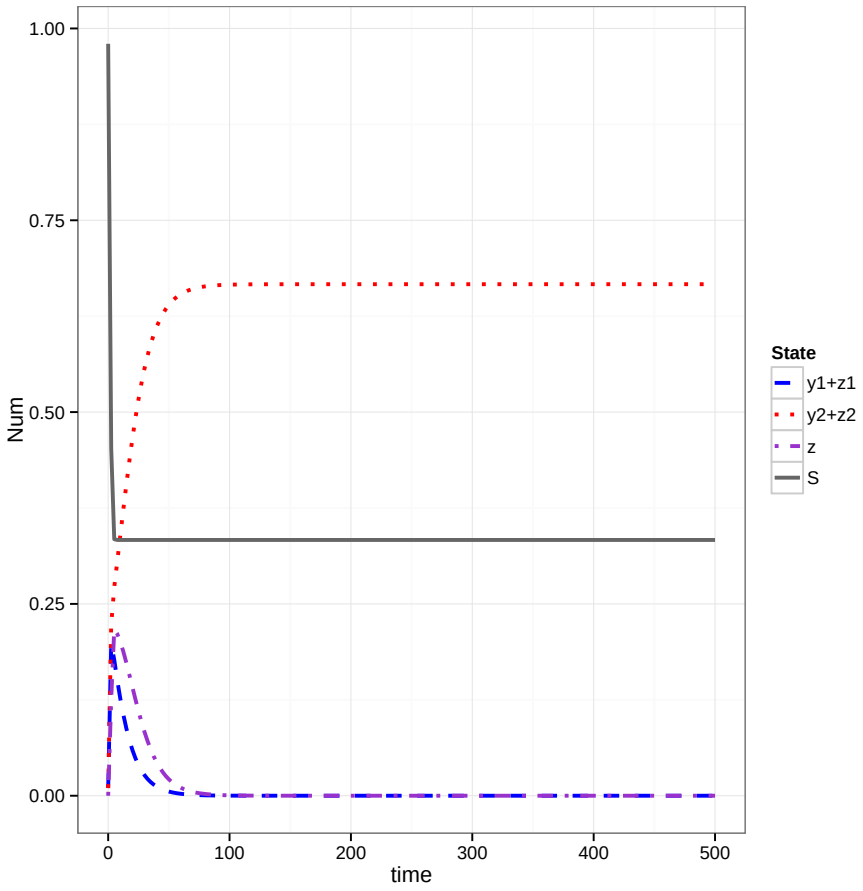

Figure C.2.7: Here we see when we allow for super-infection of the same type, followed by immunity to subsequent super-infection, we obtain symmetry in the system. We plot $y_1 + z_1$ (as they are producing the same amount of virus, blue:dashed), $y_2 + z_2$ (red:dotted), and z (purple:dash-dot). When $0.45 = p_1 < p_2 = 0.55$, we see that strain 2 dominates and strain 1 and the super-infected patches die out.

C.2.2 First-in-first-out Super-Infection Patch Model

Another method for imposing symmetry in a super-infection patch model is to allow for continual superinfection. That is, no patches become recalcitrant after some number of super-infections. Instead we patches can only "maintain" up to two strains at any one time, accounting for super-infection of the same strain twice also. We

impose a first-in-first-out (FIFO) method of superinfection. That is, if a patch is infected first with strain 1, then super-infected with strain 2, and then again super-infected with strain 2, the strain 1 is "kicked-out", and it becomes 2-2 infected. This is illustrated in figure Figure C.2.8.

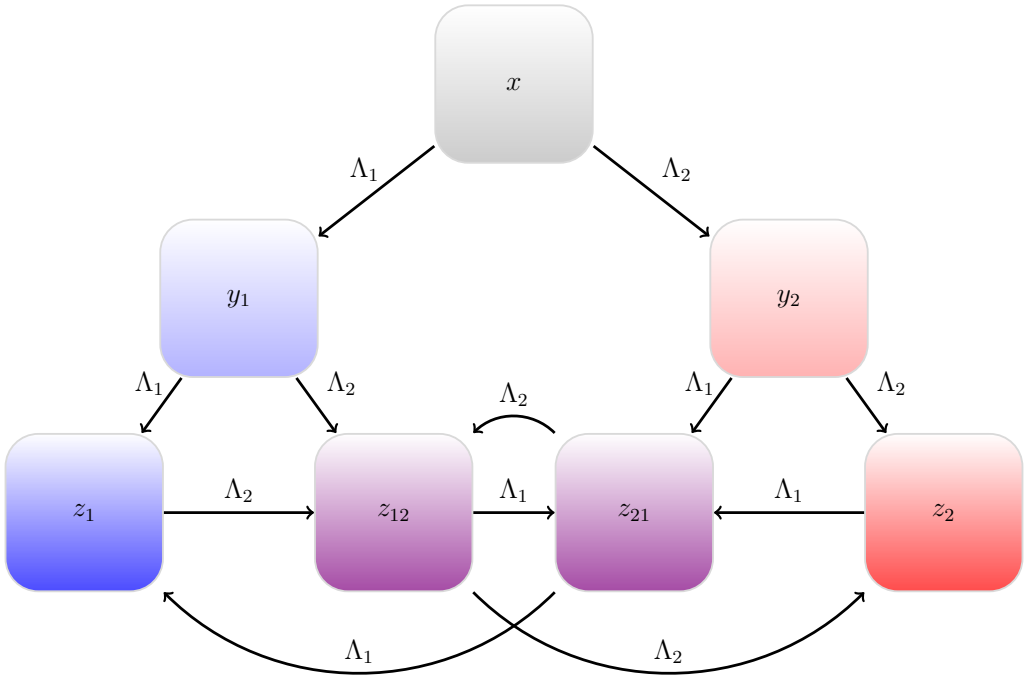

Figure C.2.8: A flow diagram illustrating the infection process and super-infection process of the First-in, First-out super-infection patch model. Susceptible patches x can be initially infected by one of two strains into y_1 or y_2. They may be super-infected by the same strain to become z_1 or z_2, or they may be super-infected by different strains to become z_{12} or z_{21} respectively. The order of the number refers to the order of infection. Then these super-infected sites can move between one another based on what strain they are infected by, via a first-in-first-out method.

We can also represent this as a system of differential equations. We set the collection of susceptible patches as x, which can then be infected by strain 1 or strain 2 to patches y_1 or y_2, respectively. After the initial infection, they can be super-infected

by the same strain and move to compartment z_1 or z_2, respectively, or they may be infected by the opposite strain and move to compartments z_{12} and z_{21}, respectively. Here, we set the order of the subscripts to mean the order of infection, *e.g.,* z_{12} means it was first infected by strain 1 then strain 2. These super-infected patches can then be infected by any strain of virus and they move to the corresponding compartment in a first-in-first-out method. For example, consider a patch is in the z_{12} class. It can be infected by strain 1, it would "kick out" the first strain 1, and replace it with strain 1, moving it to the z_{21} class. There is no qualitative difference between z_{12} and z_{21} apart from the order of infection. If the z_{12} patch is infected with strain 2 again, it will move to the z_{22} class, and be solely infected with strain 2. All infected patches can be cleared, and become subsequently susceptible again at a constant rate, which we set to 1 (this flow is not included in the diagram in Figure C.2.8). We represent this as a system of differential equations in System (C.2.14).

$$\frac{\mathrm{d}y_1}{\mathrm{d}t} = \Lambda_1 x - \Lambda_1 y_1 - \Lambda_2 y_1 - y_1, \tag{C.2.14a}$$
$$\frac{\mathrm{d}y_2}{\mathrm{d}t} = \Lambda_2 x - \Lambda_2 y_2 - \Lambda_1 y_2 - y_2, \tag{C.2.14b}$$
$$\frac{\mathrm{d}z_1}{\mathrm{d}t} = \Lambda_1 y_1 + \Lambda_1 z_{21} - \Lambda_2 z_1 - z_1, \tag{C.2.14c}$$
$$\frac{\mathrm{d}z_2}{\mathrm{d}t} = \Lambda_2 y_2 + \Lambda_2 z_{12} - \Lambda_1 z_2 - z_2, \tag{C.2.14d}$$
$$\frac{\mathrm{d}z_{12}}{\mathrm{d}t} = \Lambda_2 y_1 + \Lambda_2 z_1 - \Lambda_2 z_{12} - \Lambda_1 z_{12} + \Lambda_2 z_{21} - z_{12}, \tag{C.2.14e}$$
$$\frac{\mathrm{d}z_{21}}{\mathrm{d}t} = \Lambda_1 y_2 + \Lambda_1 z_2 - \Lambda_1 z_{21} + \Lambda_1 z_{12} - \Lambda_2 z_{21} - z_{21}, \tag{C.2.14f}$$
$$1 = x + y_1 + y_2 + z_1 + z_2 + z_{12} + z_{21}, \tag{C.2.14g}$$

This system gets quite difficult to analyze, so we illustrate the symmetry of this model by numerically solving this system in **R** using the `deSolve` package.

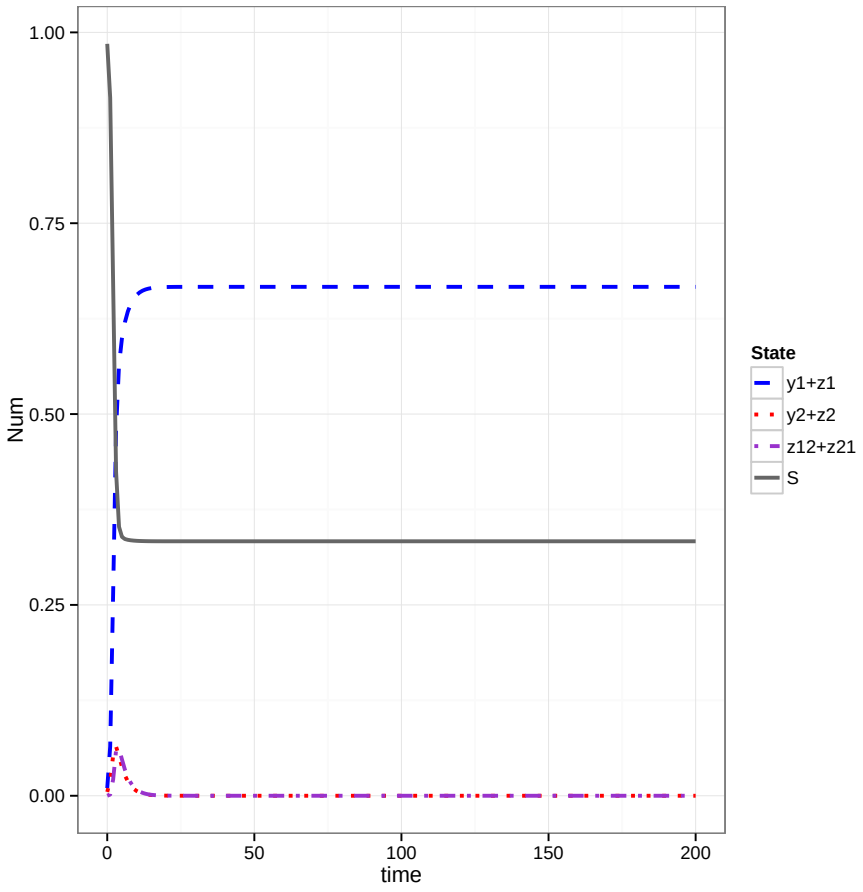

Figure C.2.9: When we allow for unlimited super-infections, including those of the same type, but recording only the most recent two super-infections (FIFO), we observe symmetry in the patch model. We plot $y_1 + z_1$ (as they are producing the same amount of virus, blue:dashed), $y_2 + z_2$ (red:dotted), and z (purple:dash-dot). When $3 = \beta_1 > \beta_2 = 2.5$, we see that strain 1 dominates and strain 2 and the super-infected patches die out.

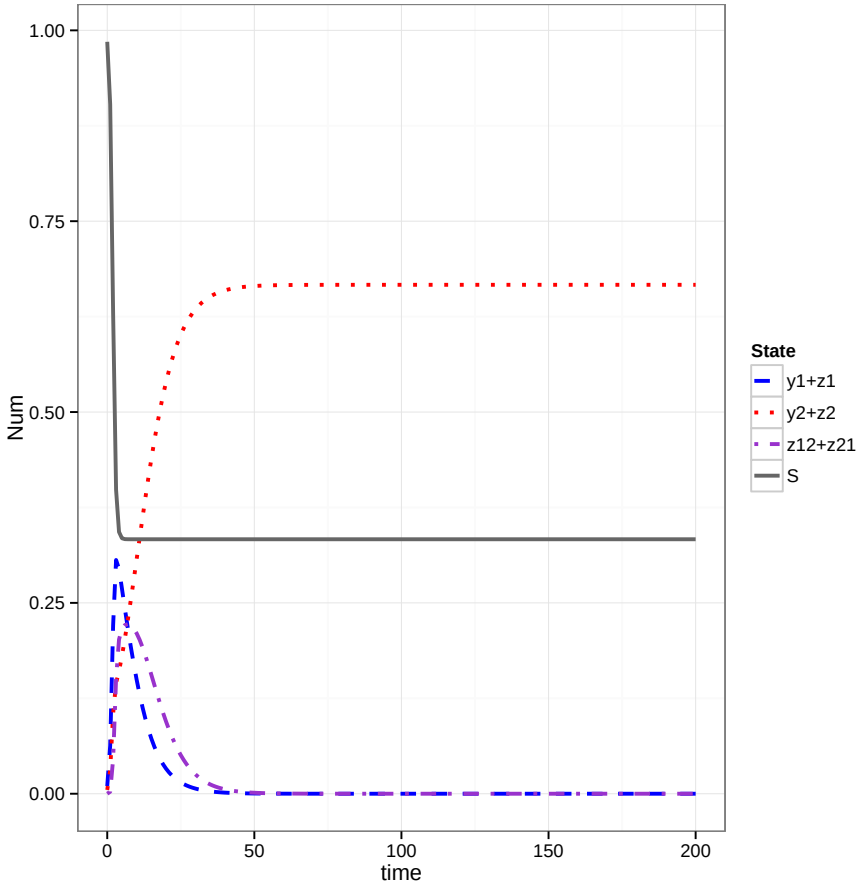

Figure C.2.10: When we allow for unlimited super-infections, including those of the same type, but recording only the most recent two super-infections (FIFO), we observe symmetry in the patch model. We plot $y_1 + z_1$ (as they are producing the same amount of virus, blue:dashed), $y_2 + z_2$ (red:dotted), and z (purple:dash-dot). When $0.45 = p_1 < p_2 = 0.55$, we see that strain 2 dominates and strain 1 and the super-infected patches die out.

C.3 Discussion

We presented two methods to avoid asymmetry in two-strain super-infection patch models. These techniques could be extrapolated to higher numbers of strains if desired. One method induces immunity to subsequent super-infection following the first

super-infection event. The other allows for an arbitrary number of super-infection events, but only keeping track of the two most recent super-infections. Depending on the assumptions of the models, one of these models may be suitable to use to avoid asymmetry. This asymmetry, unintentionally favours coexistence of both strains when the Competitive Exclusion Principle suggests otherwise. Ultimately, we hope to showcase how it is easy to unintentionally and accidentally introduce asymmetry in super-infection patch models. Thus, care should be taken when developing such models, as they may suggest coexistence between strains or species even when it is unlikely or impossible.

Appendix D

Numerical Solver and Spatial Simulation Code

All code was written in **R**. While **R** is predominately a statistical software it also has functionality for solving ordinary differential equations and running simulations while remaining non-proprietary.

D.1 Deterministic Model Solver Code for Within-Host HPV Models

The gradient functions for the system of ODEs for each model were written in **R**, and then the `ode()` function from the `deSolve` package [50] was used to numerically integrate the system. In order to implement the switch function used in the immune response delay model, we use the `event` and `root` functions of the `deSolve` package. Below are the complete gradient function, the event function, the root function, and

an example of solving the system.

```
vector.field <-
  function(t,y,parms) {
    with(as.list(c(parms,y)),{
      dB<-(-chi*B)+rho*S+alpha*(Z+M)*(E+I)/N
      dS<-chi*B-bet*V*S/N-rho*S
      dE<-bet*V*S/N-sigma*E-alpha*(Z+M)*E/N
      dI<-sigma*E-alpha*(Z+M)*I/N
      dV<- f*I-delta*V
      dZ<- switch(t,Tstart)*gam*Z*(I+E)/N-mu*Z+zeta*N
      dM<-switch(t,Tstart)*eps*Z*(1-M/(K*N))
      dCum<-dI+dE
      dTstart<-0
      res<-c(dB=dB,dS=dS,dE=dE,dI=dI,dV=dV,dZ=dZ,dM=dM,dCum=dCum,
          dTstart=dTstart)
      list(res)
    })
  }

switch<-function(t,thresh){
  x<-rep(-1,length(t))
  if(length(thresh)==2){
    off<-which(t<thresh[1]|t>thresh[2])
    on<-which(t>=thresh[1]&t<=thresh[2])
  }
  else{
    off<-which(t<thresh[1])
    on<-which(t>=thresh[1])
  }
```

```
  x[off]<-0
  x[on]<-1
  return(x)
}

#event function for switching: it sets the initial condition to be
    the same exept for
# t start which is restart to
event<-function(t,y,parms){
  B<-y[1]
  S<-y[2]
  E<-y[3]
  I<-y[4]
  V<-y[5]
  Z<-y[6]
  M<-y[7]
  Cum<-y[8]
  Tstart<-t
  return(y=c(B,S,E,I,V,Z,M,Cum,Tstart))
}



library(deSolve)

soln <- ode(
    y=c(B=B0,S=S0,E=E0,I=I0,V=V0,Z=Z0,M=M0,Cum=E0+I0,Tstart=Tstart),
    times=seq(from=0,to=tmax,by=tmax/interNum),
    func=vector.field,
    events=list(func=event,root=T),
```

```
    rootfun=function(t,y,parms){with(as.list(y,parms),Cum-C.thresh)
       },
    parms=c(chi=chi,rho=rho,sigma=sigma,
            bet=bet,f=f,mu=mu,gam=gam,
            alpha=alpha,zeta=zeta,C.thresh=C.thresh,
            delta=delta,K=K,N=N,eps=eps)
    )
```

D.2 Spatial Simulation Code

For the spatial simulation, we set up a pipeline using `Make` to run the **R** code. First we set up a parameter .R file to initialize all the necessary parameters for the simulation:

```
#parameters
alpha<-0.5
bet<-0.003
gam<-2
delta<-0.138
f<-600
chi<-0.015
rho<-0.6
sigma<-0.03
mu<-0.5
zeta<-0.01
eps<-0
Km<-0.01

count<-1
```

```
1589 16
1590 17 rowSize<-20
1591 18 colSize<-20
1592 19 organSize<-rowSize*colSize
1593 20 sStart<-round(0.1*organSize)
1594 21 eStart<-round(0.01*organSize)
1595 22
1596 23 timeMax<-365*2
1597 24 runMax<-100000
1598 25 realCountMax<-5
```

Then we have a script that defines the functions that will be used by the simulation. These functions do a number of different things including return a vector of the neighbour positions of a focal site, finding the position or coordinates of a site in a matrix, return the sites with a particular state in a neighbourhood of a focal site, return a vector of sites given a particular state from the entire organ, and provide criterion of clearance events and time to clearance for each realization.

```
1605 1 #get the list of neighbour positions of a site's position
1606 2 getNeigh<-function(site){
1607 3   site.col<-coords(site)[1]
1608 4   site<-site-1
1609 5
1610 6   vec<-c(
1611 7     (site+1+colSize)%%(colSize)+(site.col-1)*colSize,
1612 8     (site-1+colSize)%%(colSize)+(site.col-1)*colSize,
1613 9     (site-1*colSize+rowSize*colSize)%%(rowSize*colSize),
1614 0     (site+1*colSize+rowSize*colSize)%%(rowSize*colSize)
```

```
  )
  return(vec+1)
}

#get the coordinates on the matrix from a numerical position
coords<-function(pos){
  #returns a vector with the coordinates in the matrix given a
      position
  # returns row position and column position
  xy<-c(ceiling(pos/colSize),pos%%colSize)
  if(xy[2]==0){
    xy[2]<-colSize
  }
  return(xy)
}

#getting the numerical position from coordinates
positn<-function(coords){
  return((coords[1]-1)*colSize+coords[2])
}

#function to get number of states of a particular state
#in a neighbourhood of a site
getNeighStates<-function(organ,site,state){
  # transform site number to position row num, col num
  neighlist<-getNeigh(site)
  numNeigh<-neighlist[which(organ[neighlist]==state)]
  if(length(numNeigh)==0){
    numNeigh<-NA
```

```
  }
  return(numNeigh)
}

#function to get total number of sites of a particular state
getTotalStates<-function(organ,state){
  temp<-which(organ==state)
  return(which(organ==state))
}

#returns 1 if the system has cleared and 0 otherwise
#   clearance is defined as trailing zeros in the vector
is.clear<-function(vec){
    nonZeros<-which(vec!=0)
    if(length(nonZeros)>0){
        zeros<-which(vec==0)
        a<-which(max(nonZeros)<zeros)
        if(length(a)>0){
            b<-1
        }else{
            b<-0
        }
    }else{
        b<-1
    }
    return(b)
}
#A function applied to a time series data set
```

```
#returns the smallest time when a clearence event occurs (NA
    otherwise)
#A clearance event is defined as the first instance of tailing zeros
     in E
timeToClear<-function(timeSeriesData){
    with(as.data.frame(timeSeriesData),{
    nonZeros<-which(E>0)
    zeros<-which(E==0)
    if(length(zeros)>0){
        tail.zeros<-which(zeros>max(nonZeros))
        if(length(tail.zeros)>0){
            end.index<-zeros[min(tail.zeros)]
            b<-time[end.index]
        }else{#no tailing zeros
            b<-NA
        }
    }else{#no zeros in the system
        b<-NA
    }
    return(b)
    })
}
```

We then run the simulation, the following is the code for the global simulation:

```
realCount<-1
#initialize the timeSeries List
timeSeries<-list()

```

```
#start to do all the realizeations
while(realCount<=realCountMax){
  #initialize the list
  count<-1
  organ<-list()

  #organ[[1]]<-matrix(rep(1:4,rowSize*colSize/4),nrow=colSize)
  #initialize a completely healthy site
  organ[[1]]<-matrix(rep(1,rowSize*colSize),nrow=colSize)
  #randomly set abrasions

  stSuscPos<-sample(1:(organSize),sStart)
  organ[[1]][stSuscPos]<-2
  #randomly set an infection
  stExposePos<-sample(1:(organSize),eStart)
  stExpose<-c(stExposePos,getNeigh(stExposePos))
  organ[[1]][stExpose]<-3


  #get the initial counts
  HealCount<-length(getTotalStates(organ[[count]],1))
  SuscCount<-length(getTotalStates(organ[[count]],2))
  ExpoCount<-length(getTotalStates(organ[[count]],3))
  InfeCount<-length(getTotalStates(organ[[count]],4))

  ZCount<-1
  MCount<-0

  #set time to zero
```

```
#run count

#set the timeSeries data frame
timeSeries[[realCount]]<-data.frame(matrix(rep(NA,7*100),ncol=7))
colnames(timeSeries[[realCount]])<-c("time","H","S","E","I","Z","M
    ")

time<-0

while(time<timeMax&&count<(runMax+1)){
  #calculate the rates:
  c_a<-chi*HealCount
  c_r<-rho*SuscCount
  c_inf<-(bet*f/delta)*SuscCount*InfeCount/organSize
  c_gen<-sigma*ExpoCount
  c_clE<-alpha*(ZCount+MCount)*(ExpoCount)/organSize
  c_clI<-alpha*(ZCount+MCount)*(InfeCount)/organSize

  c_tot<-c_a+c_r+c_inf+c_gen+c_clE+c_clI

  #sample a time step
  u<-runif(1)
  timeStep<-log(1-u)/(-c_tot)

  time<-time+timeStep

  #Choose and event:
  event<-runif(1)
```

```
    if(event<c_a/c_tot){
      healList<-getTotalStates(organ[[count]],1)
      if(length(healList)>1){
        eventSite<-sample(healList,1)
      }else{
        eventSite<-healList
      }
      #update site
      temp<-organ[[count]]
      temp[eventSite]<-2
      organ[[count+1]]<-temp
      HealCount<-HealCount-1
      SuscCount<-SuscCount+1
    }else if(event<(c_a+c_r)/c_tot){
      suscList<-getTotalStates(organ[[count]],2)
      if(length(suscList)>1){
        eventSite<-sample(suscList,1)
      }else{
        eventSite<-suscList
      }
      #update site
      temp<-organ[[count]]
      temp[eventSite]<-1
      organ[[count+1]]<-temp
      HealCount<-HealCount+1
      SuscCount<-SuscCount-1
    }else if(event<(c_a+c_r+c_inf)/c_tot){
      suscList<-getTotalStates(organ[[count]],2)
      if(length(suscList)>1){
```

```
1786        eventSite<-sample(suscList,1)
1787      }else{
1788        eventSite<-suscList
1789      }
1790      #update site
1791      temp<-organ[[count]]
1792      temp[eventSite]<-3
1793      organ[[count+1]]<-temp
1794      ExpoCount<-ExpoCount+1
1795      SuscCount<-SuscCount-1
1796    }else if(event<(c_a+c_r+c_inf+c_gen)/c_tot){
1797      expoList<-getTotalStates(organ[[count]],3)
1798      if(length(expoList)>1){
1799        eventSite<-sample(expoList,1)
1800      }else{
1801        eventSite<-expoList
1802      }
1803      #update site
1804      temp<-organ[[count]]
1805      temp[eventSite]<-4
1806      organ[[count+1]]<-temp
1807      ExpoCount<-ExpoCount-1
1808      InfeCount<-InfeCount+1
1809    }else if(event<(c_a+c_r+c_inf+c_gen+c_clE)/c_tot){
1810      #clearance E event occurs
1811      expoList<-getTotalStates(organ[[count]],3)
1812      if(length(expoList)>1){
1813        eventSite<-sample(expoList,1)
1814      }else{
```

```
1815        eventSite<-expoList
1816      }
1817      #update site
1818      temp<-organ[[count]]
1819      temp[eventSite]<-1
1820      organ[[count+1]]<-temp
1821      ExpoCount<-ExpoCount-1
1822      HealCount<-HealCount+1
1823    }else{
1824      infeList<-getTotalStates(organ[[count]],4)
1825      if(length(infeList)>1){
1826        eventSite<-sample(infeList,1)
1827      }else{
1828        eventSite<-infeList
1829      }
1830      #update site
1831      temp<-organ[[count]]
1832      temp[eventSite]<-1
1833      organ[[count+1]]<-temp
1834      InfeCount<-InfeCount-1
1835      HealCount<-HealCount+1
1836    }
1837    time<-time+timeStep
1838    #update the immune cells
1839    if(eps>0){
1840        MCount<-max(MCount+(eps*ZCount*(1-MCount/(Km*organSize)))*
1841            timeStep,0)
1842    }
```

```
    ZCount<-max(ZCount+(zeta*organSize+gam*ZCount*(ExpoCount+
        InfeCount)/organSize-mu*ZCount)*timeStep,0)

    #write to timeSeries list
    timeSeries[[realCount]][count,]<-c(time,HealCount,SuscCount,
        ExpoCount,InfeCount,ZCount,MCount)

    count<-count+1

  }#end of realization
realCount<-realCount+1
}#end of simulation
```

After running the simulation, we aggregate the results into one list and get clearance statistics for the simulation:

```
library(abind)

#trunckate data to be the smallest dimension over all realizations
#Find minimum dimension
mindim<-runMax
for(i in 1:length(timeSeries)){
    newdim<-dim(timeSeries[[i]])[1]
    if(newdim<mindim){
        mindim<-newdim
    }
}
#truckate data to mindim
timeSeries<-lapply(timeSeries,function(x) return(x[1:mindim,]))
```

```
1870 4
1871 5 for(i in 1:(realCountMax)){
1872 6   #Reads input file name
1873 7   tmpf <- timeSeries[[i]]
1874 8   #Take the first data.frame and adjoin it into a multi-dim array
1875 9   if(i==1){
1876 0     Data.Array<-as.array(as.matrix(tmpf))
1877 1   }
1878 2   else{
1879 3   #Take data and adjoin it to the multi-dim array
1880 4     Data.Array<-abind(Data.Array,tmpf,along=3)
1881 5   }
1882 6 }
1883 7
1884 8 #If eps=0 remove M from the list
1885 9 if(eps==0){
1886 0     Data.Array<-Data.Array[,-7,]
1887 1 }
1888 2
1889 3 Data.Avg<-apply(Data.Array,c(1,2),mean,na.rm=T)
1890 4
1891 5 head(Data.Array[,,1])
1892 6 head(Data.Avg)
1893 7
1894 8 #calculate the proportion of clearance
1895 9 clear.mat<-apply(Data.Array,c(2,3),is.clear)
1896 0 clear.prop<-apply(clear.mat,1,mean)
1897 1
1898 2 clear.times<-apply(Data.Array,3,timeToClear)
```

```
1899
1900 print(clear.prop)
1901 print(clear.times)
1902 print(mean(clear.times,na.rm=T))
1903 print(var(clear.times,na.rm=T))
1904
1905 x<-list(clear.prop=clear.prop,
1906          clear.times=clear.times,
1907          meanTime=mean(clear.times,na.rm=T),
1908          varTime=var(clear.times,na.rm=T)
1909 )
```

Then we either plot the system as a time series using the `ggplot2` package [60], using `dplyr` and `dplyr` to clean up the data for plotting:

```
1912 library(tidyr)
1913 library(dplyr)
1914 library(ggplot2)
1915
1916 theme_set(theme_bw())
1917
1918 Data.Max<-apply(Data.Array,c(1,2),max)
1919 Data.Min<-apply(Data.Array,c(1,2),min)
1920
1921 Data.QuantLow<-apply(Data.Array,c(1,2),function(x) quantile(x,probs
1922     =0.05))
1923 Data.QuantUp<-apply(Data.Array,c(1,2),function(x) quantile(x,probs
1924     =0.95))
1925 print(names(Data.Avg))
```

```

gData.Avg<-gather(as.data.frame(Data.Avg),state,num,-time)
gData.Up<-gather(as.data.frame(Data.QuantUp),state,num,-time)
gData.Low<-gather(as.data.frame(Data.QuantLow),state,num,-time)

gData.Avg$up<-gData.Up$num
gData.Avg$low<-gData.Low$num

col.pal<-c("dodgerblue","darkorchid1","orange","red","black","pink1"
    )

print(
  ggplot(gData.Avg, aes(x=time, y=num))
  + geom_line(aes(color=state,linetype=state),size=1)
  + geom_ribbon(aes(fill=state,ymin=low,ymax=up),alpha=0.3)
  + scale_fill_manual(values=col.pal)
  + scale_color_manual(values=col.pal)
  + scale_linetype_manual(values=c(1,1,2,2,3,3))
)
```

Or we plot the simulation on a grid using the `lattice` package [45].

```
library(lattice)
library(gridExtra)


org.col.pal<-c("white","pink","red","firebrick")

```

```
time.list<-c(1,100,1000,2500)
par(mfrow=c(2,2))
p<-list()
for(n in time.list){

    p[[which(n==time.list)]]<-levelplot(organ[[n]],col.regions=org.
        col.pal,at=seq(0.5,4.5,by=1),
                xlab="",
                ylab="",
                main=paste("Site States at time t =",round(timeSeries
                    [[realCountMax]][n,"time"]),"days"),
                colorkey=list(col=org.col.pal,tick.number=4,at=seq
                    (0.5,4.5,by=1),
                                labels=list(labels=c("H","S","E","I","")
                                    ,cex=1.1)
                                )
    )

}

grid.arrange(p[[1]],p[[2]],p[[3]],p[[4]],ncol=2,nrow=2)
```

The local simulation uses the same process, but it has a slightly different implementation of the simulation script. This is outlined below. The parameter file, functions used in the simulation, and time series analysis and plotting are all the same for the local simulation as the global simulation.

```
realCount<-1
#initialize the timeSeries List
```

```
timeSeries<-list()

#start to do all the realizeations
while(realCount<=realCountMax){
  #initialize the list
  count<-1
  organ<-list()

  #organ[[1]]<-matrix(rep(1:4,rowSize*colSize/4),nrow=colSize)
  #initialize a completely healthy site
  organ[[1]]<-matrix(rep(1,rowSize*colSize),nrow=colSize)
  #randomly set abrasions
  stSuscPos<-sample(1:(organSize),sStart)
  organ[[1]][stSuscPos]<-2
  #randomly set an infection
  stExposePos<-sample(1:(organSize),eStart)
  stExpose<-c(stExposePos,getNeigh(stExposePos))
  organ[[1]][stExpose]<-3

  #get the initial counts
  HealCount<-length(getTotalStates(organ[[count]],1))
  SuscCount<-length(getTotalStates(organ[[count]],2))
  ExpoCount<-length(getTotalStates(organ[[count]],3))
  InfeCount<-length(getTotalStates(organ[[count]],4))

  #set up a matrix counting the number of sites
  Zmat<-matrix(rep(1/organSize,organSize),nrow=rowSize,ncol=colSize)
  ZCount<-sum(Zmat)
  Mmat<-matrix(rep(0/organSize,organSize),nrow=rowSize,ncol=colSize)
```

```
MCount<-sum(Mmat)

#set time to zero
#run count

#set the timeSeries data frame
timeSeries[[realCount]]<-data.frame(matrix(rep(NA,7*100),ncol=7))
colnames(timeSeries[[realCount]])<-c("time","H","S","E","I","Z","M
    ")

time<-0

while(time<timeMax&&count<(runMax+1)){

    #infectiousMat<-(organ[[count]]==2)*getNeighStateMatrix(organ
        [[count]],4)
    immuneEMat<-(organ[[count]]==3)*Zmat+(organ[[count]]==3)*Mmat
    immuneIMat<-(organ[[count]]==4)*Zmat+(organ[[count]]==4)*Mmat

    #calculate the rates:
  c_a<-chi*HealCount
  c_r<-rho*SuscCount
  c_inf<-bet*f/delta*InfeCount
  c_gen<-sigma*ExpoCount
  # Clearance events are equal to the immune response at a site
  # with an E or I value
  c_clE<-sum(alpha*immuneEMat)
  c_clI<-sum(alpha*immuneIMat)

```

```
    c_tot<-c_a+c_r+c_inf+c_gen+c_clE+c_clI


    #sample a time step
    u<-runif(1)
    timeStep<-log(1-u)/(-c_tot)


    #Choose and event:
    event<-runif(1)

    if(event<c_a/c_tot){
      healList<-getTotalStates(organ[[count]],1)
      if(length(healList)>1){
        eventSite<-sample(healList,1)
      }else{
        eventSite<-healList
      }
      #update site
      temp<-organ[[count]]
      temp[eventSite]<-2
      organ[[count+1]]<-temp
      HealCount<-HealCount-1
      SuscCount<-SuscCount+1
    }else if(event<(c_a+c_r)/c_tot){
      suscList<-getTotalStates(organ[[count]],2)
      if(length(suscList)>1){
        eventSite<-sample(suscList,1)
      }else{
```

```
      eventSite<-suscList
    }
    #update site
    temp<-organ[[count]]
    temp[eventSite]<-1
    organ[[count+1]]<-temp
    HealCount<-HealCount+1
    SuscCount<-SuscCount-1
  }else if(event<(c_a+c_r+c_inf)/c_tot){
    infeList<-getTotalStates(organ[[count]],4)
    if(length(infeList)>1){
        eventSiteI<-sample(infeList,1)
    }else{
        eventSiteI<-infeList
    }

    #sample a neighbour around the infected site I
          eventSiteS<-sample(getNeigh(eventSiteI),1)
          #check to see if the sampled neighbour is susceptible
          if(organ[[count]][eventSiteS]==2){
              #update site
        temp<-organ[[count]]
        temp[eventSiteS]<-3
        organ[[count+1]]<-temp
        ExpoCount<-ExpoCount+1
        SuscCount<-SuscCount-1
          }else{
              temp<-organ[[count]]
        organ[[count+1]]<-temp
```

```
2096            }
2097    }else if(event<(c_a+c_r+c_inf+c_gen)/c_tot){
2098      expoList<-getTotalStates(organ[[count]],3)
2099      if(length(expoList)>1){
2100        eventSite<-sample(expoList,1)
2101      }else{
2102        eventSite<-expoList
2103      }
2104      #update site
2105      temp<-organ[[count]]
2106      temp[eventSite]<-4
2107      organ[[count+1]]<-temp
2108      ExpoCount<-ExpoCount-1
2109      InfeCount<-InfeCount+1
2110    }else if(event<(c_a+c_r+c_inf+c_gen+c_clE)/c_tot){
2111            eventSite<-sample(1:organSize,1,prob=immuneEMat/sum(
2112                immuneEMat))
2113      #update site
2114      temp<-organ[[count]]
2115      temp[eventSite]<-1
2116      organ[[count+1]]<-temp
2117      ExpoCount<-ExpoCount-1
2118      HealCount<-HealCount+1
2119    }else{
2120      #print("Clear I")
2121            eventSite<-sample(1:organSize,1,prob=immuneIMat/sum(
2122                immuneIMat))
2123      #update site
2124      temp<-organ[[count]]
```

```
2125        temp[eventSite]<-1
2126        organ[[count+1]]<-temp
2127        InfeCount<-InfeCount-1
2128        HealCount<-HealCount+1
2129      }
2130      time<-time+timeStep
2131      #update the immune cells
2132      if(eps>0){
2133          Mmat<-Mmat+(eps*Zmat*(1-Mmat/(Km*organSize)))*timeStep
2134          Mcount<-sum(Mmat)
2135      }
2136      Zmat<-Zmat+(zeta+gam*Zmat*((organ[[count+1]]==3)+organ[[count
2137          +1]]==4)-mu*Zmat)*timeStep
2138      ZCount<-sum(Zmat)
2139
2140      #write to timeSeries list
2141      timeSeries[[realCount]][count,]<-c(time,HealCount,SuscCount,
2142          ExpoCount,InfeCount,ZCount,MCount)
2143
2144      count<-count+1
2145
2146
2147    }#end of realization
2148  realCount<-realCount+1
2149  }#end of simulation
```

D.3 Multi-type HPV Model Solver Code

The multi-type HPV model is implemented similarly to the base model. We define a gradient function in **R** and solve it using the `ode()` function in the `deSolve` package [50]. The results are printed in the basic **R** plotting function.

Below is the gradient function

```
library(deSolve)

MultiStrain.vf<-
  function(t,vars,
           parms=c(K=1,N=1)) {
    with(as.list(c(parms,vars)),{
      dB<--chi*B+rho*S+alpha[1]*Z1*(E1+I1)/N+alpha[2]*Z2*(E2+I2)/N+q
         *alpha[1]*Z1*(E2+I2)/N+q*alpha[2]*Z2*(E1+I1)/N
      dS<-chi*B-rho*S-bet[1]*S*V1/N-bet[2]*S*V2/N
      dE1<-bet[1]*S*V1/N-sigma*E1-alpha[1]*Z1*E1/N-q*alpha[2]*Z2*E1/
         N
      dE2<-bet[2]*S*V2/N-sigma*E2-alpha[2]*Z2*E2/N-q*alpha[1]*Z1*E2/
         N
      dI1<-sigma*E1-alpha[1]*Z1*I1/N -q*alpha[2]*Z2*I1/N
      dI2<-sigma*E2-alpha[2]*Z2*I2/N-q*alpha[1]*Z1*I2/N
      dV1<-f1*I1-delta[1]*V1
      dV2<-f2*I2-delta[2]*V2
      dZ1<-zeta*N/2+gam[1]*Z1*(E1+I1)/N+q*gam[1]*Z1*(E2+I2)/N-mu*Z1
      dZ2<-zeta*N/2+gam[2]*Z2*(E2+I2)/N+q*gam[2]*Z2*(E1+I1)/N-mu*Z2
      res<-c(dB=dB,dS=dS,dE1=dE1,dE2=dE2,dI1=dI1,dI2=dI2,dV1=dV1,dV2
         =dV2,dZ1=dZ1,dZ2=dZ2)
      list(res)
    })
```

```
  }
```

and then solved using the following code.

```
library(deSolve)

soln<-ode(
    y=c(B=B0,S=S0,E1=E10,E2=E20,I1=I10,
      I2=I20,V1=V10,V2=V20,Z1=Z10,Z2=Z20),
        times=seq(from=0,to=tmax,by=tmax/interNum),
        func=MultiStrain.vf,
        parms=c(alpha=alpha,bet=bet,f=f,gam=gam,
          delta=delta,zeta=zeta,chi=chi,rho=rho,
          sigma=sigma,mu=mu,q=q)
)
```

D.4 Super-infection Patch Model Code

Once again the code for the super-infection patch models is very similar, we set up three different patch model gradient functions and then solve them using the `ode()` function in the `deSolve` package [50].

The asymmetrical patch model gradient is defined here.

```
ms.patch<-
  function(t,vars,
           parms=c(K=1,N=1)) {
    with(as.list(c(parms,vars)),{
```

```
    S<-(1-y1-y2-z)
    Lam1<-beta[1]*(y1+p[1]*z)
    Lam2<-beta[2]*(y2+p[2]*z)
    dy1=Lam1*S-y1-Lam2*y1
    dy2=Lam2*S-y2-Lam1*y2
    dz=Lam2*y1+Lam1*y2-z
    dScheck = -(dy1+ dy2+ dz)
    res<-c(dy1=dy1,dy2=dy2,dz=dz, dScheck=dScheck)
    list(res)
    })
  }
```

The Same Strain Super-Infection model gradient is defined here.

```
ms.patch<-
  function(t,vars,
           parms=c(K=1,N=1)) {
    with(as.list(c(parms,vars)),{
    S<-(1-y1-y2-z-z1-z2)
    Lam1<-beta[1]*(y1+p[1]*z+z1)
    Lam2<-beta[2]*(y2+p[2]*z+z2)
    dy1=Lam1*S-y1-Lam2*y1-Lam1*y1
    dy2=Lam2*S-y2-Lam1*y2-Lam2*y2
    dz=Lam2*y1+Lam1*y2-z
    dz1=Lam1*y1-z1
    dz2=Lam2*y2-z2
    dScheck = -(dy1+ dy2+ dz+ dz1+ dz2)
    res<-c(dy1=dy1,dy2=dy2,dz=dz,dz1=dz1,dz2=dz2, dScheck=dScheck)
    list(res)
```

```
    })
  }
```

And lastly, the FIFO Super-Infection model gradient is defined below.

```
ms.patch<-function(t,vars,
  parms=c(K=1,N=1)) {
    with(as.list(c(parms,vars)),{
      #Parameters
    S<-(1-y1-y2-z12-z21-z1-z2)
    Lam1<-beta[1]*(y1+p[1]*z12+p[1]*z21+z1)
    Lam2<-beta[2]*(y2+p[2]*z12+p[2]*z21+z2)
    #ODEs
    dy1=Lam1*S-Lam2*y1-Lam1*y1-y1
    dy2=Lam2*S-Lam1*y2-Lam2*y2-y2
    dz12=Lam2*y1+Lam2*z1-Lam2*z12-Lam1*z12+Lam2*z21-z12
    dz21=Lam1*y2+Lam1*z2-Lam1*z21+Lam1*z12-Lam2*z21-z21
    dz1=Lam1*y1+Lam1*z21-Lam2*z1-z1
    dz2=Lam2*y2+Lam2*z12-Lam1*z2-z2
    dScheck = -(dy1+ dy2+ dz12 + dz21 + dz1+ dz2)
    res<-c(dy1=dy1,dy2=dy2,dz12=dz12,dz21=dz21,dz1=dz1,dz2=dz2,
        dScheck=dScheck)
    list(res)
  })
}
```

These are solved using the `ode()` function, we only show the example for the asymmetric model. Solvers for the other models are programmed accordingly.

```
library(deSolve)

soln<-ode(y=c(y1=y1,y2=y2,z=z, Scheck=Scheck),
          times=seq(0,tmax,by=tmax/inter),
          func=ms.patch,
          parms=c(beta=beta,p=p))

# add in susceptible patches
soln <- within(as.data.frame(soln), {
  S <- 1-y1-y2-z
})

#remove the Scheck
soln<-soln[,c(1:4,6)]
```

All results were plotted using the `ggplot2` package [60] and using the `dplyr` and `tidyr` packages to prepare the data frame for printing.

```
library(tidyr)
library(dplyr)
library(ggplot2)

theme_set(theme_bw())

print(names(soln))

patches <- soln %>% gather(State, Num,-time)
print(summary(patches))

```

```
2279 col.pal<-c("blue","red","darkorchid","gray40")
2280
2281 print(
2282   ggplot(patches, aes(x=time, y=Num))
2283   + geom_line(aes(color=State,linetype=State),size=1)
2284   + scale_color_manual(values=col.pal)
2285   + scale_linetype_manual(values=c(2,3,4,1))
2286 )
```

Bibliography

[1] Samuel Alizon. Co-infection and super-infection models in evolutionary epidemiology. *Interface Focus*, 3, 2013.

[2] Janet G Baseman and Laura A Koutsky. The epidemiology of human papillomavirus infections. *Journal of Clinical Virology*, 32, 2005.

[3] Sanjay Basu, Gretchen B Chapman, and Alison P Galvani. Integrating epidemiology, psychology, and economics to achieve HPV vaccination targets. *PNAS*, 105(48), 2008.

[4] Rob J De Boer, Mihaela Oprea, Rustom Antia, Kaja Murali-Krishna, Rafi Ahmed, and Alan S Perelson. Recrusitment times, proliferation, and apoptosis rates during the CD8+ T-cell response to lymphocytic choriomeningitis virus. *Journal of Virology*, 75(22), 2001.

[5] JR Carter, Z Ding, and BR Rose. HPV infection and cervical disease: a review. *Aust. N.Z. J. Obstet. Gynaecol.*, 51, 2011.

[6] Xavier Castellsagué. Natural history and epidemiology of HPV infection and cervical cancer. *Gynecologic Oncology*, 110, 2008.

[7] Anil K Chaturvedi, HA Katki, A Hildesheim, AC Rodriguez, W Quint, M Schiffman, LJ Van Doorn, C Porras, S Wacholder, P Gonzalez, ME Sherman, and R Herrero. Human papillomavirus with multiple types: pattern of coinfection and risk of cervical disease. *JID*, 203, 2011.

[8] Timothy Culp and Neil Christensen. Kinetics of *in vitro* adsorption and entry of papillomavirus virions. *Virology*, 319, 2004.

[9] Erik Dasbach, Elamin Elbasha, and Ralph Insinga. Mathematical models for predicting the epidemiologic and economic impact of vaccination against human papillomavirus infection and disease. *Epidemiol Rev*, 28, 2006.

[10] A de Jong. Frequent detection of human papillomavirus 16 E2-specific T-helper immunity in healthy subjects. *Cancer Res.*, 62, 2002.

[11] EM de Villiers, C Fauquet, T Broker, HU Bernard, and H zur Hausen. Classification of papillomaviruses. *Virology*, 324, 2004.

[12] René Descartes. *La géométrie de René Descartes*. A. Hermann, Paris, 1886.

[13] J Dillner, M Arbyn, E Unger, and L Dillner. Monitoring of human papillomavirus vaccination. *Clinical and Experimental Immunology*, 163, 2010.

[14] Elamin Elbasha and Erik Dasbach. Impact of vaccinating boys and men against HPV in the United States. *Vaccine*, 28, 2010.

[15] Elamin Elbasha, Erik Dasbach, and Ralph Insinga. A multi-type HPV transmission model. *Bulletin of Mathematical Biology*, 70, 2008.

[16] Elamin Elbasha and Alison Galvani. Vaccination against multiple HPV types. *Mathematical Biosciences*, 197, 2005.

[17] Y Woo et al. Characteriszing the local immune responses in cervical intraepithelial neoplasia: a cross-sectional and longitudinal analysis. *BJOG*, 115, 2010.

[18] J Ferlay, I Soerjomataram, M Ervik, R Dikshit, S Eser, C Mathers, M Rebolo, DM Parkin, D Forman, and F Bray. GLOBOCAN 2012 v1.0 Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]., 2013. Accessed on July 31, 2015.

[19] International Agency for Research on Cancer. IARC Monographs on the Evaluations of the Carcinogenic Risk to Humans: Biological agents - Human Papillomaviruses. *IARC*, 100B, 2012.

[20] D Forman, C de Martel, CJ Lacey, I Soerjomataram, J Lortet-Tieulent, L Bruni, J Vignat, J Ferlay, F Bray, M Plummer, and S Franceschi. Global burden of human papillomavirus and related diseases. *Vaccine*, 30S, 2012.

[21] SM Garland. Quadravalent vaccine against human papillomavirus to prevent anogenital diseases. *N. Engl. J. Med.*, 356, 2007.

[22] C Giancinti and A Giordano. Rb and cell cycle progression. *Oncogene*, 25, 2006.

[23] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25), 1977.

[24] Patti E Gravitt. The known unknowns of HPV natural history. *The Journal of Clinical Investigation*, 121(12), December 2011.

[25] Jennifer Hamborsky, Andrew Kroger, and Charles (Skip) Wolfe, editors. *Epidemiology and Prevention of Vaccine-Preventable Diseases*. Centers for Disease Control and Prevention, 13 edition, April 2015.

[26] James P Hughes, Geoff P Garnett, and Laura Koutsky. The theoretical population-level impact of a prophylactic human papilloma virus vaccine. *Epidemiology*, 13(6), November 2002.

[27] A Hurwitz. Über die Bedingungen unter Welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt. *Math. Ann.*, 46, 1895.

[28] Mimmo Iannelli, Maia Martcheva, and Xue-Zhi Li. Strain replacement in an epidemic model with super-infection and perfect vaccination. *Mathematical Biosciences*, 195, 2005.

[29] Jessica Kahn, Darron Brown, Lili Ding, Lea Widdice, Marcia Shew, Susan Glynn, and David Bernstein. Vaccine-type human papillomavirus and evidence of herd protection after vaccine introduction. *Pediatrics*, 130(2), August 2012.

[30] Mark S Lechner and Laimonis A Laimins. Inhibition of p53 DNA binding by human papillomavirus E6 proteins. *Journal of Virology*, 68(7), 1994.

[31] Joaquin Luna, Manuel Plata, Mauricio Gonzalez, Alfonso Correa, Ivete Maldonado, Claudia Nossa, David Radley, Scott Vuocolo, Richard Haupt, and Alfred Saah. Long-term follow-up observation of the safety, immunogenicity, and effectiveness of Guardasil in adult women. *PLoS One*, 8(12), December 2013.

[32] Margaret E McLaughlin-Drubin and Craig Meyers. Evidence for the coexistence of two genital HPV types. *Virology*, 321, 2004.

[33] AB Moscicki, M Schiffman, S Kjaer, and LL Villa. Chapter: 5 updating the natural history of HPV and anogenital cancer. *Vaccine*, 24, 2006. (Suppl 3).

[34] Carmen L Murall, Chris T Bauch, and Troy Day. Could the human papillomavirus vaccines drive virulence evolution. *Proc. R. Soc. Lond. B*, 282, 2014.

[35] Carmen Lía Murall, Kevin S. McCann, and Chris T. Bauch. Revising ecological assumptions about Human papillomavirus interactions and type replacement. *Journal of Theoretical Biology*, 350, 2014.

[36] Mayumi Nakagawa, William Greenfield, Andrea Moerman-Herzog, and Hannah N Coleman. Cross-reactivity, epitope spreading and *de novo* immune stimulation are possible mechanisms of cross-protection of nonvaccine human papillomavirus (HPV) types in recipients of HPV therapeutic vaccines. *Clinical and Vaccine Immunology*, 22(7), July 2015.

[37] Public Health Agency of Canada. update on the recommended human papillomavirus (HPV) vaccine immunization schedule. *PHAC*, 2015.

[38] World Health Organization. Human papillomavirus (HPV) and cervical cancer, March 2015. Accessed July 31, 2015.

[39] Marianne De Paepe and François Taddei. Viruses' life history: towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS Biol.*, 4(7), 2006.

[40] Raúl Peralta, Cruz Vergas-De-León, Augusto Cabrera, and Pedro Miramontes. Dynamics of high-risk nonvaccine human papillomavirus types after actual vaccination scheme. *Computational and Mathematical Methods in Medicine*, 2014.

[41] Eric Poolman, Elamin Elbasha, and Alison Galvani. Vaccination and the evolutionary ecology of human papillomavirus. *Vaccine*, 26, 2008.

[42] AJ Quayle. The innate and early immune to pathogen challenge in the female genital tract and the pivotal role of epithelial cells. *Journal of Reproductive Immunology*, 57, 2002.

[43] CM Roteli-Martins, P Naud, P De Borba, JC Teixeira, NS De Carvalho, T Zahaf, N Sanchez, B Geeraerts, and D Descamps. Sustained immunogenicity and efficacy of the HPV-16/18 AS06-adjuvanted vaccine. *Human Vaccines & Immunotherapeutics*, 8(3), 2012.

[44] EJ Routh. *A Treatise on the Stability of a Given State of Motion Particularly Steady Motion*. MacMillon, 1877.

[45] Deepayan Sarkar. *Lattice: Multivariate Data Visualization with R*. Springer, New York, 2008. ISBN 978-0-387-75968-5.

[46] Mirte Scherpenisse, Rutger M Schepp, Madelief Mollers, Chris JLM Meijer, Guy AM Berbers, and Fiona RM van der Kils. Characteristics of HPV-specific antibody responses induced by infection and vaccination: cross-reactivity, neutralizing activity, avidity and IgG subclasses. *PLoS one*, 8(9), September 2013.

[47] NF Schlect, S Kulaga, J Robitaille, S Ferreira, M Santos, RA Miyamura, E Duarte-Franco, TE Rohan, A Ferenczy, LL Villa, and EL Franco. Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. *JAMA*, 286(24), December 2001.

[48] Phillip H Shaw. The role of p53 in cell cycle regulation. *Path. Res. Pract.*, 192, 1996.

[49] CDC Fact Sheet. Genital HPV infection, 2014. Accessed on July 31, 2015.

[50] Karline Soetaert, Thomas Petzoldt, and R. Woodrow Setzer. Solving differential equations in R: Package desolve. *Journal of Statistical Software*, 33(9):1–25, 2010.

[51] Margaret Stanley. Immune responses to human papillomavirus. *Vaccine*, S4S1, 2006.

[52] Margaret Stanley. HPV-immune response to infection and vaccination. *Infectious Agents and Cancer*, 5(19), 2010.

[53] Margaret Stanley, Douglas R Lowy, and Ian Frazer. Prophylactic HPV vaccines: underlying mechanisms. *Vaccine*, 24S3, 2006.

[54] Margaret Stanley, Ligia A Pinto, and Connie Trimble. Human papillomavirus vaccines - immune responses. *Vaccine*, 30S, 2012.

[55] Margaret A Stanley. Epithelial cell response to infection with human papillomavirus. *Clin. Microbiol. Rev.*, 25(2), 2012.

[56] JE Tota, AV Ramanakumar, M Jiang, J Dillner, SD Walter, JS Kaufman, F Coutlée, LL Villa, and EL Franco. Epidemiological approaches to evaluating the potential for human papillomavirus type replacement post vaccination. *American Journal of Epidemiology*, 178(4), 2013.

[57] Helen Trottier and Eduardo L Franco. The epidemiology of genital human papillomavirus infection. *Vaccine*, 24, 2006.

[58] Minus van Baalen and Maurice W Sabelis. The dynamics of multiple infection and evolution of virulence. *The American Naturalist*, 146(6), 1995.

[59] P van den Driessche and James Watmough. Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.*, 180, 2002.

[60] Hadley Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009.

[61] OM Williams, KW Hart, ECY Wang, and CM Gelder. Analysis of CD4+ T-cells responses to human papillomavirus (HPV) type 11 L1 in healthy adults reveals a high degree of responsiveness and cross-reactivity with other HPV types. *J. Virol.*, 76(15), 2002.

[62] Long Fu xi, Zoe R Edelstein, and Craig Meyers et al. Human papillomavirus types 16 and 18 DNA load in relation to coexistence of other types, particularly those in the same species. *Cancer Epidemiol Biomarkers Prev*, 18, 2009.

[63] Eun-Kyoung Yim and Jong-Sup Park. The role of HPV H6 and E7 oncoproteins in HPV-associated cervical cacinogenesis. *Cancer Res Treat.*, 37(6), 2005.