Insights into the Emergence of Novel Infectious Diseases to Humans

Ruben Josef Kubiak
Linacre College
University of Oxford

Thesis submitted for the degree of
Doctor of Philosophy

Hilary 2012

Für Angelika, die mir die Möglichkeit gibt meine Träume in der Welt zu verwirklichen und dabei immer zu mir hält.

ACKNOWLEDGEMENTS

First, I would like to thank Angela McLean for all her support, expert advice, patience, and excellent supervision. Julie Ball helped with friendly advice on administrative matters. Nim Pathy helped me with some tricky mathematics, and Helen Fryer provided me with some great British tea experiences and answers to my odd questions about the English language. Michael Bonsall and Charles Godfray gave me insights into theoretical ecology, which was very helpful for defining my modelling approach. I am also grateful to the Oxford Martin School (James Martin Fellowship) and the Environmental and Physical Sciences Research Council (EPSRC scholarship) for financial support. Finally, I would like to thank all my friends and family for supporting me throughout my studies.

"We must not forget that when radium was discovered no one knew that it would prove useful in hospitals. The work was one of pure science. And this is a proof that scientific work must not be considered from the point of view of the direct usefulness of it. It must be done for itself, for the beauty of science, and then there is always the chance that a scientific discovery may become like the radium a benefit for humanity."

- Marie Curie, *Lecture at Vassar College, 14 May 1921*

STATEMENT OF CONTRIBUTION

Parts of this thesis include collaborative research between Ruben J. Kubiak (RJK), Nimalan Arinaminpathy (NA), and Angela R. McLean (ARM). In the following, the contribution of each collaborator is stated for each research chapter:

- Chapter 3 presents research which was conceived and designed by: RJK, NA, and ARM. The mathematical models were developed by: RJK and NA. The programming and simulations were done by: RJK. Tools for analysis were provided by: RJK. The data was analysed by: RJK. The peer-reviewed article was written by: RJK, NA, and ARM.

- Chapter 4 presents research which was conceived and designed by: RJK. The mathematical models were developed by: RJK. The programming and simulations were done by: RJK. Tools for analysis were provided by: RJK. The data was analysed by: RJK.

- Chapter 5 presents research which was conceived and designed by: RJK and ARM. The mathematical models were developed by: RJK and ARM. The programming and simulations were done by: RJK. Tools for analysis were provided by: RJK. The data was analysed by: RJK.

- Chapter 6 presents research which was conceived and designed by: RJK and ARM. The mathematical models were developed by: RJK. The programming was done by: RJK. Tools for analysis were provided by: RJK. The data was analysed by: RJK and ARM. The peer-reviewed article was written by: RJK and ARM.

ABSTRACT

Novel infectious diseases in humans are of great concern to public health authorities and researchers in epidemiology. Zoonotic pathogens in particular have the potential to cause epidemics without any or little warning. In this thesis, I investigate evolutionary and environmental conditions, and the interactions between both, which facilitate the zoonotic emergence of novel pathogens. I start with a list of the mechanisms and processes which might influence a zoonotic emergence, and identify some unsolved problems. I address these with multiple, theoretical models. First, I use a village-city model with different adaptation scenarios to examine the influence of spatial heterogeneity on the emergence process. I derive general analytical results for the statistical properties of emergence events, including the probability distribution of outbreak sizes. My results suggest that, for typical connection strengths between communities, spatial heterogeneity has only a weak effect on outbreak size distributions, and on the risk of emergence per introduction. Next, I extend the research on environmental conditions by looking at pathogen specialisation in multi-host systems. I derive threshold connectivities for which generalist pathogens, which infect multiple species and might therefore be more dangerous to cross into the human species, can sustain transmission and are not dominated by specialists, which can only cause sustained transmission chains in a single host species, but are able to cause emergences with little warning. My third research chapter is interested in the effect of the loss of biodiversity. I analytically derive expected prevalences for fast growing and slow growing species. If fast growing species tend to perform better in degraded environments, my analytical results suggest that the overall prevalence level of infectious diseases will rise as environments degrade, which facilitates the chance of zoonotic jumps. In my last research chapter, I examine the actual impact of a novel, emerging infectious disease. I use data from the recent 'Swine flu' epidemic in England to estimate epidemiological parameters of the infectious agent. My results suggest that the majority of infected cases showed no or only mild symptoms. This reveals that more data than just the estimated number of cases are necessary to fully evaluate the danger of a possible zoonotic, emerging infectious disease. I conclude by discussing my results and the implications which these might have.

CONTENTS

CHAPTER 1

THE BIOLOGY OF ZOONOTIC INFECTIOUS DISEASES

1.1 An Introduction to Emerging Infectious Diseases

Emerging infectious diseases (EIDs) are one of today's important research areas. A better understanding of underlying factors and processes of infectious disease emergence could drive novel drug therapies and enhancements in public health. But EIDs are only a fraction in the spectrum of infectious diseases as the number and variety of infectious disease is so large. Recent estimates have listed 1,407 (Woolhouse and Gowtage-Sequeria, 2005) to 1,415 (Taylor et al., 2001; Woolhouse, 2002) species known to be pathogenic to humans. Therefore, a definition of EIDs is a necessary starting point. This definition needs to be broad and specific at the same time. Diseases like HIV need to have fulfilled this definition when they first occurred, or still fulfil it when spreading into new regions where they were previously unknown. But the definition also needs to include the re-emergence of previously circulating diseases which were considered almost eradicated like measles, but also drug-resistant tuberculosis.

The scientific literature offers a range of different and sometimes contradicting definitions. For example, Morse (1995) defines EIDs as newly appearing infections, or existing ones which rapidly increase in incidence or geographic range. Being a very broad definition, it includes novel infectious diseases which are not necessarily increasing in incidence. An example for this is the marburg virus, which is neither endemic nor epidemic in the human population. This pathogen has previously been introduced into the human population and caused some outbreaks. However, these outbreaks have never yet 'emerged' into a large scale phenomenon, even though the marburg virus would fulfil the definition of Morse (1995). Furthermore, Cliff et al. (2009) argue that this definition is still too narrow, because it excludes the recognition that the knowledge about an emergence might just be a result of improved medical research and surveillance. Yet, the knowledge about it does not change the fact that an infectious disease is circulating within a population, increasing or decreasing its incidence, or adapting to a drug treatment.

As a result, I deem the above mentioned definition as not sufficient for the purpose of this thesis. Woolhouse and Dye (2001) offer a definition which fits much better. Following them, an EID is:

> "an infectious disease whose incidence is increasing following its first introduction into a new host population or whose incidence is increasing in an existing host population as a result of long-term changes in its underlying epidemiology."

Here, the change in the underlying epidemiology could be the result of a variety of changes, for example changes in the biology of the pathogen, host demography, host behaviour, or in the ecology of the transmission process. This definition therefore covers scenarios like the re-introduction of malaria into populations from which it had been eradicated a long time ago, the introduction of newly infectious diseases like SARS, or the emergence of drug-resistant pathogen strains. On the other hand, it excludes just the appearance of a novel pathogen as sufficient condition for emergence. It also does not depend on our knowledge of the pathogen's existence.

Still, an increase in medical and public health knowledge will favour surveillance and control. Research on infectious diseases allows new insights, and is the basis for our knowledge about EIDs. The scientific literature estimates that $12 - 13\%$ of all infectious diseases known to humans are EIDs by applying Woolhouse and Dye (2001)'s definition (Woolhouse, 2002; Woolhouse and Gowtage-Sequeria, 2005). But EIDs are not homogeneously distributed over all taxa. For example, Woolhouse (2002), and with updated data Woolhouse and Gowtage-Sequeria (2005), show that RNA viruses are the most prominent taxon in human EIDs. They have a much higher nucleotide substitution rate which might allow for faster adaptation and increased chances of generating the necessary mutations necessary for an emergence. It is therefore not surprising that pathogen evolution is overall the fifth-highest risk factor for EIDs to develop, considering for example its importance in drug-resistance (Woolhouse and Gowtage-Sequeria, 2005). Taylor et al. (2001) support these findings by measuring risk factors for human disease emergence by reviewing the published literature on infectious diseases. Their results show that the emergence of an infectious disease is most likely if the pathogen is a virus or a protozoa while helminths show the lowest potential to emerge.

Jones et al. (2008) use a different approach by counting the actual number of emergence events per pathogen rather than by looking at the distribution of pathogens known

to cause emergences. Using a dataset with 355 emergence events between 1940 and 2004, they show that over half of all emergence events are caused by bacteria or rickettsia, highlighting the large number of emergences caused by drug-resistant strains. Furthermore, EIDs are significantly correlated with socio-economic, environmental, and ecological factors such as population density or population growth. This result is also reflected in the worldwide distribution of EID events. While comparatively few EID events happened in Africa and South America, most emergences started in North America and Europe (Jones et al., 2008). Nonetheless, it is quite possible that these data have some reporting bias as rich, western countries have more resources to investigate and report EIDs. But an emergence of a drug-resistant strain can only happen in a country where this drug is actually in use, leading to a heterogeneous distribution with emphasis on rich, western countries.

One of the most important risk factors influencing the emergence potential of an infectious disease is the zoonotic status. EIDs tend to encompass broad host ranges, often including several mammalian or even non-mammalian orders (Woolhouse, 2002; Woolhouse and Gowtage-Sequeria, 2005). Zoonotic pathogens are twice as likely to become an EID (Taylor et al., 2001; Woolhouse and Gowtage-Sequeria, 2005). Jones et al. (2008) reveal that 60.3% of all EIDs from 1940 to 2004 resulted from zoonoses, showing the importance of increasing our knowledge of zoonotic infectious disease.

1.1.1 A Brief Overview of Zoonotic Infectious Diseases

Zoonotic diseases are defined as those which can pass the inter-species barrier from its reservoir host species, the donor species, into the human host species, the recipient species (Keeling and Rohani, 2007). Estimates show that approximately two-thirds to three-quarters of all emerging human pathogens are zoonotic (Taylor et al., 2001; Woolhouse, 2002; Woolhouse and Gowtage-Sequeria, 2005; Jones et al., 2008), revealing the importance of a better understanding of zoonotic emergences for public health.

Some researchers argue that many zoonotic pathogens cause either mild infections or severe disease, but rarely intermediate symptoms (Palmer et al., 1998; Frank and Jeffrey, 2001). The idea behind this reasoning is that most zoonotic pathogens are not very well

adapted to the human physiology and are therefore similar to commensal infections which rarely cause disease. Or, because of the missing adaptation, cross into host tissues where they cause life-threatening symptoms, for example the blood-brain barrier with a subsequent meningitis. The basis for this is the theoretical prediction that parasites adapt to minimise virulence while maximising the transmission potential (May and Anderson, 1983). However, more recent studies have shown that this theoretical assumption might not necessarily be true, because it is neglecting the possible trade-off between virulence and transmission potential of a given pathogen (Anderson and May, 1992; Boots and Sasaki, 1999, 2003). It is believed that evolutionary selection should favour higher transmissibility, but the exact relationship between virulence and transmission potential remains unknown.

Furthermore, zoonotic infectious disease can come from a wide range of taxa and impose different levels of virulence. In the following, I will present some examples of well-known zoonotic diseases:

- Anthrax is a bacterial infection of *Bacillus anthracis*, common in a wide range of species. In humans, it typically arises from carcasses of diseased herbivore mammals (Frank and Jeffrey, 2001). While the risk of human-to-human transmission is very low (Keeling and Rohani, 2007), recent attention has been drawn to its possible use as a weapon of biological warfare for terrorists (Wein et al., 2003), which raised the public's awareness of this zoonotic pathogen. Public attention was further raised by 34 cases of an anthrax infection with seven deaths within a Scottish heroin user community in 2010 (DEFRA - Department for Environment, Food and Rural Affairs, 2011).

- Ebola virus is one of the most notorious zoonotic diseases due to its high mortality within humans. It causes a rapid onset of haemorrhagic fever and has a case fatality rate of 0.25 - 0.9 (World Health Organization, 2011). It has been argued that the need for very close contact and the severity of the disease prevent the infectious disease from emergence (Keeling and Rohani, 2007). Previous outbreaks occurred in rural parts of Africa. Therefore, it remains unknown if an outbreak can be contained once it reaches areas with high population density or transportation hubs. Leroy et al.

(2005) and Biek et al. (2006) identified an animal reservoir of the virus in central African fruit bats with the possibility of more, still unknown reservoirs.

- In the United Kingdom, campylobacter infection is currently the most common zoonotic disease. Originally an infectious disease of many animal species, its 18 zoonotic outbreaks caused over 70,000 laboratory confirmed human cases in 2010 and are believed to have originated in poultry farms (DEFRA - Department for Environment, Food and Rural Affairs, 2011). However, DEFRA - Department for Environment, Food and Rural Affairs (2011)'s case estimates predict that the actual number of cases might have been up to ten times higher. The 'faecal-oral' transmission route within the human population allows for control of the infectious disease through increased hygiene within livestock and human communities.

- Highly pathogenic influenza A(H5N1), also known as 'bird flu', has recently received lots of public attention. Its natural host reservoir is in poultry and waterfowl with zoonotic episodes of very limited outbreak size in the human population (Beigel et al., 2005; Abdel-Ghafar et al., 2008).

Though all these diseases are of zoonotic origin, they do not transmit well between humans. Nevertheless, the risks of a zoonotic emergence are severe. Released into a naïve human population, a novel pathogen poses a great risk to human health. Possible scenarios include pandemics without adequate medical treatment such as antiviral therapies or vaccines. History tells us that such events are rare but possible.

The most prominent example is the 'Spanish flu' pandemic of 1918. Neither its origins nor its epidemiological features are clearly known (Taubenberger and Morens, 2006). It probably originated in birds (Reid et al., 1999), and Taubenberger and Morens (2006) estimate that over one third of the world population, equalling approximately 500 million people, were infected with clinically apparent symptoms. An unusually high case fatality rate of more than 10% has been estimated for the Spanish flu which caused approximately 50 million illness related deaths (Johnson and Mueller, 2002).

While much less virulent, the subsequent influenza pandemics of 1957, known as 'Asian flu'; 1968, known as 'Hong Kong flu'; and 1977, known as 'Russian flu', were all potent re-

minders of the capacity of the influenza virus to cross the species barrier into humans and cause a severe threat to human health (Horimoto and Kawaoka, 2005). The 'Swine flu' pandemic of 2009, being the most recent influenza pandemic, is believed to have evolved out of several strains circulating in swine (Dawood et al., 2009). Smith et al. (2009) show that the pandemic influenza strain is a triple-reassortant virus, whose eight genomic segments can be followed back to genome samples from the 1970's. Some genome segments originated from avian influenza A(H1N1), some from classic swine influenza A(H1N1), and others from seasonal, human influenza A(H3N2). While estimates suggest that it was not as virulent as the 1918 influenza pandemic (Fraser et al., 2009), it still demonstrates the severe danger of zoonotic diseases to human health, because even less virulent influenza pandemics lead to large numbers of casualties and great stress on the public health systems just by the sheer impact of the large number of infected cases. Furthermore, its prevalence reminds us of what harm the next influenza pandemic might inflict if its virulence were comparable to that of the 1918 pandemic.

Many other pathogens share this pandemic potential: the SARS coronavirus outbreak of 2003 (Lipsitch et al., 2003; Brockmann et al., 2005; Skowronski et al., 2005) has been linked to bats and palm civets (Guan et al., 2003; Li et al., 2005). While SARS did not cause a pandemic, it is believed that only the rapid worldwide response with isolation of infected human cases prevented a pandemic (World Health Organization, 2006).

In the rest of this chapter, I will concentrate on the barriers which might hinder a zoonotic pathogen from emerging into the human population. I present key challenges a pathogen has to overcome, as well as additional factors influencing the emergence of a zoonotic infectious disease. Specifically, I ask: What are the necessary steps for a species jump? What are the underlying factors influencing an emergence? What are the necessary adaptations to human physiology? What are the possible effects of human host heterogeneity? Which factors have an influence on the development of EIDs? I will start with general considerations about species jumps, followed by more detailed insights into the evolution of pathogens.

1.2 Characterisation of a Successful Emergence

1.2.1 Key Conditions for a Zoonotic Emergence

The invasion of the human, recipient species is a challenging task for a zoonotic pathogen. For example, influenza A(H5N1) preferentially binds to α2,3-linked sialic acid receptors found on cells in the gastrointestinal tract of waterfowl, while it is believed that A(H5N1) needs to be able to bind to α2,6-linked sialic acid receptors in the human repository tract for sustained human-to-human transmission (Olofsson et al., 2005; Kuiken et al., 2006; Abdel-Ghafar et al., 2008). This missing but probably necessary adaptation could explain why influenza A(H5N1) infections in humans have not led to a widespread epidemic. Furthermore, two research groups from the Netherlands and Japan have allegedly been able to identify four mutations in the A(H5N1)'s haemagglutinin, which would allow for sustained transmission between ferrets and probably humans. However, the publication of these research results is yet under embargo due to possible biosecurity risks.

This example illustrates one of the problems a pathogen has to overcome on its way to emerge into the human host species. Overall, I classify four different so-called 'key conditions' which cover all possible problems a zoonotic infectious agent has to master on its way to emergence:

(1) Exposure of the (human) recipient species to the pathogen. This is the first step necessary for a successful emergence into the recipient species. The pathogen needs to be present in the donor species, and a contact suitable for transmission between a diseased individual of the donor species and a recipient host must be established (Kuiken et al., 2006). Environmental and ecological influences have a large impact on the strength of the exposure (McMichael, 2004; Weiss and McMichael, 2004), and changes in these conditions might cause a novel pathogen to invade the human population (Cohen, 2000). These changes in environmental and ecological conditions might happen in the human population, but also in the pathogen's animal reservoir. There, a change towards higher prevalence usually increases chances of cross-species transmission (Parrish et al., 2008).

(2) Ability to infect the (human) recipient species. The second step after exposure consists of the pathogen's actual ability to infect an individual of the recipient species and cause outbreaks of limited size. For this, any pathogen needs to be sufficiently adapted to the recipient host physiology to successfully enter and replicate within the novel host, but also to leave the host and cause secondary human infections (Kuiken et al., 2006; Parrish et al., 2008). In some cases, these processes might happen within the recipient, human host population - in other cases, these adaptations might have already happened in the donor species (Holmes, 2009a). Especially the later variant is dangerous from a public health perspective as it offers little warning in terms of failed introductions. Overall, this condition corresponds to the second and third of Wolfe et al. (2007)'s five stages, which describe the path from a pathogen exclusively infecting animals towards a human-only pathogen. Stage two defines the sporadic infections of humans exclusively from an animal source, and it separates exclusive infectious diseases in animals from the ones which actually cross the species barrier. The third stages describes pathogens that cross into the human population, but generate only outbreaks of limited size.

(3) Sufficient transmission in (human) recipient species. Once the pathogen has actually infected a recipient host it must overcome the challenge of sufficient human-to-human transmission. This is the bottleneck for any kind of emerging disease, and a failure in sustained transmission can be caused by insufficient adaptations. A pathogen might already be well adapted to the human host and capable of sustained transmission at its first encounter like in the previous condition, develop this ability while causing limited outbreaks, or never cause sustained transmission (Holmes, 2009a). But also environmental conditions like host demographics might hinder a perfectly adapted pathogen from emerging (May et al., 2001; Newman, 2002).

(4) Chance, or the effect of stochasticity. This is probably the most abstract condition. It is not a challenge like the others but rather the notion that all of the above conditions are stochastic processes. For example, a pathogen might be well adapted and environmental conditions allow for emergence, but it goes extinct at the beginning of the

emergence in the human population just by pure chance (May et al., 2001; Antia et al., 2003; Woolhouse et al., 2005). Since the zoonotic transmission and first emergence into humans will almost always affect a very small number of people, stochastic effects cannot be ignored.

For a successful emergence to occur a zoonotic pathogen must pass each of these conditions. If a pathogen fails in any of these four, it will ultimately not emerge into the recipient host population.

1.2.2 Main Factors Influencing an Emergence

The possibility and likelihood of fulfilling each of the four key conditions depends on many underlying factors able to influence the zoonotic emergence of a pathogen. For example, a change in climate might lead to a range expansion of zoonotic, mosquito-born diseases (Epstein, 2001).

Cohen (2000) gives an excellent overview of six main factors leading to any kind of EID, i.e. drug-resistant EIDs as well as zoonotic EIDs. Using the six main factors, I extract the relevant ones for zoonotic EIDs. All these can be ordered into two broad categories, making up my two main factors:

(i) **Environment** is referring to everything surrounding the emergence, and concentrates on all processes outside a host. It contains the host interactions between and within donor and recipient species. Therefore, it is influenced by many factors such as behaviour, ecology, and demographics. Even public health measures, or the lack of these, play into this main factor as they have a great influence on the interactions once a possible emergence threat is identified. SARS is a prominent example of the impact of public health interventions on EIDs, because some argue that only the fast implementation of quarantining infectious hosts and the eradication of poultry markets in Hong Kong prevented a pandemic (World Health Organization, 2006).

(ii) **Adaptations and change of pathogen** includes all factors influencing within-host processes. This factor is sub-divided into **a)** all evolutionary mutations happening in

the donor species (pre-jump) which are beneficial to the emergence process. Such mutations can be beneficial or neutral to the fitness of the pathogen in the donor species, but also deleterious under certain conditions (Holmes, 2009a). It also contains **b)** all evolutionary mutations happening in the human species (post-jump) which are beneficial in the emergence process. These mutations can only be beneficial to the fitness of the pathogen in the recipient species as they would otherwise not be beneficial for emergence (Pepin et al., 2010).

It is important to separate the second factor into pre-jump and post-jump even if the effect, mutations with a beneficial impact on emergence, is the same. But the outcome of pre-jump adaptations might allow for an infectious agent which generates an emergence 'out of nowhere', while a pathogen adapting post-jump might give warning with limited outbreaks. Each of these two main factors or a combination of them can play a role in overcoming the four key conditions. Interestingly, Cohen (2000) points out that EIDs are often the consequence of change:

> "The recurring theme throughout all of these factors that influence the emergence of infectious diseases is change. During the twentieth century, tremendous societal and technological changes occurred, and these are likely to continue or accelerate in the twenty-first century. It is useful to examine how these changes have influenced infectious disease emergence so that we can anticipate some of the challenges of the new century."

Obviously, evolutionary mutations are changes by definition. But other, sometimes slow changes might have an impact on the occurrence of zoonotic EIDs. For example, changes in world-wide travel towards travelling longer distances in shorter time, influencing the human within-species interactions, greatly increase the possible spread of any infectious disease (Brockmann et al., 2005).

While change in ecology and environment can be the trigger for an emergence (see condition (1) and factor (i)), the pathogen's infectiousness to humans remains an absolute constraint (see condition (2) and factor (ii)). However, even if capable of infecting the human host species, zoonotic pathogens are usually, although not always, significantly less infectious to humans compared to their animal reservoir (Woolhouse et al., 2005). To overcome this barrier can be substantial, implying that much higher doses might be required

to infect the human host. The rabies virus is an example of such a process in non-human species jumps. The dose of virions from foxes necessary to infect cats and dogs has been shown to be up to a million times greater than that required to infect other foxes (Woolhouse et al., 2005).

The next step is sustained human-to-human transmission (see condition (3) and factor (ii)) as an increase in incidence is the defining element of an emergence. Scientific literature offers insights into some zoonotic infectious agents which did not make this step. For example, five cases, four of them fatal, of a new arenavirus strain, the Lujo virus, were detected in South Africa, causing great concern in 2008 (Briese et al., 2009). It was able to enter the human recipient species and produced secondary human cases. Nonetheless, it was not able to sufficiently transmit between humans to cause a sustained chain of transmissions. The virus might not have been sufficiently adapted, the environment of the human host population might have prevented subsequent transmission, or stochastic effects (see condition (4)) might have concealed an inherent ability for sustained transmission.

Overall, the process of emergence arises from complex interactions between pathogens, hosts, and the environment they share. The next section focuses on the environmental and ecological aspects of these interactions.

1.3 The Role of the Environment

All factors influencing a zoonotic emergence can be divided into two broad categories, as shown in the previous section 1.2.2. The first main factor, the environment, is the totality of everything excluding the pathogen and its adaptations. It contains the ecology of the donor and recipient species, as well as the interaction between both. Host demography, social behaviour, or spatial community structure influence the environment. A change in environmental conditions could be a necessary condition for emergence, but it might also hinder the chances of cross-species jumps. Overall, change seems to play a key role in the emergence of novel diseases (see section 1.2.2 and (Cohen, 2000)) as it might increase the likelihood of emergence (Epstein, 2001; Keesing et al., 2006, 2010; Hoskisson and Trevors, 2010).

In general, the cross-species jump of a pathogen partly resembles the ecological invasion of a novel patch. In this ecological context, researchers often talk about the 'spillover' of pathogens into the recipient species without subsequent emergence, which is often described as 'source to sink' migration. Obviously, a spillover can only occur if a connection between different species has been established. Unfortunately, experimental evidence of the environment's influence on such between-species connections is rare. Studying ecological interactions between two species of mammals under controlled laboratory conditions requires a very large-scaled set-up, which incurs high costs and maintenance. This problem can be overcome by using bacteria and phage as model systems (Dennehy, 2009). Their small size and fast generation-time allows for feasible experimental studies in a laboratory setting, which can give valuable insights into basic principals and test theoretical predictions. However, one needs to be aware of the possible shortcomings of bacterial model systems.

As a result, the majority of our knowledge about environmental influences on zoonotic infectious disease emergence arises from evaluating the conditions after a species jump into the human population has occurred. Here, I use some of this knowledge to subdivide the term 'environment' in relation to zoonotic EIDs into four sub-topics.

1.3.1 Agriculture, Technology, and Industry

Taking a look at human history, McMichael (2004) identifies past major transitions in the structure of human society as the basis for the subsequent spread of zoonotic EIDs. For example, the change from neolithic hunter-gatherers towards an agriculturally based society dramatically increased the contacts between livestock, rodents common in households, and humans. It is speculated that measles emerged about 7,000 years ago from Rinderpest of cattle as a result of the rise of agriculture (Weiss and McMichael, 2004). Thus, the agricultural revolution changed the within-species dynamics of the donor species, cattle, the recipient species, humans, and the interaction between both.

As this example shows, the impact of agriculture and its industrialisation should not be underestimated. Indeed, Woolhouse and Gowtage-Sequeria (2005) rank changes in land

use and agricultural industrialisation first in a list of drivers for emerging infectious diseases. Wolfe et al. (2007) argue that many 'crowd epidemic diseases' were only able to enter the human population because the advent of agriculture increased human food supply and allowed for larger settlements. Infectious diseases like measles, rubella, and pertussis need several hundred thousand people to naturally sustain itself in the human population. If the population size is lower, not enough new hosts are born to re-fuel the number of susceptibles. Only a single wave would be possible after which an infectious disease like Rubella would face extinction.

Agriculture has changed substantially since its introduction. It grew into a whole food industry and, together with technological advances in processing and preservation, allowed for a dramatic increase in *Escherichia coli O157:H7* infections (Cohen, 2000). Food poisoning with zoonotic infectious agents has been rising widely over the last two decades (McMichael, 2001). This is in line with observations from the United Kingdom where the two most common zoonotic infectious diseases in 2010, salmonella and campylobacter, are acquired through the food-chain (DEFRA - Department for Environment, Food and Rural Affairs, 2011).

The change of land use in Asia is another example of industrial influences: after the second world-war, the increase of rice production towards industrial sized farming facilitated an increase in Hantaan virus carrying rodent, corresponding to an increase of zoonotic incidence in humans (Cliff et al., 2009). In the UK, BSE, also known as 'mad cow disease', is probably one of the most notorious examples of agriculture's influence on zoonotic infections. It had a significant impact on the agricultural industry and raised fears that it might be persistent in multiple species (Kao et al., 2002). Only the industrialised cannibalism, whereby neural tissue and bone meal from slaughtered cattle were fed back to cattle, allowed for a massive epidemic with infectious prions, and subsequent human zoonotic infections with vCJD (Weiss and McMichael, 2004).

Other areas of technology facilitated cross-species jumps in different ways. The widespread use of cooling towers, whirlpools, and showers provides temperatures that promote the survival and proliferation of *Legionella pneumophila*, a pathogen unknown before its first zoonotic outbreaks in the 1970s (Cliff et al., 2009). Technological advances

in medicine and public health aid the emergence of infectious diseases in a different way (Smolinski et al., 2003): medical therapies produced from animal-cell substrates can be a vital danger to human health. Syringe and needle are the most efficient 'medical vector' as the propagation of the SV40 virus through polio vaccine demonstrated (Weiss and McMichael, 2004).

But changes in technology and industrialisation should not solely be seen as a threat. Technological developments yield immeasurable benefits in containing zoonotic pathogens. For example, improvements in municipal water systems had a dramatic effect on cholera infections (Cliff et al., 2009), and in medicine, the development of antibiotics and antiviral drugs serve to control and prevent zoonotic pathogens.

1.3.2 Host Demographics and Behaviour

Not only new technologies and the rise or fall of industries change the environment. We, the human hosts, change the environment for pathogens with our behaviour and the structure of our society. Change in human demography and society is in fact the number two cause for emerging and re-emerging infectious diseases (Woolhouse and Gowtage-Sequeria, 2005). Ageing, for example, often contributes to susceptibility to certain diseases (Smolinski et al., 2003). Seasonal influenza is associated with higher incidence in the elderly (Molinari et al., 2007), and it is estimated that over a quarter of the US population will be older than 65 years by 2040 (Cohen, 2000), which is a five-fold increase compared to the beginning of the twentieth century. Shifting the age distribution might not only change the prevalence of known infectious agents, but also be the basis for a successful emergence of a zoonotic infectious disease. Two-income families become more likely, increasing the use of day care facilities or kindergartens. These places are ideal hotspots for infectious disease transmission, and might aid the emergence process of a zoonotic pathogen.

Host demographics are only one part of the changing environment. Our behaviour might facilitate transmission in another way. The use of drugs can open new transmission routes for pathogens, and changes in sexual behaviour might make the difference from unsustainable to sustainable human-to-human transmission. HIV is probably the

best example for an originally zoonotic infectious disease, which emerged into the human population and is now endemic in many parts of the world (Schrag and Wiener, 1995). The virus' main route of transmission is through sexual contact, or, especially in the developed countries, drug users sharing needles and injection equipment (Anderson and May, 1992). It is believed that social and economic factors led to a change in sexual behaviour towards multiple sexual partners in Africa (Ezzell, 2000), which allowed for the widespread distribution of the infectious disease. Based on this notion, behavioural change is one of the most promising measurements to reduce risk of HIV exposure (Prochaska et al., 1994).

1.3.3 Travel and Trade

Human behaviour also changed travel and trade. Historical evidence from the time of the imperialist expansion ca. 500 years ago shows the devastating impact of European expansion into new regions. The Spanish arrival in South America led to severe epidemics within the native population, and is often connected with the fall of the Inca empire (McMichael, 2004). By estimating the origin of human infectious diseases, Wolfe et al. (2007) show that a majority of human pathogens originated in the Old World and were introduced into the Americas much later. However, this was not a one-way street. Subsequent trade between the contingents introduced chagas disease, a tropical parasitic disease infecting mammals and humans through Reduviid bugs, to Europe.

Unsurprisingly, travel and trade are the seventh and ninth most important driver for emerging infectious diseases (Woolhouse and Gowtage-Sequeria, 2005). In today's world, millions of people fly around the globe. Cars allow us to travel distances in hours which took us days a hundred years ago. This has dramatic effects on disease spread. For example, the 'Black Death', a pandemic of the bacterium *Yersinia pestis*, took from 1348 till 1350 to sweep from Italy to Sweden (Hays, 2005). In comparison, the recent 'Swine Flu' pandemic took less than a month to reach the United Kingdom from Mexico (Dawood et al., 2009; Health Protection Agency, 2009c). The West Nile virus was unknown in North America and originated, as the name suggests, in Africa with sporadic episodes in the Middle East and Europe (Cohen, 2000). It has been suggested that aeroplanes unwittingly

transported the pathogen to New York and caused an outbreak in 1999 (McMichael, 2004). Today, the virus is endemic in the USA with animal reservoirs in birds and horses (Ezenwa et al., 2006).

The examples above show that human mobility drastically increased over the last few hundred years. Indeed, estimates suggest that spatial mobility of the average human increased 1000-fold since the year 1800 (Smolinski et al., 2003). An infectious host can board an aeroplane in Europe and arrive infectious in South America, scaling a new outbreak to a worldwide problem in less than a day. SARS has been an example for such a behaviour, where the pathogen was transported from Asia to North America in a short time period. In such a case, it is necessary to evaluate all contacts every infectious host had in an attempt to break the transmission chain. Using data on social contacts between hosts, Meyers et al. (2005) and Lloyd-Smith et al. (2005) show that so-called 'superspreaders' had an important influence on the SARS outbreak in 2002. These individual hosts cause a disproportionately high number of secondary cases and can shape the overall behaviour of an epidemic. It is likely that they were driving the SARS outbreak, and the majority of infected people caused less than one secondary case on average (Meyers et al., 2005).

Data gathering is a demanding but important task to shed light on human host contact structures. Brockmann et al. (2006) present an influential study of human travel patterns by following the path of dollar bills. They show that human travel follows a power-law, which can be modelled with scale-free random walks, known as Lévy flights. Power-laws in human movement have the interesting property that the majority of individuals travel only short distances, while a small number of individuals travel very far. For example, over 50% of all dollar bills travelled less than ten kilometres in 14 days, while approximately 7% of all bills moved 800 kilometres within the two weeks (Brockmann et al., 2006). Assuming that human travel pattern resembles the movement pattern of dollar bills, individuals can transport pathogens over a long distance in a short time. Other studies show similar results using aviation passenger data (Hufnagel et al., 2004; Colizza et al., 2006; Cliff et al., 2009), commuting data from the USA (Balcan et al., 2009), or contact tracing methods (Kiss et al., 2006).

1.3.4 Ecological Changes

The change in human mobility, advances in technology, and the resulting changes in behaviour can also be seen in the spatial distribution of human communities. It is believed that the first settlements occurred with the start of agriculture and allowed for 'crowd diseases' (McMichael, 2004), as described in section 1.3.1. Some of these settlements evolved into larger towns of which some became big cities. It is estimated that around the year 1800 roughly 5% of the human population lived in an urban environment. Today's urbanisation leads to ever bigger cities with an estimated 50% of the population in urban areas (Smolinski et al., 2003; McMichael, 2004), significantly changing the spatial distribution of human hosts. Furthermore, this figure is expected to rise to 65% by 2030 compared to the average community sizes of 100 - 200 human hosts in pre-agricultural times (McMichael, 2004).

Not only the urbanisation, but how the growth happens is another matter of concern for public health. Estimates suggest that migration from rural areas contributes approximately 40% of urban growth (Smolinski et al., 2003). Rural-to-urban migration, especially in cities without adequate infrastructure, yields several risks for infectious disease emergence. Most often, recent arrivals live in very dense areas without access to health services and inadequate supply of clean water (Smolinski et al., 2003). Furthermore, the migration itself might transport a zoonotic pathogen from its rural source into a crowded city with conditions favourable for its emergence.

Dengue virus is an example of an infectious agent with vector-borne transmission with zoonotic cycles between humans and primates in tropical Africa (Monath, 1994). Over time, dengue's vector species, *Aedes aegypti*, adapted to colder climates and followed humankind into urban areas (Rogers et al., 2006). Now, dengue is the most prominent vector-borne viral disease in humans with an estimated 80 million cases per year (McMichael, 2004).

Urbanisation and resulting changes in spatial heterogeneity change the ecology of the human host population. However, these changes might not always be for the worse. For example, it has been estimated that the urbanisation trend in Africa tends to decrease

malaria transmission and severe disease (Hay et al., 2005). Overall, humankind's impact on animal or plant ecological systems is manifold. Climate change, deforestation, and pollution have a dramatic effect on many ecosystems (McMichael, 2004; Cliff et al., 2009). Like with *Aedes aegypti*, it might bring the natural habitat of vectors or rodents nearer to human hosts. Or it might even drive whole species extinct. This loss of species diversity has a potentially increasing or decreasing effect on infectious disease transmission. For example, it appears that lower between-species than within-species transmission is a necessary condition for lowered host diversity to reduce disease risk (Keesing et al., 2006), and urbanisation might increase the between-species contacts significantly. On the other hand, the scarcity of a species facilitating pathogen transmission will lower the prevalence. Conversely, if this species reduces the competition for a disease-carrying vector, hence increasing the number of pathogen-carrying vectors, the prevalence of the disease will increase. Case studies with hantaviruses in rodents show a clear evidence of this effect (Keesing et al., 2010). By experimentally removing non-host species, the mean density of seropositive rodents increased significantly. Such an effect has also been shown in zoonotic disease transmission. Case numbers of human hosts infected with the West Nile virus have been shown to correlate negatively with avian species diversity (Ezenwa et al., 2006).

Nonetheless, the impact of species diversity on zoonotic disease emergence can also be of an enhancing nature. Results with plant pathogens suggest an opposite effect of biodiversity. Plant communities lacking the species *Avena fatua* show less prevalence of the Barley Yellow Dwarf virus than very diverse communities with higher *Avena fatua* densities despite not being the sole disease transmitting species (Power and Mitchell, 2004). Jones et al. (2008) estimate a small, but significant positive correlation between host species richness and the probability of zoonotic emergence from a wildlife species. One explanation might be that host species richness has a direct impact on the pathogen species richness. Decreasing species richness might reduce the number of host reservoirs, which are fuelling spillover infections, or even reduce the overall number of pathogen species, which might go extinct with their respective host.

This duality in the effect of biodiversity has also been reported using a solely theoretical approach (Dobson, 2004). Models describing infectious disease spread with density-

dependent transmission yield that an increase in species diversity will always lead to an increased transmission potential, because the number of contacts between infectious and susceptible hosts will always increase. In contrast, models with frequency-dependent transmission dynamics result in reduced within-species transmission as the overall number of transmissions gets split between within- and between-species transmissions. A general result from theoretical epidemiology is that decreased transmission reduces the probability for outbreaks to occur (May et al., 2001). In such a case, increased species diversity would decrease disease establishment.

Biodiversity shows well how ecological changes can have enhancing or suppressing effects on zoonotic EIDs. But also changes to re-instate previous environments might facilitate zoonotic pathogens. The emergence of the lyme disease in the midwestern USA has been linked to reforestation programs of former farm land (Smolinski et al., 2003). Having its natural reservoir in rodents and deer, the infectious disease gets transmitted by a tick vector (Hartfield et al., 2011). Through the reforestation, white-tail deer populations increased significantly and, through the tick vector, spread the disease into the human population.

All the above examples show the environment's effects on the cross-species jumps of pathogens, and its chances to emerge in the human population. But it is only one part in the complex interaction between hosts, pathogens and environment. Another part is the interactions of pathogen and host, which are usually within-host processes such as adaptations to the novel host physiology. The following section discusses barriers a pathogen needs to overcome, and the mechanisms which allow for adaptations.

1.4 Pathogen Evolution and Adaptation

This section deals with the within-host processes, which can facilitate the emergence of a zoonotic infectious disease. Zoonotic pathogens are sometimes insufficiently adjusted to the novel recipient species to cause an emergence without further evolutionary adaptations. Indeed, Dennehy et al. (2010) are able to show the impact of the difference in recipient and donor species on the necessary adaptations by using bacteria and phage as

model system. They alter the genetic difference between the donor and recipient bacteria species, and show that pathogens need to undergo larger adaptations, expressed in relative fitness within the other species, when the genetic difference between the species becomes larger. These genetic adaptations are of highest interest to researchers, as they are often seen as the key for improved surveillance and risk assessment of potential EIDs. This also sparked the idea of 'genetic markers', which could allow us to assess the risk of zoonotic emergence based on the pathogen's genome (Pepin et al., 2010). As yet, this remains a distant but tantalising aspiration.

Although not all necessary adaptive steps for each zoonotic pathogen are known, and previous studies suggest that the within-host interplay of wildtype and mutant strains is complex (Iwasa et al., 2004), it is possible to define general molecular hurdles that pathogens must overcome in order to emerge in the human population (see challenge (2)). Here, I present some of the obstacles of pathogen adaptation in detail by using zoonotic viruses, the largest taxa of emerging pathogens (Woolhouse and Gowtage-Sequeria, 2005), as examples.

In general, virus emergence can be restricted at many different levels, including receptor binding, entry or fusion, trafficking within the cell, genome replication, and gene expression (Parrish et al., 2008). While it may matter if evolutionary adaptations need to happen in a specific order or not (Traulsen et al., 2007; Gokhale et al., 2009), it does not change the basic biological processes I will present here.

1.4.1 Necessary Steps for Adaptation

In general terms, a virus needs to accomplish five steps to actually infect the human host population (Webby et al., 2004a; Holmes, 2009a): (a) infection of an appropriate host cell, (b) replication inside the host cell utilising host factors, (c) exit from the host cell, (d) sufficient evasion of the host immune system, and (e) exit from the host and transmission to another. It is obvious from the number of zoonotic viruses that succeeded in steps (a - d), but not (e), that the adaptive changes necessary for a zoonotic virus to replicate in the human host are independent of, but necessary for, those required for successful and

sustained human-to-human transmission. Especially RNA viruses are a constant cause of concern (Holmes, 2009b): the Simian Foamy virus, influenza A(H5N1), and Lujo virus have all crossed the inter-species barrier from animals to humans, but have yet to achieve the next step of sustained human-to-human transmission. By contrast, influenza A(H1N1) and HIV are examples of zoonotic RNA viruses that have recently emerged and become endemic in humans.

1.4.1.1 Infection of an Appropriate Host Cell

Viruses are intracellular parasites; the need to enter an appropriate host cell is a prerequisite to replication. For this, they must breach entry barriers, such as mucus, alveolar macrophages or epithelium, and find their way to tissues in which they can replicate. For example, chimpanzees are comparatively resistant to experimental respiratory exposure to human influenza strains (Olofsson et al., 2005). A possible explanation is the existence of special virus binding mucins in the muscus of the respiratory tract.

Sometimes infection is systemic, incorporating multiple tissues in the host body. In other cases, it can be restricted to certain tissue and therefore body parts. For example, an influenza virus usually enters the human host through the respiratory tract with the need for an internal transport system to spread from there to other susceptible tissues. Thus, it needs to enter the blood or lymph system, be successfully transported, and exit at appropriate tissue-blood/lymph junctions (Smith and Sweet, 1988). For example, it depends on the amino acid sequence at the cleavage site of the precursor hemagglutinin if an infection is localised or systemic in poultry, while viral factors yielding systemic infections in humans are less clear (Kuiken et al., 2006). Furthermore, the ability to enter appropriate tissue, or evolve novel tissue tropisms, is also necessary for other types of pathogens, i.e. parasites (Read and Skorping, 1995).

Once in appropriate tissues, the interaction of a virus protein with its corresponding cell receptors initially mediates the entry of a virus particle into the host cell. In many viruses, this virus-cell interaction is well understood and can be a primary determinant of host range (Webby et al., 2004a). This can be expressed with a rule-of-thumb: if the

appropriate receptor is present, the virus may replicate; if it is absent, the virus will not. For example, the emergence of influenza A viruses in the human host population seems, at least in part, to be mediated by cell entry interactions. The main reservoir, and therefore most likely donor species, of influenza A viruses is waterfowl from which viruses sporadically transmit to human hosts. Because avian influenza viruses preferentially bind to different host cell receptors, sialic acids, compared with influenza viruses of human origin, they need to adapt to humans to form stable linages. Hemagglutinin molecules on the viral coats of avian influenza viruses, which are gastrointestinal infections in waterfowl, such as A(H5N1), preferentially bind to α2,3-linked sialic acid receptors on the animal host cell membrane (Abdel-Ghafar et al., 2008). However, human influenza viruses bind to α2,6-linked sialic acid receptor, the predominant receptor for influenza in the human trachea. Given the original host environment of waterfowl, the α2,3-linked sialic acid preference makes biological sense as the gastrointestinal tract of waterfowl contains predominantly α2,3-linked rather than α2,6-linked sialic acid receptors.

This difference may explain why replication of avian influenza viruses in humans generally tends to be restricted. Furthermore, research on SARS has revealed similar cell receptor restriction. It has been shown that the susceptibility of a given cell type to infection correlates well with levels of a specific SARS receptor molecule using SARS pseudotyped HIV particles (Nie et al., 2004). Gaining the ability to bind to a new receptor can be a complex process, requiring multiple changes in the viral genome. The receptor binding motif of SARS, a corona virus, is largely missing from other coronaviruses common in bats. This motif may have been established by recombination with other coronaviruses in the animal reservoir and subsequent mutations (Li et al., 2006).

In general, viruses use various host molecules as receptors. Consequently, cell entry is very host-specific for uncommon binding sites, whereas some viruses can attach to numerous cell types in different species when the receptor is globally conserved. Arenaviruses are such a class of viruses with broad host range and cell tropism, using the widely conserved α-dystoglycan protein as binding site (Meyer et al., 2002). By contrast, coronaviruses enter the host cell by using a specific interaction of their viral spike protein with glycoproteins on the cell's surface (Holmes et al., 2001). This naturally limits coron-

aviruses to a restricted host range. Nonetheless, mutations might change the viral spike protein, allowing it to bind to novel glycoproteins, and increasing the possible host range to new species as it happened with the SARS coronavirus. Indeed, the viral spike protein of a mouse coronavirus can bind to non-mouse cells after alteration of just a few residues in the viral spike's amino-terminal region (Thackray and Holmes, 2004). As mentioned before, the infidelity of the replication process of many viruses (RNA viruses in particular with their error prone RNA-polymerase) could lead to the quick accumulation of the necessary adaptive mutations in an alternative host.

However, while receptor specificity is a necessary condition, it does not seem to be sufficient for host cell entry (Matrosovich et al., 2004). Thus, more factors might be involved than the simple rule of thumb for receptors suggests.

1.4.1.2 Replication Inside the Host Cell Utilising Host Factors

Penetration of an appropriate tissue and entry into a host cell is only the first challenge for an emerging virus. The next step is to successfully co-opt host cell processes for viral replication. The virus has to uncoat and transport its genetic material to the appropriate cellular compartment. It can be seen as a hostile take-over of the host cell's replication machinery to produce copies of its viral genome, the virion coating, and assemble everything to produce a new virion.

The limited genome size of viruses is a general barrier for coding capacity. As a result, most viruses, except a few larger DNA viruses, have to rely heavily on host cell mechanisms to successfully replicate. Thus, a substantial adaptive pressure acts on zoonotic viruses to effectively use the cell proteins provided by the novel host. If the required changes in the viral genome are minor and paired with a high mutation rate, such as typically seen in RNA viruses, then the chances are high that the virus will adapt and be able to utilise the within-cell mechanisms. On the other hand, the requirement for major changes or low mutation rates render the chances of adaptation minimal.

Baigent and McCauley (2003) illustrate this fact with avian influenza infecting mouse cells. While it is possible to infect mouse cells - the virus can bind to appropriate receptors

and is able to enter the host cell - the within-cell replication is not functional. Probably, the reason can be found in the viral polymerase protein PB2 which is different in avian and mammalian influenza, and experimentally changed avian influenza viruses seem to support this view. It has been shown that an influenza A(H5N1) virus with a mutated viral polymerase is able to infect mice with increased virulence and viral spread to extrarespiratory tissue (Hatta et al., 2001), a clear sign of functional reproduction within mice cells. Another example of increased replication by adaptation to host cells is the Hepatitis C virus infection and replication in human hepatoma cell lines. *In vitro* adaptation experiments revealed a particular sensitive region for mutations, which can significantly optimise the replication rate of the viral genome with a single amino acid substitution (Blight, 2000).

These are just two examples of how evolutionary adaptation can allow or improve replication in novel host cells. Similar to the mutations that can facilitate receptor usage and cell entry, the number and type of mutations to achieve within-cell replication might vary depending on virus and host. Nevertheless, this does not rule out that only a small number of amino acid substitutions are needed to successfully allow replication.

1.4.1.3 Exit from the Host Cell

The first two steps present the stages of a viral infection from host entry to replication within a host cell. Logically, the next step for a virus is to exit the cell to infect other host cells, or shed from the host and get transmitted to another host. While the need for a successful exit seems to be trivial, the actual mechanisms can be complex.

In hosts infected with influenza, the virions assembled within the host cells are bound to the cells' sialosaccharides by their hemagglutinin (Kuiken et al., 2006). Thus, the influenza virus needs a mechanism to detach the virions from the cell - viral neuraminidase. It cleaves, like 'scissors', the sialosaccharides on the cell surface. Consequently, influenza viruses share the same preferences in sialosaccharides in cell exit as in receptor binding. For example, neuraminidase on the avian influenza virion surface prefers α2,3-linked sialic acids, while human influenza's neuraminidase has an affinity for α2,6-linked sialic acids (Baigent and McCauley, 2003). Obviously, mutational changes might alter this preference,

allowing avian influenza viruses to increase host infectiousness in humans. However, it remains unclear how many molecular substitutions would be necessary, or how most other viruses overcome the cell exit in novel hosts.

HIV cell exit is another example. Freed (1998) gives an overview of this part of the pathogen's life cycle: after the viral RNA is synthesised from the viral DNA in the host cell's nucleus, it gets transported to the cytoplasm. Envelope glycoproteins of the virus, synthesised from the viral RNA, move to the plasma membrane via the secretory pathway. Viral Gag and Gag-pol polyproteins recruit two copies of the single-stranded viral RNA and assemble it into a complex Gag protein structure. In addition, the Gag polyproteins induce a bud and fill it with the Gag protein complex using the plasma membrane and envelope glycoproteins as future viral surface. Once budding is complete, the virion pinches off the host cell as an immature HIV. Shortly after, viral protease cleaves the Gag and Gag-pol polyprotein precursors to Gag and Pol proteins. This cleavage also condenses the core and matures the virion to an infectious agent.

1.4.1.4 Evasion of the Host Immune System

In the previous sections, I presented the obstacles for viruses in infecting novel hosts; namely host cell entry, replication in the cell and exit from the cell. What I did not mention is that a virus has to succeed in all these tasks while the host's immune system tries to clear the organism from the intruder. Janeway et al. (2005) give an excellent introduction into the immune system which I will use as basis for my review.

The immune system is highly complex and can be differentiated into two categories: the innate and the adaptive immune response. The innate immune system, which is also called non-specific immune system, is the evolutionary older part and consists of several layers of protection (Beck and Habicht, 1996). It is separated from the adaptive immune system in that it is the first immune response evoked, it is non-specific towards the pathogen it encounters, and it does not have any immunological memory. The actual anatomical barrier is usually counted as being the first layer of the innate immune system, which I describe in greater detail in 1.4.1.1, and the first obstacle a pathogen needs to overcome. The

innate immune system also includes the inflammatory response, the complement system, and some of the leukocytes as further layers. Phagocytes, for example, 'consume' other cells or pathogens. They remove host cells, which died either by an infection or apoptosis, or bacteria by wrapping its plasma membrane around the other cell until it is enveloped. Then, they use lysosomes filled with enzymes and acids to 'digest' the engulfed cell.

How can the immune system correctly identify and clear foreign intruders while not attacking the host's own, non-infected cells? It does so by dividing cells and micro-organisms into three groups: self, non-self, and missing-self. Self-cells are all the host's own healthy ones. Conversely, non-self-cells are of foreign origin. Missing-self cells are the host's own cells, but in an infected state. The immune system uses chemical compounds to differentiate cells between these groups. The innate immune system is usually triggered when pathogens are identified by special pattern recognition receptors as non-self, which recognise components that are conserved among broad groups of pathogens, or when infected cells use chemical alarm signals. For example, natural killer cells react to low levels of the cell-surface marker human leukocyte antigen (HLA) - a situation of missing-self - which can arise in viral infections of the host cell. But instead of targeting the invader directly, a natural killer cell induces an apoptosis in the infected cell to interfere with the reproduction cycle of the pathogen.

One of the most prominent classes of such signal molecules of pathogen origin are antibody generators (antigens). These chemical substances binds to specific immune receptors, similar to receptor interaction at cell entry, and are important for the adaptive immune response. In general, antigens are differentiated into exogenus antigens, which come from extracellular pathogens like bacteria, and endogenous antigens, which are produced by intracellular pathogens such as viruses replicating within a host cell. Dendritic cells, which are part of the innate immune system, have the main function of processing exogenus antigen material and presenting it on their surface, being a messenger between the innate and adaptive immune system. Similar to the innate immune response, the adaptive immune response relies on the capacity to distinguish between foreign pathogens and the host's own cells. Host's cells produce and present self-antigens on the cell's membrane, which are different to the foreign non-self-antigens on pathogenic bacteria, or the endogenous

viral antigens presented on infected host cells (missing-self).

The adaptive immune system's cells are a type of leukocytes, the lymphocytes with the major types B-cells and T-cells. It is the evolutionary younger part of the immune system and gets activated after the innate immune system encountered an invader (Litman, 1996). Lymphocytes circulate in the peripheral blood and the lymphatic system. They use antigens to differentiate between self, non-self, and missing-self. While all nucleated cells have the capability to present antigen on their surface and trigger an adaptive immune response, there are some specialised antigen-presenting cells such as dendritic cells. Their main function is to present the exogenous antigens to CD4+ helper T-cells in the lymph nodes by coupling them to the HLA.

The endogenous antigen presentation, which is the relevant one for viral infections, works in a different way. Human host cells have enzymes specialised in presenting proteins on the cell surface by coupling them to HLA. This is a form of missing-self and alerts CD8+ cytotoxic T-cells (CTLs), also known as 'killer' T-cells. Naïve CTLs get activated if they encounter such an endogenous antigen/HLA complex. It is a primer for the CTLs to start rapid replication with the progeny specifically binding to this specific antigen/HLA complex. Once bound to it, the CTL release cytotoxins which lyse the host cell, destroying it and stopping the viral replication. This process is mediated by CD4+ 'helper' T-cells which provide stimulus to CTLs.

Furthermore, CD4+ T-cells activate macrophages and B-cells. B-cells are a major component of the adaptive immune response. In their naïve form, B-cells get triggered by a specific foreign antigen once they encounter it. Thereby, they engulf the antigen, digest it, and present it as an antigen/HLA complex on their surface. If these now specialised B-cells encounter matured CD4+ T-cells, primed by the same foreign antibody, they start to rapidly multiply and mature into plasma cells. These are generally short-lived, but produce large amounts of antigen specific antibodies, which are glycoproteins used to neutralise the pathogen.

The specificity of the immune response is one of the differences between the innate and adaptive immune system. Another one is the capability to form an immunological memory. Approximately 10% of T-cells and B-cells (matured to plasma cells) convert to

so-called memory cells. These cells are long-lived and stay primed to the specified foreign antigen, dramatically shortening the response time in case of re-infection with the same pathogen. These cells are responsible for host immunity, and make vaccines possible.

However, an immune response to a viral infection already starts within the infected host cell. Many cell types have a built-in antiviral response mechanism: interferon (INF), which was first discovered in embryonated chicken eggs by Isaacs and Lindenmann (1957). In humans, INF has been classified into two categories (with an ongoing discussion about a possible third type) based on the type of receptor which they use (Webby et al., 2004a): type I binds to INF-α and INF-β receptors on the cell surface, and is expressed by most cell types in response to viral infections. It makes several contributions to the defence against viral infections. First, it induces a state of resistance to viral replication within the infected and neighbouring cells. By binding to the receptors, the transcription of several host-cell proteins is triggered which interfere with the viral replication. Second, it facilitates the adaptive immune response by its induced increase of HLA to make the cell more visible to CD4+ T-cells, and increases the visibility to natural killer cells, although the underlying process remains unclear as natural killer cells belong to the non-specific innate immune response (Janeway et al., 2005). Type II, another protein which binds to INF-γ receptors, is mainly expressed in some T-cells and natural killer cells to activate inflammatory responses within the host with rather low antiviral effects.

Experiments on INF type I expression with some natural and inactivated viruses show that viral replication within the host cell induces INF production, but components expressed during viral replication actively try to repress the production of INF (Isaacs and Burke, 1958). Other viruses interfere with INF signalling capabilities, and yet others inhibit the functionality of INF-induced proteins (Webby et al., 2004a). Indeed, it is only logical to assume that viruses try to adapt to and overcome the host immune system and antagonise INF responses. If the virus has effective mechanisms to interfere and inhibit INF host responses, it gains the upper hand and its replication ensues. On the other hand, if the virus is not able to inhibit INF production and release, the immune system might repress viral replication and the infection will get cleared out of the host.

Mechanisms to counter host immune responses have been described for a variety of

viral EIDs. The influenza virus produces the viral NS1 polypeptide to act as an INF antagonist by sequestration of double-stranded RNA and inhibition of host mechanisms to process post-transcriptional mRNAs (Baigent and McCauley, 2003). For influenza A(H5N1), it has also been shown that NS1 helps the virions to replicate in INF treated cell cultures (Seo et al., 2002), although it remains unclear if this is a general property for NS1 in influenza viruses. Other viruses use different systems to restrict the host's INF response. For example, the Nipah virus uses its V protein to disturb the subsequent antiviral response by changing the sub-cellular distribution of STAT1 or STAT2 complexes (Rodriguez et al., 2002). These complexes are proteins involved in downstream INF signal transduction, mediating cellular responses to INF type I and II. Probably the best known example for a successful adaptation to the human immune system by an EID is HIV. It not only evades the immune system, it specifically infects immune cells, mainly CD4+ T-cells (Janeway et al., 2005).

These examples describe the interaction of viruses with the human immune system, leading to the question of non-human immune systems which is especially important in zoonotic EIDs. Schultz et al. (2004) show that INF is used in most, if not all, vertebrates with sometimes large differences in the INF systems, especially between mammals and birds. Furthermore, some form of innate immune response is common even in unicellular organisms like bacteria (Beck and Habicht, 1996). On the other hand, the adaptive immune system evolved in early jawed vertebrates (Litman, 1996). Thus, the virus-immune system interactions are largely species-specific, presenting an essential adaptation step in the successful emergence of zoonotic viruses.

1.4.1.5 Exit from the Host and Transmission to Another

For a virus in a novel host, it is necessary to overcome all of the above steps. But from an epidemiological viewpoint, the bottleneck is transmission between hosts. The pathogen can only succeed within the new recipient host population if the virus is capable of sufficient between-host transmission. Therefore, the last step is exit from the original host and transmission to another. For example, if a pathogen is best suited to infect tissue in the

upper respiratory tract, such as influenza, its most efficient route of transmission would be airborne. In such a case, the pathogen would need to exit the previous host in a way that allows airborne transmission. While this requirement seems to be easy to fulfil for respiratory infections, the route of exit might be more complex for diseases targeting tissue without direct contact to the environment, such as the the nervous system.

1.4.2 Adaptation by Mutation, Recombination, or Reassortment

Viral diversity can be generated through three processes: mutation, recombination, and reassortment. All three mechanisms change the genome of the pathogen and can have an influence on the phenotype with possible changes to the pathogen's epidemiological properties. Figure 1.1 provides a schematic overview of these three mechanisms.

Furthermore, these three mechanisms have a direct influence on the processes called 'antigenic drift' and 'antigenic shift', most commonly known from influenza (Scholtissek, 1995). As described in section 1.4.1.4, some parts of the immune system recognise a virus through its antigens on the virion surface. For influenza viruses, these are hemagglutinin and neuroaminidase (Bouvier and Palese, 2008). Antigenic drift describes the continuous process of change in the antigenic profile with a continuous number of small changes in the genome over time, most likely caused by mutations. This process is usually associated with the seasonal change of influenza strains (Bouvier and Palese, 2008). On the other hand, antigenic shift describes an abrupt change in the antigenic profile, which usually occurs by the interaction of two different strains through reassortment or recombination. The result is a novel strain with a mixture of the ancestral surface antigens, causing sudden jumps in influenza strain evolution, and leading to a strain with pandemic potential. For example, the 'Asian flu' pandemic from 1957 arose through reassortment of different influenza strains (Scholtissek et al., 1978).

It is worth noting that the adaptive steps might happen in the donor species, the recipient species, or a third species, a so called 'mixing vessel'. Especially in combination with reassortment and recombination, mixing vessels might explain the sudden occurrence of novel, fully adapted viruses without any warning signs (Arinaminpathy and McLean,

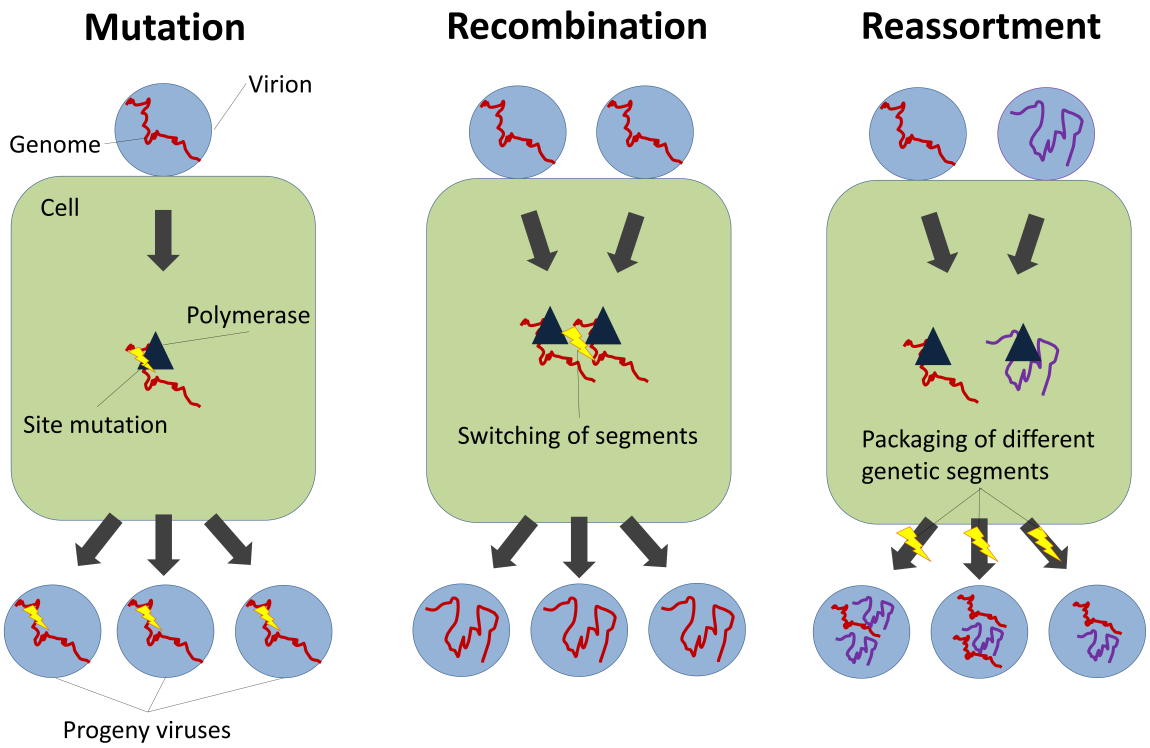

Figure 1.1: Different molecular processes for generating viral diversity. Viruses have three mechanisms to generate genome diversity at the replication stage. **Mutations** of single sites are a result of the lacking proof-reading capabilities of viral polymerase. Mutations can happen at every replication and are the most common cause of genomic changes in viruses (Webby et al., 2004a). **Recombinations** incorporate the genetic material of a foreign virion through mechanisms such as template switching. This type of genomic change is much less likely than mutations. **Reassortments**, which can occur during dual infection of a cell with two different segmented genome viruses, swap whole gene segments during the packaging stage of new virions. These three mechanisms are not exclusive and may result in new biological properties of the virus, which can help overcome the necessary adaptation steps.

2009).

1.4.2.1 Adaptation by Mutation

Although viruses have multiple defining characteristics, perhaps the most important from the perspective of EIDs is their capacity for mutation. Estimates of mutation rates in RNA viruses are in the range of $10^{-1} - 1$ mutations per genome and replication cycle (Duffy et al., 2008; Holmes, 2010). In comparison, bacteria usually exhibits a range of $10^{-4} - 10^{-3}$ mutations per genome and replication (Gago et al., 2009), several orders of magnitude below RNA viruses. Figure 1.2 gives an overview of mutation rates versus genome sizes for some biological entities.

For many viruses, the main source of mutation is the RNA polymerase (Duffy et al., 2008). It lacks the proof reading abilities usually common with DNA polymerase in bacteria or eukaryotes (Webby et al., 2004a), yielding such a remarkably high error rate. Factoring in that the majority of all mutations are not of beneficial nature, it is plausible to assume that viruses, and RNA viruses in particular, can only survive this enormous burden of errors by an equally remarkable within-host and between-host reproduction capability. Mutation rates can be a crucial parameter in EIDs as they have a direct influence on the ability to adapt to novel host species. Therefore, the large number of viruses, especially RNA viruses, in the group of zoonotic EIDs is not surprising.

1.4.2.2 Adaptation by Recombination

Recombination changes the viral genome in a different way. Rather than altering single sites in the genome through errors in replication, it allows for acquisition or exchange of foreign genetic material into the virus genome. In contrast to mutations, this can only happen in a cell infected by more than one virus. Recombination occurs through splicing of the foreign component into the viral genome (Worobey and Holmes, 1999). This process can be a homologous crossing over of nucleic acids strains, which replace each other in the respective genome of two nearly identical RNAs, or a non-homologous recombination resulting in deletions, insertions, and repetitions between two RNAs that have a short

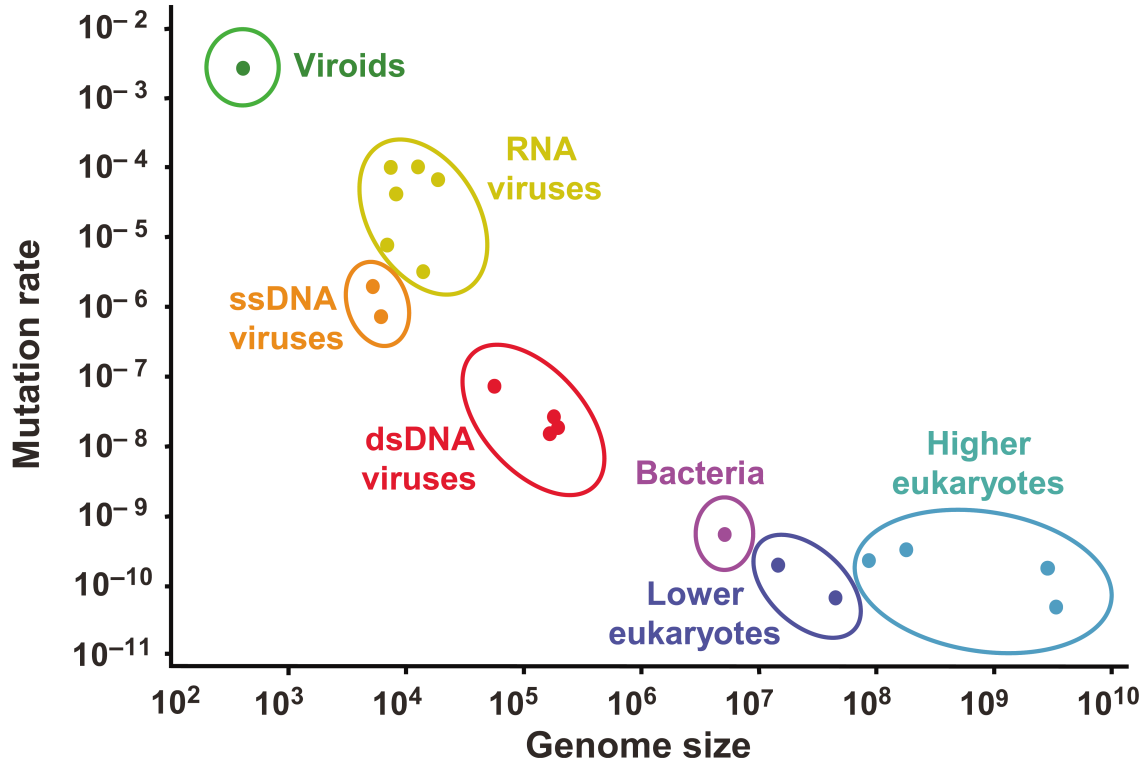

Figure 1.2: Overview of per site mutation rates per replication cycle for viruses and per generation for bacteria and eukaryotes versus genome sizes. Viroids, plant pathogens with minimal RNA genomes, exhibit the highest mutation rates while also having the smallest genome size. RNA viruses have a larger genome size with smaller per-site mutation rates, but the number of mutations per genome is comparable. Single-strained DNA viruses, and especially double-strained DNA viruses, show smaller mutation rates per-site with less mutations per genome. Bacteria and eukaryotes have the largest genome sizes and smallest mutation rates. The mutation rate is a crucial parameter in EIDs as it has a direct influence on the ability to adapt to novel host species. Therefore, the large number of RNA viruses in the group of zoonotic EIDs is not surprising. Source: Gago et al. (2009).

antiparalell stretch of complementarity (Roossinck, 1997).

As with mutations, the genome changes might be non-beneficial. Nevertheless, recombination has the potential to change the phenotype in a way that allows for novel epidemiological properties, which facilitate the likelihood of between-species jumps. This process is most often observed in RNA viruses such as coronaviruses or influenza (Nagy and Simon, 1997).

1.4.2.3 Adaptation by Reassortment

Reassortment is the last of the three possible genome changing mechanisms. It is a similar process to recombination with its acquisition of foreign genetic material, because it requires a host cell to be infected by two different virions. However, it does not alter the RNA segments directly. Instead, reassortment is the packaging of different RNA segments into one viral envelope (Roossinck, 1997). Multiple different processes play a role in reassortment, such as assortment of chromosomes and chromosomal crossover (Alberts et al., 2009). For example, influenza A(H1N1) consists of eight genome segments, which have been reassorted from different human, swine and avian influenza lineages (Smith et al., 2009). If a host cell gets infected with two different viruses such as two influenza strains each with eight genome segments, reassorted offspring might have up to $2^8 = 256$ different genomic combinations. However, studies with the influenza type A virus indicate the existence of a selective rather than random mechanism for packaging genome segments into virions, which would narrow the expected genomic diversity of progeny virus (Fujii et al., 2003).

Reassortment in genome segments can be the reason for an antigenic shift in influenza and result in novel pandemic strain such as influenza A(H1N1). Day et al. (2006) estimate the number of necessary reassortments to explain the number of influenza pandemics of the last 250 years. Their model suggests that either the number of co-infections or the probability or reassortment need to be much higher than previously expected, even though empirical data is missing to validate these estimates. Webby et al. (2004a) point out that little is known about the full conditions and likelihood of reassortment, except about the

necessary compatibility of viral gene segments (Webby et al., 2004b).

1.5 The Venezuelan Equine Encephalitis Virus as Example

One of the few examples of well understood host range change is the Venezuelan Equine Encephalitis Virus (VEEV) (Holmes, 2006). Anishchenko et al. (2006) were the first to fully explain the underlying factors and mechanisms facilitating the observed host species jump. Here, I will review the key findings of Anishchenko et al. (2006) and relate them back to the insights previously presented in this chapter.

The VEEV is a mosquito-borne RNA alphavirus enzootic in rodents. Transmitted to humans and horses, it causes severe, sometimes fatal, disease. While other arboviruses like the West Nile virus produce incidental spillovers into the human population, VEEV sometimes causes epidemics with an incidence of hundreds of thousands of infected human hosts in the Americas. What allows for these explosive outbreaks of VEEV? The answer is a combination of evolutionary adaptation and environment, as described in section 1.2.2. The wildtype strain is avirulent to horses, but a single site mutation in the pathogen's genome is sufficient to produce a virulent, epidemic-capable strain of VEEV. Given the high mutation rates of RNA viruses (see 1.4.2.1), a high frequency of this mutation is expected. Indeed, this mutation happens on average in every transmitting mosquito, given an average RNA virus per site mutation rate of 10^{-4} and an average mosquito viral load of 10^6. Nonetheless, epidemics are relatively rare.

This leads to another factor: the environment. The enzootic disease is usually transmitted by mosquito vectors of the genus*Culex (Melanoconin)*, which usually feed on rodents such as rats. Rarely, these mosquitoes have blood-meals at horses. Instead, floodwater mosquitoes of the genus *Aedes* and *Psorophora* usually feed on horses, and it is believed that these mosquitoes are able to transmit VEEV as vectors. Furthermore, humans living in proximity to horses are at risk of infection, because of the mosquitoes' wide host range.

Therefore, environmental factors rather than adaptive difficulties seem to be the bottleneck for a successful emergence. The density and proximity of donor and recipient hosts species, paired with the abundance of floodwater mosquitoes usually common in marsh

or pasture habitats, which is influenced by rainfall, climate, and competition, seem to be the main factors determining the zoonotic emergence into the human population.

VEEV is an excellent example of the previously described complex interactions between environment, hosts, and pathogen. Neither evolution nor the environment play an exclusive role in the generation of zoonotic EIDs. It is the right balance which allows for this. It might be that an infectious agent gained all necessary adaptations in the donor species and can jump between animals and humans like Dengue (Monath, 1994). Other possibilities include that only minimal adaptations are necessary, like with VEEV (Holmes, 2006). But neither virus can jump species if the environment does not provide the necessary vector. On the other hand, environmental conditions might allow for zoonotic infections with a subsequent emergence, but the pathogen appears to be not well enough adapted to the human physiology, as is assumed for influenza A(H5N1) (Abdel-Ghafar et al., 2008). It is a valuable reminder for every researcher in the field of epidemiology and public health that a failed emergence can be the result of both - or just chance, even if the conditions would have been sufficient for emergence. Even more, changes might have a devastating or beneficial effect, both in the environment as well as in the pathogen's genome. Unfortunately, judgements of the effects are often only made in retrospect, which highlights the need for predictive theories and models.

CHAPTER 2

A DIVE INTO INFECTIOUS DISEASE MODELLING

2.1 The Use of Theoretical Models in Infectious Disease Research

Researchers in epidemiology and public health often look for answers to large scale problems: how fast will an epidemic advance? Who should be vaccinated? Which age group is especially at risk? Answers to these questions are most important. Because of the scale of an epidemic, even slight improvements to public health strategies can save many lives. But the scale of epidemics and their severity also cause problems in our efforts to replicate, predict, and explain the behaviour of infectious diseases with experimental studies. One possibility is to use model systems. But many aspects cannot be fully evaluated in an experimental setting. That is the point where theoretical models can offer valuable insights, and allow us to predict and explain the risks of infectious diseases without harming any creature. Indeed, modelling infectious diseases has a considerable history, and the academic literature on this subject is extensive (Anderson and May, 1992; Keeling and Rohani, 2007).

Hamer (1906) was the first to use simple but precise mathematical formulations of the transmission of infectious disease. Over twenty years later, Kermack and McKendrick (1927) published the first 'modern' mathematical model of infectious disease transmission incorporating a necessary threshold of susceptible hosts for sustained disease transmission. Subsequently, advances in mathematics and biology led to new insights and continuing improvements in the mathematical understanding of epidemiological processes, and a large variety of model types (May, 2004). The best way to classify theoretical models is by the problem they are trying to resolve. While this does not necessarily resemble the most common nature of model classification, it usefully focuses upon the research question, and is therefore the method of choice here. Models can represent within-host dynamics or between-host dynamics, they can be stochastic or deterministic, they can incorporate multiple strains or multiple populations, just to name some possible differences. While some models look at within-host dynamics in zoonotic emergences (Iwasa et al., 2004), I will concentrate on models looking at between-host dynamics as they are of greatest relevance to public health questions.

The sheer number of different between-host dynamic models is explained by the trade-

off between model complexity, realism, and relevance. In general, a model gets more complex as realism is increased. For example, models usually simplify factors like demography or contact structures. Otherwise, they would be computationally infeasible. Hence, theoretical models are usually a balance between necessary realism and minimising complexity, while increasing relevance. In some situations, models can become more realistic and more complex but the overall relevance is not increased, because adding realism might not increase model usefulness (Dodd and Ferguson, 2007). Indeed, many questions researchers address can be best explained with simple models (Arino et al., 2006; Keeling and Rohani, 2007). Ideally the scientific question will determine the appropriate levels of realism and complexity for a model (Wearing et al., 2005).

The main benefit of epidemiological models is the possibility to explore a vast range of situations and containment interventions without putting any human in danger (Louz et al., 2010). For example, Longini et al. (2005) examine in detail the possible advantages and disadvantages of containment strategies just after the emergence of a novel infectious disease, while Ferguson et al. (2005) try to model a single country in detail to predict the spread of an infectious disease on a national level. Theoretical modelling has also been used to simulate response strategies to biological warfare (Wein et al., 2003). A common finding is that intervention measures are most effective at the source when case numbers are still small (Wallinga et al., 2010). But modelling also allows for testing public health measures during an epidemic. For example, it can be used to evaluate vaccination strategies (Brown and White, 2010), especially under the stress of limited vaccine supply (Bansal et al., 2006). Furthermore, theoretical modelling is also used in plant diseases (White and Gilligan, 1998).

In light of problems with investigating zoonotic EIDs in experimental settings, and paired with the advantages of theoretical models, my research largely relies on theoretical models, which is a common approach in the field of EIDs (Antia et al., 2003; Yates et al., 2006; Arinaminpathy and McLean, 2009; Louz et al., 2010). Here, I focus on between-host transmission models, and give a brief overview of their use in public health and theoretical epidemiology. I will start with the introduction of the most important parameter to measure the pathogen's transmission potential, and then introduce compartmental models be-

fore I concentrate on more complex models incorporating heterogeneities. Subsequently, I will review the literature on theoretical models of zoonotic EIDs and their findings.

2.2 The Basic Reproductive Number

In the first chapter, I discussed the environmental conditions pertinent to an infectious agent, and the biological processes which allow a pathogen to adapt to its surrounding and increase its fitness. In general, 'fitness' describes how well a pathogen performs in its environment, which means in this context how well the pathogen transmits between hosts. The pathogen's transmission potential is therefore a measure of its fitness and level of adaptation, and builds the basis for the definition of emergence (see section 1.1).

Nowadays the most commonly used term to define the reproductive capability of a pathogen in a given population is the basic reproductive number R_0 (Anderson and May, 1992; Heffernan et al., 2005; Keeling and Rohani, 2007). It defines the average number of secondary infections arising from a single case in a completely susceptible population (Anderson and May, 1992). It can also be expressed as $R_0 = n\,T\,t$, where n represents the average number of contacts each infectious host has per unit time, T is the average probability of transmission between an infectious and susceptible host, and t is the mean infectious time (Lipsitch et al., 2003). The number of infectious hosts can only grow and cause an emergence for $R_0 > 1$. On the other hand, it will decline and cause only an outbreak if $R_0 < 1$. It is therefore similar to a birth-death process. A population will only grow if there are more births than deaths, otherwise it will shrink.

Further, the number of infectious hosts will stay constant for $R_0 = 1$. Again, this is similar to a birth-death process where the population size stays constant if a new individual is born for every one that dies. Nonetheless, this is only correct if transmission is a deterministic process, leading to exactly one new infectious host for every recovered or removed one. If the process is stochastic - as every realistic transmission process is - the pathogen will eventually die out due to stochastic fluctuations, leaving $R_0 > 1$ as a necessary condition for sustained transmission or emergence of an infectious disease (Anderson and May, 1992). This condition has been observed with many well studied infectious diseases.

For example, the Measles virus often causes outbreaks but no epidemic, because the basic reproductive number is below one due to mass-vaccination (Jansen et al., 2003).

Nonetheless, $R_0 > 1$ is only a condition, not a guarantee for emergence and sustained transmission. As mentioned before, the process is prone to stochastic variation. The fate of an emergence is unpredictable, especially at low infectious numbers, as the probability of a failed emergence, or probability of extinction, follows $p_{\text{ex}} = (1/R_0)^x$ where x is the number of infectious hosts within a homogeneous host population (May et al., 2001). This formulation is similar to the standard Yule-Kendall process, which describes the survival process of family names. Each infectious host has a chance of $1 - 1/R_0$ to cause an emergence. It is sufficient if only one of the x infectious hosts causes a transmission chain without extinction for a pathogen to emerge, because all infectious hosts are independent of each other. Hence, the overall probability of extinction is the combined probability of extinction for each infectious host.

Because R_0 incorporates the average number of secondary infections, it depends on the biological properties of the infectious agent, as well as the environment it spreads in. Both aspects influence the infectious agent's reproductive capability, and it might be impossible to define the exact influence of each of the two factors on R_0. Nevertheless, it is reasonable to assume that a directly transmitted pathogen will spread better, and therefore have a larger R_0, in densely populated areas than in rural, sparsely populated communities.

2.3 Basic Compartmental Models

A model of an epidemiological process needs to make some simplifying abstractions in its description of individuals. One possibility is to describe all individual hosts by their infectious state, as is done in compartmental models. There, hosts are accumulated according to their appropriate state, and the flux between states is explicitly modelled. They are one of the least complex but most helpful classes of models, also known as Susceptible-Infectious-Recovered (SIR) models. The first of these models was proposed by Lowell Reed and Wade Hampton Frost in the 1920s (Keeling and Rohani, 2007), and compartmental models have since become the 'workhorse' of theoretical epidemiology (Newman, 2002). Figure 2.1

gives an overview of some of the most common compartmental models.

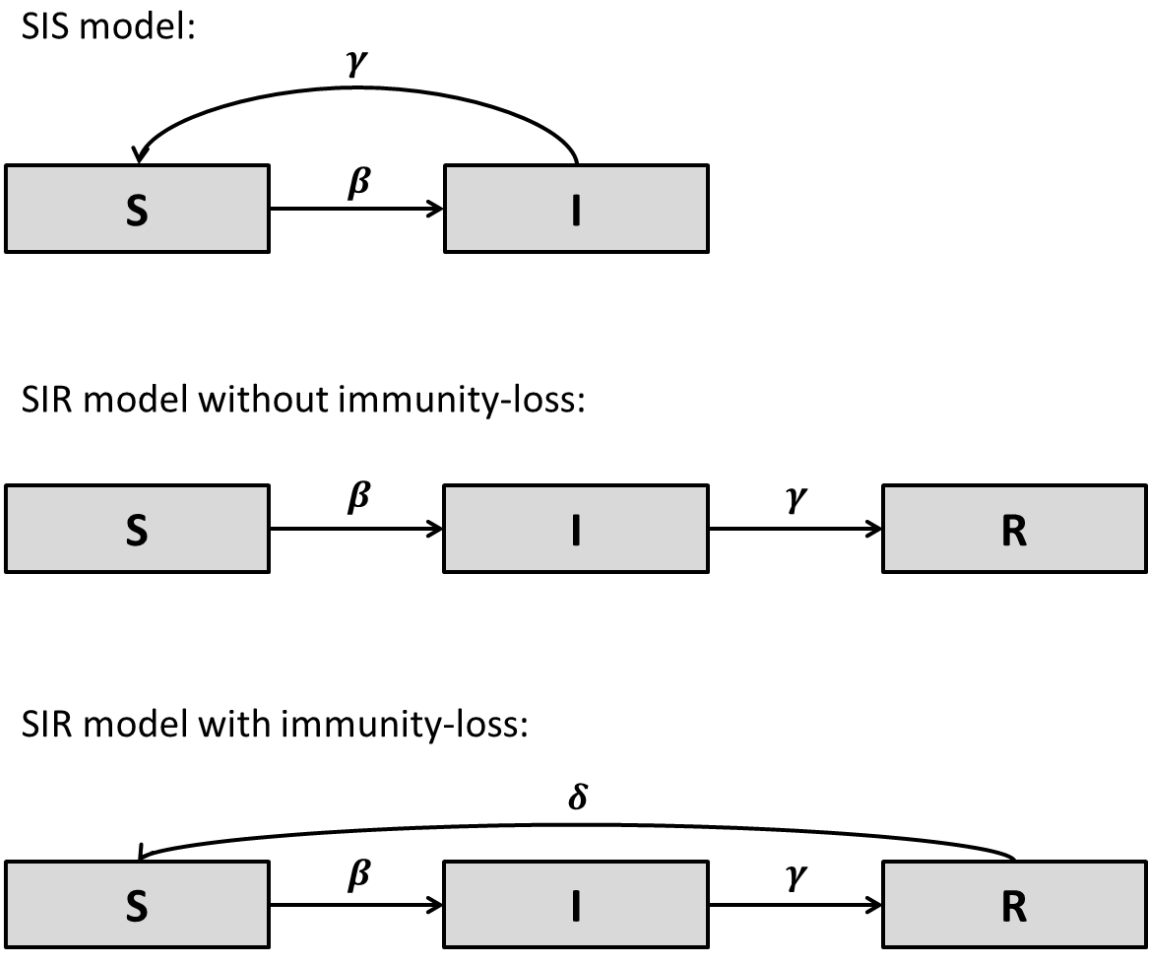

Figure 2.1: Three common types of model. The first is an SIS model where infectious hosts recover back into the susceptible state. This model does not contain any form of immunity. On the other hand, the SIR model without immunity-loss incorporates permanent immunity. Infectious hosts recover into the recovered class. This prohibits re-infection, and is used to model infectious diseases inducing life-long immunity. The last model allows for temporary immunity, determined by the immunity-loss rate δ. Like in the other SIR model, infectious hosts recover and gain immunity. But immunity is not life-long anymore, and hosts lose immunity at per capita rate δ.

2.3.1 Mathematical Formulation using Differential Equations

SIR models assume that every population can be divided into three states, suceptible (S), infectious (I), and recovered (R). Furthermore, SIR models are full contact models, which means every individual is in in contact with every other. In addition, the contact between

an infectious individual and a susceptible individual can lead to a new infection - the susceptible host becomes infectious. The transition is characterised by the rate β, the infectious rate. It incorporates the probability of establishing a contact and subsequently transmitting the infectious disease (Antolin, 2008). Hence, various within- and between-host factors are packed into a single parameter to express a rate of transmission. On the other hand, infectious hosts leave the infectious state through recovery. Again, various factors can have an influence on this process, which are packed into a single rate γ, called the recovery rate. It is the rate with which infectious individuals recover and acquire immunity. The basic SIR model can be expressed with differential equations

$$\begin{aligned} \frac{dS}{dt} &= -\beta \frac{I\,S}{N} \\ \frac{dI}{dt} &= \beta \frac{I\,S}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I \\ \text{with } \; N &= S + I + R = \text{const.} \end{aligned} \tag{2.1}$$

This formulation is called mass action (Anderson and May, 1992; Dobson, 2004; Keeling and Rohani, 2007). The transmission is frequency-dependent, and S, I, and R represent the actual number of hosts. Another way to express the SIR model is using pseudo-mass action (McCallum et al., 2001; Keeling and Rohani, 2007). This approach models density-dependence in infectious disease transmission. Written in differential equations, a pseudo-mass action model reads

$$\begin{aligned} \frac{dS}{dt} &= -\beta\, I\, S \\ \frac{dI}{dt} &= \beta\, I\, S - \gamma\, I \\ \frac{dR}{dt} &= \gamma\, I \\ \text{with } \; N &= S + I + R = \text{const.} \end{aligned} \tag{2.2}$$

The difference between frequency- and density-dependence lies in the relationship between population size and contact rate. If both models have the same population size N, the contact rate of the density-dependent model must be N times smaller than the one

in the frequency-dependent model to yield the same number of transmissions. Therefore, the difference in the models becomes important if the population sizes changes, or if a range of population sizes are modelled in an attempt to parameterise empirical infectious disease data (Dobson, 2004).

The basic SIR model is like a one-way street: once infectious, hosts can only recover with full immunity, excluding the possibility for re-infection. It is therefore, more accurately expressed, the SIR model without immunity-loss. The SIS model is very similar to the basic SIR model. But it does not incorporate any form of immunity (Keeling and Rohani, 2007). Infectious hosts recover back into the susceptible state, instantaneously ready for a possible re-infection. Expressed with differential equations, the model reads in frequency-dependent formulation

$$\begin{aligned} \frac{dS}{dt} &= -\beta \frac{I\,S}{N} + \gamma I \\ \frac{dI}{dt} &= \beta \frac{I\,S}{N} - \gamma I \\ \text{with} \quad N &= S + I = \text{const.} \end{aligned} \tag{2.3}$$

These models are the two extremes of a possible immunity induced by infection. The SIR model with immunity-loss allows for a middle ground. It consists of the same three stages as the previous SIR model, but like in the SIS model recovery from infection does not induce life-long immunity. Expressed with differential equations, the model dynamics read

$$\begin{aligned} \frac{dS}{dt} &= -\beta \frac{I\,S}{N} + \delta R \\ \frac{dI}{dt} &= \beta \frac{I\,S}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I - \delta R \\ \text{with} \quad N &= S + I + R = \text{const.} \end{aligned} \tag{2.4}$$

Here, I make use of another, novel rate. δ is the immunity-loss rate and describes the flux of recovered back into the susceptible hosts class. As with β and γ, many within-host processes can have an influence on this rate.

While all of these three described models allow for modelling an epidemic, only the SIS model and the SIR model with immunity-loss allow for endemic or multi-seasonal modelling if no process to generate new susceptible hosts is added. Hence, an important extension to these models is to incorporate birth and death processes. These allow for models with demography. Note that all the formulations presented here are deterministic models.

2.3.2 Estimating R_0 and Other Useful Calculations

It is possible to define a relation between the basic reproductive number and compartmental models. As shown in equation (2.1), a single infectious host causes on average β new infections per time unit. In addition, the host is on average infectious for $1/\gamma$ time units, the infectious period. Therefore, the average number of new infections caused by an infectious host in frequency-dependent models is (Anderson and May, 1992)

$$R_0 = \frac{\beta}{\gamma} \tag{2.5}$$

As previously, the basic reproductive number is defined for a naïve population. But the fraction of susceptible hosts decreases in every SIR model as a result of infections. It leads to a probability smaller than one for an infectious host having contact with a susceptible host. Using the susceptible host fraction, I can define the effective reproductive number

$$R_{\text{eff}} = R_0 \frac{S}{N} = \frac{\beta}{\gamma} \frac{S}{N} \tag{2.6}$$

The effective reproductive number becomes the basic reproductive number for $S/N \to 1$. It is directly apparent that the initial number of susceptible hosts needs to be above the critical threshold of $S(0) > N/R_0$ for $R_{\text{eff}} > 1$. If the number of susceptible hosts is below this threshold, every infectious host causes less than one new infection and the disease dies out.

Nonetheless, this is eventually the fate for every infectious diseases with life-long im-

munity. Rewriting equation (2.1) under the use of equation (2.5) yields

$$S(t) = S(0)\, e^{-R(t)R_0/N} \tag{2.7}$$

which is just the integral form of (2.1). But it yields further insights: given that the number of recovered hosts can never be larger than the total population size, $R(t) \leq N$, the number of susceptible hosts can never drop below e^{-R_0}. Before this point, the chain of transmissions eventually breaks due to the decline in infectious and susceptible hosts. Unfortunately, no explicit algebraic solution of the final number of susceptible hosts is possible because of the transcendental nature of equation (2.7), but numerical root finding methods allow for a very good approximation of the final number of susceptible hosts.

The number of infectious hosts can also be extracted from equation (2.1). I use the fact that $S/N \approx 1$ at the initial phase of an epidemic. The integral form of the infectious host number reads

$$I(t) \approx I(0)\, e^{(R_0-1)t} \tag{2.8}$$

This relation is especially useful as it allows for an estimation of the basic reproductive number given the initial growth of an epidemic.

Nevertheless, researchers must be fully aware of the underlying assumptions, and therefore strict limitations, of compartmental models if they use these models to estimate epidemiological parameters like R_0 (Wearing et al., 2005). Furthermore, many different approaches to estimate R_0 exist in the scientific literature (Heffernan et al., 2005).

2.3.3 Deterministic or Stochastic? - The Gillespie Algorithm

The compartmental models presented thus far are deterministic models. These ordinary differential equations will always describe the same outcome given constant parameters and initial conditions. However, one would not expect such behaviour from a real epidemic. Indeed, it is impossible to predict if a certain host will get infected or not, even if all epidemiological parameters and the number of initial cases are known. The spread of infectious diseases is a stochastic process, and deterministic models only represent an

average of a large range of outcomes (Keeling and Ross, 2008). Therefore, deterministic models might be unsuitable if stochastic events are important or even the cause of interest.

This is usually the case if the the number of infectious hosts is small. As mentioned in section 2.2, the probability of extinction is $p_{ex} = (1/R_0)^x$ where x is the number of infectious hosts within a homogeneous host population (May et al., 2001). Therefore, a standard SIR model with one initially infected host and $R_0 = 2$ will always produce an epidemic of constant size, even if the actual probability of extinction is $p_{ex} = 0.5$. Hence, this type of compartmental models only describe the average behaviour of an epidemic, ignoring all failed introductions.

The probability of extinction is especially interesting to know for EIDs. These diseases are, by definition, not widely distributed before their emergence, and the probability of extinction gives direct information on the likelihood of an emergence. Furthermore, the emergence of zoonotic infectious diseases might incorporate pathogen stages with unsustainable between-host transmission where stochastic mutations are necessary to cross the fitness valley.

It is possible to extend deterministic compartmental models with stochastic terms to overcome this shortcoming (Brockmann et al., 2005; Keeling and Ross, 2008). Another possibility is stochastic simulations based on a deterministic compartmental model. In a highly influential article, Gillespie (1977) was the first to describe an algorithm to translate deterministic dynamics of differential equations into stochastic simulations. Even though the original algorithm has been based on a system of chemical reactions, the algorithm can also be used for SIR models, because of the similarity between molecular dynamics and infectious processes in compartmental models.

The Gillespie algorithm is one of the most popular ways to derive stochastic results from an originally deterministic model (Keeling and Rohani, 2007). The key idea of the algorithm is to determine the next event (see the appendix for an implementation of the Gillespie algorithm). For a simple SIR model, possible events are the infection of a previously susceptible host, or the recovery of a previously infectious host. Both event probabilities are functions of the infectious rate and recovery rate. These probabilities of all possible events can be expressed in a probability density function for the next event. Furthermore,

the waiting time distribution for the next event, or in plain words when the next event will happen, is an exponential decay function dependent on the next event probabilities.

The Gillespie algorithm consists out of five main steps:

(1) initialise the number of hosts in each compartment, initial time, and random number generator

(2) calculate the probability density functions for the next event and waiting time

(3) generate a pair of random numbers to determine the next reaction to occur as well as the elapsed time interval

(4) add the elapsed time interval and re-calculate the number of hosts in each compartment according to the occurred event

(5) go back to step (2) unless the next event probabilities are zero or the simulation time has been exceeded

This is a simple, fast, and rigorous way given the reliable unit-interval uniform pseudo-random number generators widely available in modern programming languages. The algorithm stops iterating itself if the probability for a next event to happen is zero. For example, this is the case if an epidemic eventually dies out due to the reduced numbers of susceptible hosts. This equals the case when there are zero infectious hosts left in a SIR model.

2.4 Modelling Epidemics with Branching Processes

One way to look at epidemiological processes is by only looking at overall number of infectious hosts as in compartmental models. Another approach is to look at each single infected host. Nowadays, such models are often described as agent-based models (Keeling and Rohani, 2007). Here, I follow this definition and include all models looking at the individual host in the category of agent-based models. This also includes models based on branching processes. Epidemiological models based on branching processes are by definition stochastic models. Nonetheless, they can produce deterministic results, such as

the average value and variance of epidemic sizes. Historically, they have been used in other fields like physics to describe electron avalanches in Geiger-Müller tubes, but also in biology for survival processes of mutant genes.

The classical Galton-Watson branching process is the starting point for many epidemiological models (Jacob, 2010), and previous applications range from epidemiological models for the common cold (Becker, 1981) to epizootic diseases in various environments (Trapman et al., 2004). More complex models resemble particular compartmental models, such as the SIS model (Jacob and Viet, 2003), or allow for deeper insights in the initial phase of an epidemic (Trapman and Bootsma, 2009). Another way to describe branching processes is with Markov chains. Jansen et al. (2003) use this approach to estimated the effective reproductive number of measles through outbreak sizes in the United Kingdom. While $R_{\text{eff}} < 1$, the reproductive number has been increasing over years, because of reduced MMR vaccine uptake. They also predict that if vaccine uptake were to decrease further, a re-emergence of measles is likely.

This important study reveals the usefulness of theoretical models. They allow us to estimate relevant parameters for public health, which then yield practical insights into the likely course of future events. In general, branching process models are valuable in increasing our understanding of the spread of infectious diseases, and public health efforts for disease control. They are especially useful if the number of individuals is small and if the main focus of interest lies in the numerous possible outcomes of a pathogen introduction. On the other hand, branching process models usually have a much greater mathematical complexity compared to compartmental models, which drastically increases for any kind of host heterogeneity (Karlin and Taylor, 1975). Therefore, this type of model is not a good choice if only the average epidemiological behaviour is of interest.

2.4.1 Probability Generating Functions

Before going into detail with branching processes, a short introduction to probability generating functions (pgfs) is useful. The idea behind pgfs is to provide all different characteristics of a probability distribution within one function (Wilf, 1994). A pgf usually has the

following form

$$G_0(s) = \sum_{j=0}^{\infty} p_j s^j \tag{2.9}$$

with p_j as event probabilities. Note that the sum over all probabilities holds

$$G_0(1) = \sum_{j=0}^{\infty} p_j = 1 \tag{2.10}$$

Given the pgf, the probability distribution can be obtained by repeated differentiation

$$p_j = \frac{1}{j!} \left[\frac{d^j G_0}{ds^j} \right]_{j=0} \tag{2.11}$$

Furthermore, pgfs contain all moments of a probability distribution. The first moment, the mean, is given by the first derivative of the pgf with $s = 1$

$$\langle j \rangle = \sum_{j=0}^{\infty} j p_j = \left[\frac{dG_0}{ds} \right]_{s=1} = G_0'(1) \tag{2.12}$$

This can be generalised to moments of higher order

$$\langle j^n \rangle = \sum_{j=0}^{\infty} j^n p_j = \left[\left(s \frac{d}{ds} \right)^n G_0(s) \right]_{s=1} \tag{2.13}$$

Another advantage of pgfs is their ease of use with powers, $[G_0(s)]^n$. If I assume $n = 2$, it is

$$\begin{aligned} G_0(s)\, G_0(s) &= \sum_{k,l=0}^{\infty} p_k p_l s^{k+l} \\ &= \left[\sum_{j=0}^{\infty} p_j s^j \right]^2 \\ &= [G_0(s)]^2 \end{aligned} \tag{2.14}$$

An extension to powers of higher order is analogous.

2.4.2 Basic Mathematical Formulation of Branching Processes

Following Karlin and Taylor (1975), the basic process is universal and starts with the initial number of infectious host, X_0. Each individual host infects k new ones, independently of the others, with probability p_k. Therefore, the average number of new infections per infectious hosts equals the basic reproductive number

$$R_0 = \sum_{k=0}^{\infty} k p_k \tag{2.15}$$

In addition, the totality of all newly infected hosts, the first generation of offsprings, is denoted by X_1. This process continues, and the i-th generation, with a size of X_i, is composed of descendants of the $(i-1)$-th generation.

Using the independence of individuals, I can rewrite a process with $X_0 = n$ initially infected hosts into n processes with $X_0 = 1$ initially infected hosts. Then, the i-th number of offspring is defined as

$$X_i = \sum_{r=1}^{X_{i-1}} \zeta_r \tag{2.16}$$

where ζ_r are independently identically distributed (iid) random variables with

$$\Pr\{\zeta_r = k\} = p_k \quad \text{with} \quad \sum_{k=1}^{\infty} p_k = 1 \tag{2.17}$$

As previously shown, pgfs are a convenient way to handle probability distributions. The pgf for the number of offspring is

$$\varphi(s) = \sum_{k=0}^{\infty} p_k s^k \tag{2.18}$$

and

$$\varphi_i(s) = \sum_{k=0}^{\infty} \Pr\{X_i = k\} s^k, \quad \text{for } n = 0, 1, 2, \ldots \tag{2.19}$$

Since ζ_r are iid random variables with the common pgf $\varphi(s)$, the sum $\zeta_1 + \ldots + \zeta_j$ has the

pgf $[\varphi(s)]^j$. It follows that the i-th generation pgf is

$$\varphi_i(s) = \sum_{j=0}^{\infty} \Pr\{X_{i-1} = j\} [\varphi(s)]^j = \varphi_{i-1}(\varphi(s)) \tag{2.20}$$

This result can be used to evaluate the expected generation size EX_i. As a standard result from pgfs shows (see equation (2.12)), it is

$$EX_i = \varphi_i'(1) \tag{2.21}$$

Then differentiating equation (2.20) and setting $s = 1$, I obtain

$$EX_i = \varphi_i'(1) = \left[\varphi'(1)\right]^i = EX_1^i \tag{2.22}$$

as a standard result. Furthermore, the variance can be obtained in a similar way, also using standard characteristics of pgfs.

2.5 Modelling Host Heterogeneity

These models are useful tools, yet they ignore any host heterogeneity, which in many cases matters greatly. For example, spatial heterogeneity can reduce disease persistence (Hagenaars et al., 2004), increase temporal heterogeneity of an epidemic (Watts et al., 2005), or may need different public health control measures than homogeneous populations (Roberts et al., 2007). However, the most important point might be the impact of spatial heterogeneity on the size of an epidemic. Here, compartmental models can yield some insights.

Equation (2.7) gives the the epidemic size which is $1 - S(\infty)$. While it is impossible to derive an explicit expression for the epidemic size in this type of compartmental models, (2.7) shows that it is solely determined by the basic reproductive number and population size. Given two populations with a fixed N, the epidemic size solely depends on R_0, a measurement of the transmission potential. As shown in the first chapter, the number of transmission events of an infectious agent depends on its biology and the environment it

spreads in. It is natural to expect that a directly transmitted pathogen in a rural, sparely populated community has a lower R_0 than it has in a densely populated city with the same population size. Therefore, (2.7) would yield a larger epidemic size in the city than in the rural community, showing the effect of spatial host heterogeneity on the epidemic size. While this result is based on a compartmental model, more complex models show the same qualitative results (Tildesley and Keeling, 2009).

In the following, I will present two common approaches to modelling spatial host heterogeneity in more detail: meta-population models which are similar to the previously introduced SIR models, and a different approach known as network models.

2.5.1 Meta-Population Models

On a large scale, meta-population models can be used to model towns and villages connected with each other (Keeling and Rohani, 2007). While heterogeneity in the number of contacts looks at direct neighbours and is best addressed with networks, meta-population models use the approach that the probability of contact is higher with hosts from the same community than with hosts from other ones. However, the contact structure is homogeneous within such communities - often called sub-populations. Previously, meta-population models have been used to reveal spatial heterogeneity's influence on the persistence of disease (Hagenaars et al., 2004) as well as control efforts (Riley and Ferguson, 2006). For example, May and Anderson (1984) show that fewer hosts need to be vaccinated to generate herd-immunity in a heterogeneous population compared to a homogeneous one using a meta-population model.

The most simple meta-population model consists of just two connected populations, forming a meta-population. In this form, the connection is realised by allowing for infections from both populations, which is expressed in a combined 'force of infection' (May and Anderson, 1984; Lloyd and May, 1996). Such coupling does not incorporate moving hosts between the populations, which greatly reduces the model's complexity. Incorporat-

ing the most simple meta-population structure into (2.1) reads

$$\begin{aligned} \frac{dS_i}{dt} &= -\lambda_i \frac{S_i}{N_i} \\ \frac{dI_i}{dt} &= \lambda_i \frac{S_i}{N_i} - \gamma_i I_i \\ \frac{dR_i}{dt} &= \gamma_i I_i \end{aligned} \tag{2.23}$$

$$\text{with } \lambda_i = \sum_j \beta_{ij} I_j \text{ and } N_i = S_i + I_i + R_i = \text{constant}$$

Here i denotes the different populations and β_{ij} represents the rate with which an infectious hosts from population j will infect a susceptible host in population i. This allows for calculating the force of infection λ_i which is a representation of the coupling strength.

The beauty of meta-population models is that host heterogeneity can be incorporated into the standard SIR model without loosing most of its simplicity. This teaches us that the effects of host heterogeneity can be discovered with simple methods, and gives evidence to the notion that models do not need to be overly complex to be relevant.

Keeling and Rohani (2002) show that in most cases the coupling through a force of infection yields comparable dynamics as a model of explicit host movement would. Mechanistic models use an explicit number of moving hosts (Hufnagel et al., 2004), but are computationally more demanding. Again, this shows the impact of realism's relevance. Only if the movement rates become similar to infectious rates does population coupling start to show differences from the mechanistic model. This is because when the length of stay becomes similar to the infectious period, the likelihood that a host returns home before recovery is small, therefore limiting transmission within its home population.

The meta-population approach can also be used to hierarchically structure populations (Watts et al., 2005). The idea is that each single population does not necessarily need to be a community, it could be a household with other households on the same hierarchical level. All households within a community could then form another population on the next level, connected to other communities filled with household populations. Such an hierarchical model allows for greater realism as hosts usually have more and longer contacts within a household than in other environments. Further, it is possible to design networks

of meta-populations with a threshold connectivity to allow for infectious disease spread (Vazquez, 2007; Colizza and Vespignani, 2008). These models might be the appropriate choice if one wants to model the spread of an infectious disease through communities, which yield a different probability of extinction compared to homogeneous compartmental models (Vazquez, 2007). However, extensive and complex analytical expressions are necessary to determine properties of these type of models, making it demanding to gain any extra information.

2.5.2 Network Models

Meta-population models are a valuable tool for community-based structures. But even hierarchical meta-populations do not include spatial structure on an individual scale (Watts et al., 2005). One way to accurately represent such connections is with networks (Newman, 2003), often called graphs in mathematics. As with branching processes, networks look at the individual host level, and are therefore agent-based models rather than compartmental models. The simple idea behind network models is to represent each host as a node and draw edges between two nodes if a certain type of connection exists. Usually, this connection represents contacts appropriate for disease transmission. It follows that these contact networks are highly dependent on the pathogen and its mode of transmission, as well as the host environment. For example, one would expect fewer connections in sparsely populated regions than within cities. Such contact networks represent all possible trajectories with which an epidemic could advance through a host population.

The number of contacts a host has, hence the number of edges of a node, is called the 'degree' k. Usually, types of networks are separated by the underlying degree distribution, which is the probability distribution of host contact numbers. The probability of a random node having a degree k is denoted as p_k. However, connections are necessary but not sufficient for transmission. There exists a certain probability of transmission, T, which describes the likelihood that a host infects another one through their connection. Furthermore, each network type has a transmission threshold, corresponding to $R_0 = 1$, which

determines if epidemics are possible. This critical transmissibility can be derived using

$$T_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} \tag{2.24}$$

with $\langle k \rangle$ as mean degree and $\langle k^2 \rangle$ as mean square degree (Newman, 2002; Meyers et al., 2005). The critical transmissibility therefore depends on the average number of connections and the actual distribution of connections. Given the relation between T_c and $R_0 = 1$, the transmissibility can be expressed as a function of the basic reproductive number

$$T = R_0 \, T_c = R_0 \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} \tag{2.25}$$

Anderson and May (1992) reach a related result for the basic reproductive number of sexual transmitted diseases. Not only the mean number of partners, but also the variance of the partner distribution influences R_0. Their exact expression for the reproductive threshold is different to (2.24), because they use a meta-population model instead of a network of individuals. Nevertheless, the quality of the result with the disproportionately high influence of only a few hosts with a large number of contacts is present in both types of model.

However, there are some distinctive difference between the basic reproductive number and the tansmissibility. First, R_0 incorporates all biological and environmental aspects, and might therefore have different values in an urban or a rural environment (see chapter 1 and section 2.2). T, on the other hand, is independent of any host-contact heterogeneity, which is explicitly modelled with the network approach. Furthermore, R_0 offers an explicit number of secondary infections, while T is a only probability of transmission given an established connection. Knowing the network structure is a mandatory requirement to understand what the transmissibility means in terms of infected hosts.

On the other hand, the decomposition of R_0 into a transmissibility and a host-contact part (see (2.25)) explicitly shows the heterogeneity in disease transmission. Using epidemiological data, Meyers et al. (2005) show that so-called 'superspreaders' had an important influence on the SARS outbreak in 2002. These findings were also noted by Lloyd-Smith

et al. (2005) using stochastic branching process models. Superspreaders are individual hosts who cause a disproportionately high number of secondary cases, and can shape the overall behaviour of an epidemic. For example, it is likely that superspreaders drove the SARS outbreak and the majority of infected people caused less than one secondary case. Indeed, superspreading seems to be a general feature in infectious disease spread with a special impact on EIDs, reducing the overall likelihood of emergence, but enhancing the overall dynamics of successful emergences (Lloyd-Smith et al., 2005).

Networks have a clear advantage in easily representing host contact heterogeneities. They are powerful tools to discover the impact of spatial host contact structures, and help increase realism in predicting the effect of public health strategies. Nevertheless, they have some drawbacks.

One of the disadvantages in using networks is the large amount of accurate information needed to reproduce a real contact structure (Keeling and Eames, 2005; Kiss et al., 2005) - a condition which is rarely met. Researchers usually use theoretical host population structures like scale-free distributions which are based on strong assumptions. Furthermore, while theoretical random networks can easily construct host contact heterogeneity on a small scale, incorporating larger scale structures such as clustering on a community level, or allocating individuals to geographical locations, is a much more demanding task.

Often, simpler models are more effective and appropriate to use. Arino et al. (2006) show that compartmental models can be a more robust and relevant choice than more complex ones, like networks, if epidemiological parameters are not well known. This is usually the case for EIDs where knowledge is limited during the initial phase of emergence.

2.6 Overview of Zoonotic EID Models

Previously, I introduced theoretical modelling as an efficient and valuable tool to gain insights into complex biological systems. Here, I will focus on the use of zoonotic EID models to increase our understanding of the underlying biological processes, as well as their implications for public health. I will start with simple, homogeneous models and

review more complex models, including heterogeneity, in later sections.

2.6.1 Homogeneous Models of Zoonotic Emergences

In terms of the basic reproductive number, the pathogens of interest are assumed to have $R_0 \leq 1$ within human populations. This means that zoonotic pathogens need an evolutionary adaptation to the recipient species' physiology to transmit and survive within the novel, human host population (Webby et al., 2004a). New infectious cases mainly arise from between-species transmissions, and evolutionary adaptations of the wildtype strain are necessary to accomplish an increased transmissibility with $R_0 > 1$ amongst humans.

Antia et al. (2003) were the first to model the probability of a successful emergence of a zoonotic, adapting novel disease as a function of R_0 and the mutation rate μ. A branching-type formulation was used to incorporate the stochastic behaviour of a zoonotic emergence with a Poisson distributed number of secondary infections with mean R_0. Figure 2.2 shows a schematic representation of the adaptation process. Only very few introductions eventually lead to an emergence, and the probability of emergence per introduction, p_{em}, is very low, especially if the wildtype strain has $R_0 \ll 1$. In addition, the evolutionary routes of adaptation have a significant impact. p_{em} is drastically reduced if adaptation occurs in multiple numbers of adaptive steps, caused by the increase in the necessary number of mutations. Furthermore, the basic reproductive numbers of these adaptive steps, also called intermediate strains, play a vital role.

Though $R_0 < 1$ for all intermediate strains, the increase in fitness by mutation can vary highly. Especially evolutionary routes with a fitness valley, corresponding to necessary mutations with a decrease in fitness, lead to a decrease in the probability of emergence compared to models with a steady fitness increase.

Overall, one of the key findings of Antia et al. (2003) is that the basic reproductive numbers of not fully-adapted, intermediate strains contain important information about the risk of pathogen adaptation and emergence. In particular, intermediate strains with R_0 close to one, but still below one, have larger mean outbreak sizes, causing more infectious hosts to potentially allow for mutation and adaptation.

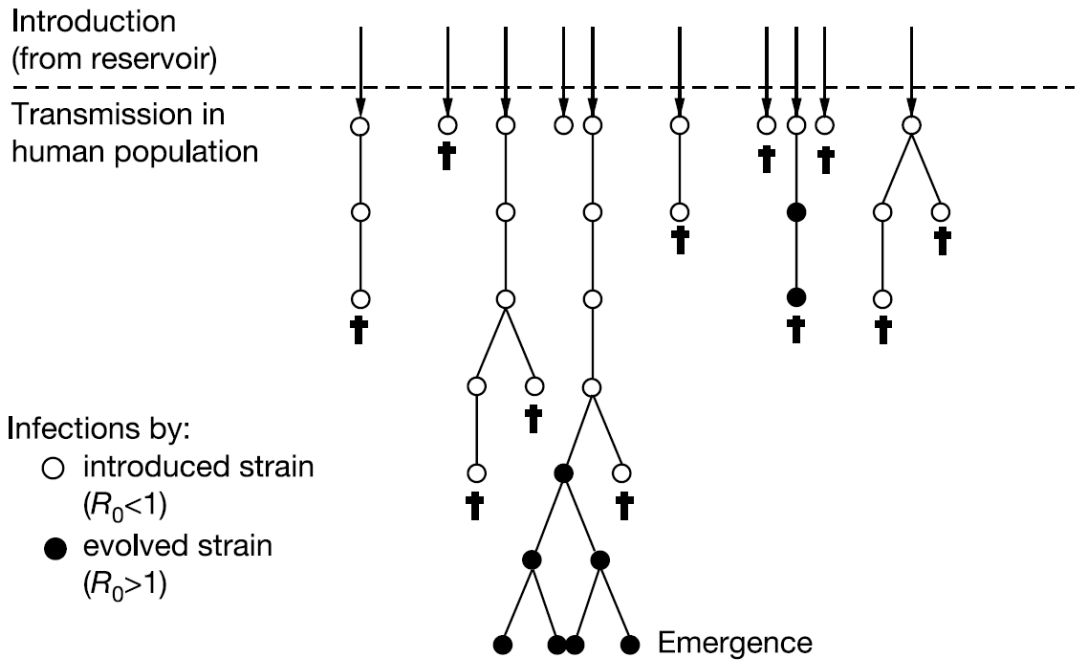

Figure 2.2: Schematic representation of the adaptation model. New introductions from the animal reservoir are followed by chains of transmission and mutations within the human host population. Open circles represent wildtype strain and filled circles evolved strains. Crosses indicate extinction. Source: Antia et al. (2003).

The disease life history can be another influencing factor (André and Day, 2005). Human-to-human transmission is the bottleneck for increased adaptation to the human recipient species, but within-host adaptation plays an important role in this. Long lasting infections cause a much higher risk as evolutionary adaptations become more likely, caused by the longer within-host selection process. This effect intensifies if the rate of adaptation increases during the course of an infection, for example caused by an increased rate of generation of mutations. Nevertheless, the changes in the pathogen's genome ultimately need to lead to an increased fitness measured in between-host transmission to be beneficial to emergence. This includes any adaptations to within-host processes such as the host immune system. They increase the pathogen's fitness within a single host, but are only relevant if they also increase the number of secondary infections.

Regardless of the evolutionary course, introductions lead in most cases to an extinction with a finite outbreak size (Antia et al., 2003; Arinaminpathy and McLean, 2009). The outbreak size depends on the reproductive numbers of the wildtype and intermediate strains. For each strain, the average outbreak size is given by $1/(1 - R_0)$ for $R_0 < 1$ (Becker, 1981). While it is possible to predict this characteristic, the fate of any given introduction is highly unpredictable (Arinaminpathy and McLean, 2009). Conversely, this means that even a large number of failed introductions neither rules out the possibility of a successful adaptation and emergence, nor does it allow accurate estimates of the probability of emergence. Figure 2.3 shows the upper bound on the probability of emergence per introduction given N unsuccessful introductions. The gain in information on the probability of emergence is especially small for large numbers of failed introductions (Arinaminpathy and McLean, 2009). Further, the stochastic nature of disease transmission paired with the high uncertainty in the probability of emergence lead to poor ability to accurately forecast the behaviour of ongoing outbreaks.

These homogeneous models have implications for public health. A warning system for EIDs would need to be capable of differentiating between an emergence or extinction at low numbers of infectious hosts to be effective (Longini et al., 2005; Ferguson et al., 2005; Wallinga et al., 2010). It would be possible to design 'safe systems' which are able to warn before almost all emergences, but only at the cost of many false alarms. Such over-sensitive

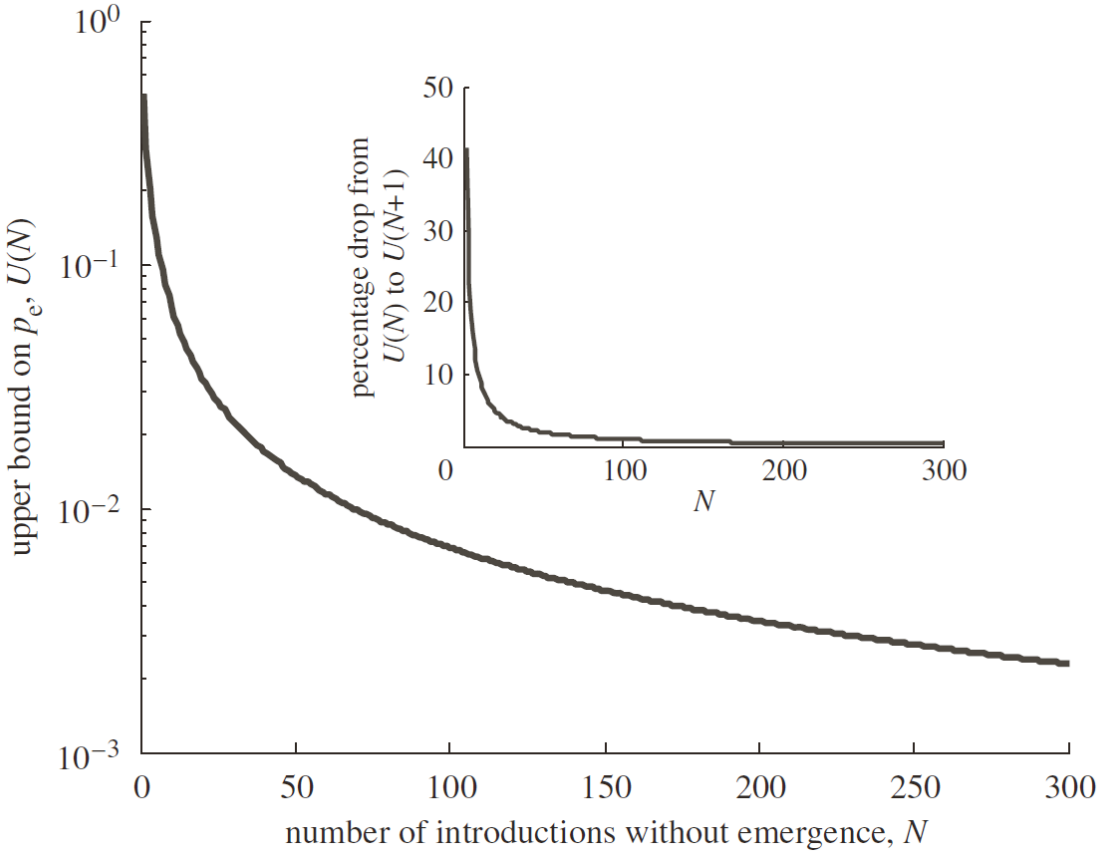

Figure 2.3: Upper bound on the probability of emergence $p_{\text{em}} = p_e$ given that N introductions have occurred without emergence $U(N)$. Calculated as the maximal value of p_e giving at least a 50 per cent probability of observing N introductions without emergence. Although giving infinitesimally small values for p_e for large N, the curve does not reach $p_e = 0$ for any finite N. Inset shows the relative information gain acquired as N increases, measured as the percentage drop from $U(N)$ to $U(N+1)$. Source: Arinaminpathy and McLean (2009).

systems would soon be rendered useless as the public would not take the warnings seriously. Unfortunately, most simple warning system do not fulfil the necessary criteria and produce too many false alarms, or are only able to give accurate warnings at very late stages of the emergence where public health measurements like quarantining would be infeasible to achieve (Arinaminpathy and McLean, 2009). For example, it has been estimated that in the case of a possible novel avian influenza emergence, interventions such as household quarantine and anti-viral treatment need to be initiated at most 21 days after the first case (Longini et al., 2005).

Another, previously neglected, possibility is 'two-way' between-species transmissions. Reluga et al. (2007) look at the importance of transmission back into the animal population during the emergence process. They use a branching process model similar to the one previously used by Antia et al. (2003), but incorporate subsequent species-jumps between the recipient and donor species. Their results indicate that if between-species jumps are rare, or the transmissibility of the already partly-adapted strain is low within the donor species, the probability of a successful emergence converges to the systems without multiple between-species jump. On the other hand, if these zoonotic jumps are sufficiently common and non-wildtype pathogen transmission within the donor species is significant, the probability of emergence can increase significantly. Influenza might be an example for this as seen in the genetic origins of the influenza A(H1N1) pandemic strain (Smith et al., 2009).

2.6.2 The Effect of Heterogeneity on the Emergence Process

The previous section reviewed studies of zoonotic emergences in unstructured human host populations. Every randomly picked infectious host has the same chance of infecting every randomly picked susceptible host within the whole population - as in all basic compartmental or branching process models. It follows directly that every host has the same offspring probability distribution. Variations in the number of offsprings are only caused by stochastic fluctuations. Nonetheless, a homogeneous host population is only an approximation. As discussed in section 2.5, many studies have shown that humans form

very heterogeneous population structures with their social contacts (Watts and Strogatz, 1998; Eubank et al., 2004; Read et al., 2008), economic networks (Faust, 1999), and travel behaviour (Hufnagel et al., 2004; Brockmann et al., 2006; Balcan et al., 2009).

But what happens if the host population is not homogeneous? What is the effect on the offspring distribution? What is the effect on the emergence of a disease? What are the effects on the epidemiological parameters? Here, I will address these questions by reviewing the published literature.

Host heterogeneity describes some characteristics of the uniqueness of every human host. Models can incorporate heterogeneity by using parameters like age, genetic diversity, or spatial host distribution. Heterogeneous models are relevant if the outcome is highly dependent on the underlying heterogeneity. For example, this is the case for infectious diseases when the risk of serious illness, or even death, depends on the host's age at infection. Rubella is a well known infectious disease of this kind (Anderson and May, 1992). In general, the pathogen causes mild infections. In pregnant women, however, it might cause the severe congenital rubella syndrome in the offspring. Hence, models of vaccination strategies should incorporate this heterogeneity in sex and age as only women in a specific age group tend to get pregnant.

Another type of heterogeneity is the host contact distribution. It is a part of spatial heterogeneity as seen in network models (see section 2.5.2). In general, it appears to be a common property of infectious diseases that only a small proportion of the population generates most infections as seen with SARS (Meyers et al., 2005; Skowronski et al., 2005). Figure 2.4 gives an overview of this heterogeneity for several diseases. In general, this effect is called superspreading (see section 2.5.2 for more details on superspreaders). Lloyd-Smith et al. (2005) estimate the effect of host heterogeneity on disease emergence of a fully-adapted strain. Their main findings show that heterogeneity decreases the probability of an emergence, but increases the tempo of an emergence once it happens.

Yates et al. (2006) extend these results to pathogen evolution, and incorporated host contact heterogeneity using previously published adaptation models. They incorporate heterogeneity in four factors: infectivity, mixing, susceptibility, and frequency of contacts. These factors determine the reproductive number R_0 for each host. They define 'types' of

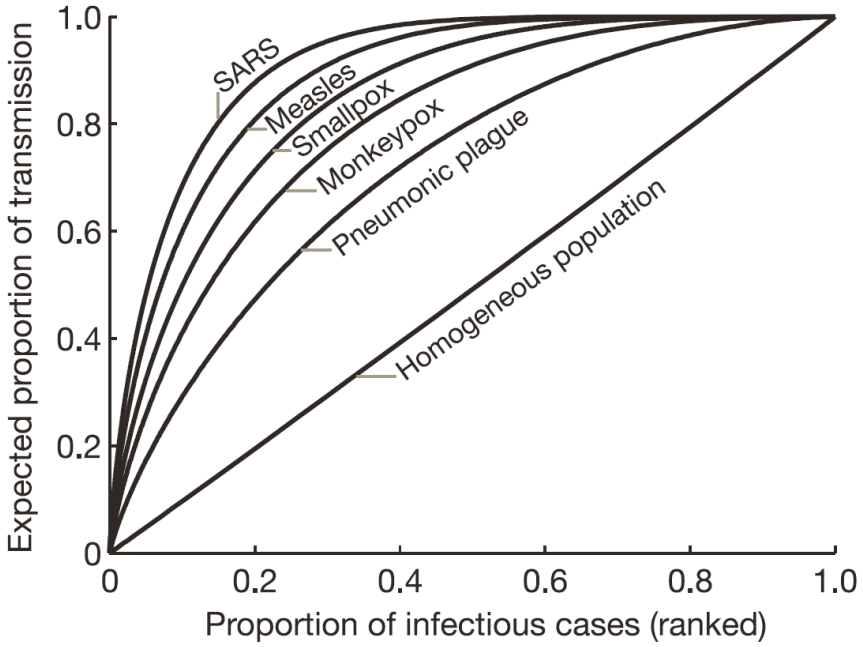

Figure 2.4: Variation in the infectivity of some diseases. Expected proportion of all transmission due to a given proportion of infectious cases, where cases are ranked by infectiousness. For a homogeneous population, this relation is linear. For five directly transmitted infections, the line is concave owing to heterogeneity in disease transmission. Source: Lloyd-Smith et al. (2005).

hosts with type-specific R_0 according to the different properties in the four factors. These types are called: *assortative mixing*, where each infectious host primarily infects hosts of the same type; *dissortative mixing*, where each infectious host primarily infects hosts of a different type; *heterogeneity in infectivity*, which models superspreading by allowing a small proportion of the population to be 20 times more infectious than the rest; and *homogeneous*, which represents a homogeneous populations with one single R_0. As with Lloyd-Smith et al. (2005), host heterogeneity always reduces the probability of emergence. The same is mostly true for an adapting pathogen (see Figure 2.5). While heterogeneity usually lowers the probability of emergence, dissortative mixing can slightly enhance the probability of emergence if superspreaders are also present.

These findings have an important impact on public health measures. Identification of high-risk groups can drastically reduce the stockpile of medications needed for treatment (Yates et al., 2006), and reducing the R_0 of superspreaders can drastically reduce overall disease transmission (Keeling and Eames, 2005; Meyers et al., 2005).

Spatial heterogeneity can be subdivided according to its scale. Host populations can be heterogeneous in the host contact structure on the level of their neighbours, or on a community level with many connections within communities and only few between them.

The first one, contact structure on a small scale, represents heterogeneity in the number of contacts each individual host has. Therefore, it also determines the number of secondary infections. Such host populations have multiple offspring probability distributions while in homogeneous populations variation in the number of offspring is a random process with one underlying offspring probability distribution. For example, network models incorporate host contact heterogeneity through their degree distribution (see section 2.5.2). This degree distribution is the crucial point for the number of offsprings. An infectious host with two contacts is not able to infect four other hosts, and an infectious host with 100 contacts is very unlikely to not infect any of its neighbours. Especially local spatial contact structures seem to have important implications for the fate of an emergence (Keeling, 1999; Yates et al., 2006).

Previously, pathogen evolution on networks has only been studied in general (Read and Keeling, 2003), but Alexander and Day (2010) join a branching-process model of zoo-

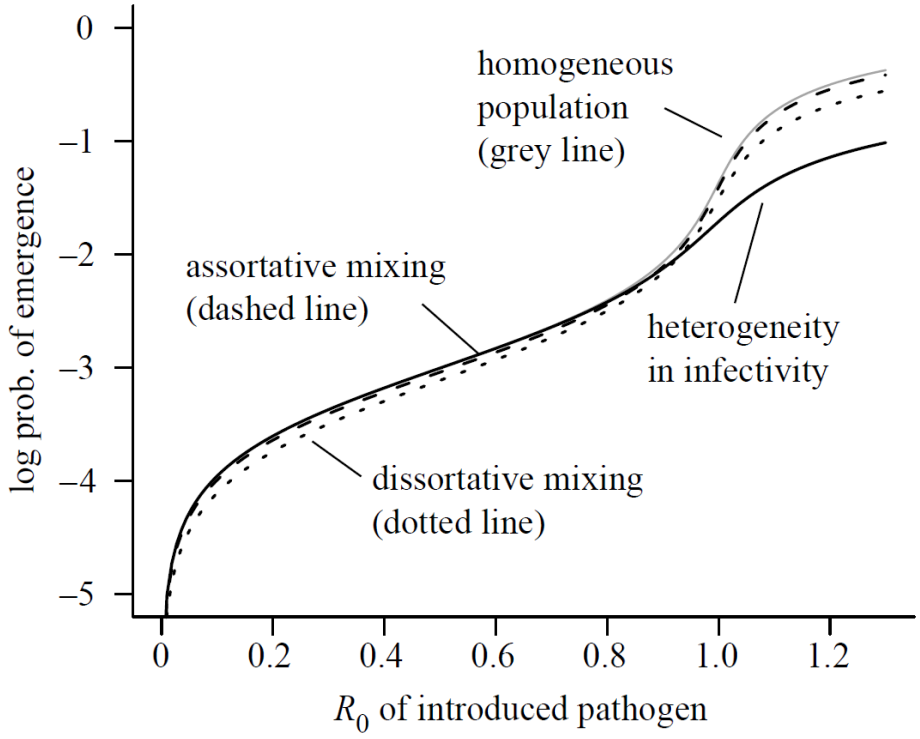

Figure 2.5: Probability of one-step pathogen evolution and emergence occurring in the presence of host heterogeneity as derived by calculations. The grey line is the base line with the homogeneous case. To model heterogeneity, the population is split in a 9 : 1 ratio with the smaller group being 100 times more susceptible to the pathogen. The dashed line represents assortative mixing, where hosts are more likely to have contact with others from the same group, and the dotted line is dissortative mixing, where hosts are more likely to have contact with hosts from the other group. The solid black line represents a different form of heterogeneity. Here, the 9 : 1 split remains but the smaller group is 20 times more infectious. This type of heterogeneity diverges most, for larger R_0, compared to the base line. Source: Yates et al. (2006).

notic emergence with a network model to study zoonotic emergence on networks. They explore the effect of different network structures with two different adaptation routes: a one-way route without back-mutation, and a two-way route where back-mutations are possible. Their findings suggest that back-mutations usually have little influence on the overall emergence process, except if the strain transmissibility is near the critical threshold equivalent to $R_0 = 1$. This result is in accordance with findings for homogeneous populations (see section 2.6.1 and (Reluga et al., 2007)). However, larger adaptation steps, for example through simultaneous mutations, can have a drastic effect on the probability of emergence. This is in line with previous results leading to the notion that the emergence process can be significantly increased if the order of the mutations does not matter (Gokhale et al., 2009).

Furthermore, the approach of Alexander and Day (2010) allows comparison of different host contact structures through different degree distributions. Their study yields two main insights: first, the host contact structure can lower or increase the necessary transmission potential of the novel pathogen. A characteristic feature of networks is that one basic reproductive number might yield significantly different transmission probabilities between two connected hosts if the network types are different. Second and in contrast to Lloyd-Smith et al. (2005), host heterogeneity decreases the probability of disease extinction.

An explanation for this discrepancy can be found in the respective model structures. While Lloyd-Smith et al. (2005) model contact heterogeneity purely as a function of the secondary infections, networks incorporate heterogeneity on an incoming and outgoing level. A host with a large number of contacts not only has a higher chance of infecting more contacts if infectious, but is also at higher risk of getting infected than a host with a smaller degree.

As outlined in section 2.1, it is important to choose the appropriate mathematical model for the right type of heterogeneity. Effectively, network models are hard to handle if the aim is to incorporate both large and small scale spatial heterogeneity. It is possible to incorporate large scale spatial heterogeneity into networks through clustering. But networks will always also show the effect of the local contact structure, too.

These examples from the literature on zoonotic EIDs show the necessity to define which kind of heterogeneity is relevant to the proposed research question. Accordingly, a model needs to be designed in a way that represents the relevant heterogeneity, while keeping unnecessary complexity to a minimum.

2.7 Does the Adaptive Process Matter?

All the studies of zoonotic EIDs reviewed here model evolutionary adaptations to the recipient physiology in a very simplistic way. No model specifies the actual mutations on a within-host level, therefore neither describing the exact changes in geno- or phenotype (see section 1.4), nor the process of achieving the mutations (see section 1.4.2). The reasoning behind this simplistic approach lies clearly in the constraints of complexity. However, it raises the question of realism, and if this might have an effect on the validity and relevance of the models.

For example, avian and human influenza have the ability to produce novel strains by viral reassortment with swine as the most prominent possible mixing vehicles (Webby et al., 2004a), because swine share receptors in the upper repository tract with humans and birds, showing intermediate susceptibility to human and avian influenza. Indeed, the pandemic influenza A(H1N1) strain of 2009 is a triple-reassortant virus with a genome consisting of human, swine, and avian influenza parts (Dawood et al., 2009).

Do these findings invalidate the theoretical models, which assume fixed mutation steps measured in the basic reproductive number? I argue that this approach does not invalidate the results. As previously stated (see section 2.6), only mutations beneficial to fitness, and therefore the reproductive number, can ultimately lead to an emergence. It is irrelevant if this mutation happened by antigenic shift or drift, reassortment or recombination. In all cases, it is an event which leads to a possible change in the phenotype with a certain probability. And that is exactly what the theoretical models describe. Hence, these models are, with some modification like the incorporation of vectors or mixing vessels, very flexible in their application.

A more important question is the actual route of adaptation; that is, the fitness gain

per adaptive steps. While few studies have looked at the optimal speed and trajectory of evolution (Traulsen et al., 2007; Gokhale et al., 2009), our knowledge is, in general, limited.

2.8 Gaps In Our Knowledge

The existing literature yields fascinating insights into the evolution and emergence of zoonotic infectious agents. However, there are still many aspects of zoonotic EIDs which we do not understand yet. For example, previous studies such as Antia et al. (2003) do not incorporate any spatial host heterogeneity. Furthermore, it has been unclear how to identify or even stop a previously unknown EID at its beginning, even though first steps have been taken in this direction (Arinaminpathy and McLean, 2009). Another unresolved question is how environmental changes can influence the risk of an emergence, including ecological changes through intensified agricultural land use or mass farming, as well as such recent threats as global climate change. I will address some of these unknown aspects with four main research questions:

(i) How can zoonotic EIDs come 'out of nowhere'?

(ii) What environmental conditions can facilitate inter-species jumps of an infectious disease?

(iii) Does a change in the species mixture facilitate zoonotic EIDs?

(iv) What was the impact of an exemplar novel infectious disease in a human population?

Each of my research projects presented in this thesis concentrates on at least one of these questions in an attempt to increase our overall understanding of how zoonotic pathogens can evolve and emerge into the human host population.

CHAPTER 3

MODELLING THE ZOONOTIC EMERGENCE OF A NOVEL INFECTIOUS DISEASE

3.1 Introduction

As discussed in chapter 2, previous work (Antia et al., 2003; Arinaminpathy and McLean, 2009) has studied models of within-host evolution and between-host transmission in which an initially poorly transmitting pathogen acquires adaptations to human hosts, following repeated zoonotic introductions until it achieves pandemic potential. These make the natural, simplifying assumption that the host population is homogeneous, and changes in epidemiological parameters entirely reflect adaptations in the biology of host infection. In reality, however, factors such as human contact patterns (Woolhouse et al., 2005) and other host heterogeneity (Lloyd-Smith et al., 2005; Yates et al., 2006) may also shape the risk and tempo of emergence events. We concentrate here on the first of these factors, an area which has hitherto received little attention in the context of zoonotic emerging pathogens. Most of the results in this chapter have been published as Kubiak et al. (2010).

Our aim is to develop a mathematical model to explore under what regimes such 'ecological' structures could have an effect on emergence, and include 'evolutionary' factors governing the biology of infection. As reviewed in chapter 2, various types of models exist which include host heterogeneity. Here, we focus on the impact of large scale spatial heterogeneity on the adaptation of a novel, zoonotic disease.

In the following sections we give an overview of the modelling approach. We then present new analytical results for the simple models studied previously, which ignored spatial host population structure. We use these expressions to answer the following questions: if we knew how a zoonotic pathogen would adapt to human physiology, could we anticipate its emergence? How reliable would such predictions be? Furthermore, can we predict which zoonoses will cause outbreaks which do not turn into epidemics? Next, we ask: how large does a single, finite host population have to be, for population size to have a negligible effect? We then incorporate spatial heterogeneity by separating the human host population into communities. We present a model in which a small village is connected, by human travel, with a large city as an example of the general case of two interconnected communities. We use this model to ask: how strong do community interconnections have to be for us to safely ignore the separation of a population into spatially structured com-

munities, such as cities and villages? We then review available commuting data to ask how these thresholds compare with typical human mobility patterns? We close with a discussion of the implications of these findings for public health monitoring systems.

3.2 Modelling Evolutionary Adaptation

As discussed in section 1.4, many EIDs need to adapt to their novel recipient species. While the necessary adaptations might be different in likelihood and locus of effect (such as cell entry or replication within the cell), they eventually increase fitness within the novel host. One of the most useful ways to measure pathogen fitness is the basic reproductive number as it describes the bottleneck of reproduction - the transmission process (see section 2.2).

Here, we build on a model of evolution and emergence originally presented by Antia et al. (2003) in which a zoonotic pathogen infects humans, and initially has very poor onward transmissability. Thus for people who are infected by animals the basic reproductive number, R_0, is well below one ($R_0 \ll 1$). We call this the first reproductive number α_1 for the wildtype strain. Occasionally, during such zoonotic infections, the pathogen acquires genetic changes that increase its ability to pass to other humans. During any chain of transmission the pathogen might adapt sufficiently that it achieves such ease of human-to-human transmission that $R_0 > 1$ and an epidemic becomes possible. Such a process can be characterised by a vector of reproductive numbers (α_i with $i = 1, \ldots, n$), and a vector of mutation probabilities (μ_i with $i = 1, \ldots, n$) where $n - 1$ denotes the number of adaptive steps necessary to reach the fully adapted strain n.

In general, every introduction has only two possible outcomes: emergence or extinction. Extinction happens if the novel pathogen dies out, because it fails to adapt for human transmission or just by stochastic extinction. Hence, the introduction only leads to a limited number of infectious hosts, which we refer to as the 'outbreak size'. Conversely, a novel pathogen of zoonotic origin 'emerges' if it is sufficiently adapted for human transmission and begins to spread in a self-sustaining way. Formally, in an unlimited host population the cumulative number of infectious hosts is unbounded as time goes to infinity.

3.3 Theoretical Description of a Zoonotic Emergence

3.3.1 Probability of Emergence

In the special case $N \to \infty$ and $S/N \to 1$ it is possible to calculate the probability of emergence per introduction into the human host population, given the evolutionary course of the pathogen, and the mutation rate with which the pathogen adapts. Our derivations start with assuming one homogeneous human host population of infinite size. This assumption will be relaxed later. To calculate the probability of emergence, we define next event probabilities of infection p_i, mutation m_i, and recovery r_i for each individual infected host. Therefore, the probabilities for what type of event will come next for each infectious host are

$$\begin{aligned} p_i &= \frac{\alpha_i}{\alpha_i + \mu_i + 1} \\ m_i &= \frac{\mu_i}{\alpha_i + \mu_i + 1} \\ r_i &= \frac{1}{\alpha_i + \mu_i + 1} \end{aligned} \tag{3.1}$$

Note that in general, we can extend this adaptation process to arbitrarily many adaptive steps. The mutations are uni-directional towards the adapted strain. This is a constraint of using branching process formulations to model this system (see section 2.4 for more details). But this assumption is also biologically plausible, because only mutations which will eventually increase fitness can lead to an emergence (see section 1.4 for more details), hence all other mutations have only little or no influence on the probability of emergence per introduction.

Our derivations use a multi-type branching formulation to describe the extinction probability in the process of emergence. In general, our n-type branching process is given by n pgfs

$$G_i(s_i, \ldots, s_n) = \sum_{j_i, \ldots, j_n = 0}^{\infty} \phi_i(j_i, \ldots, j_n) s_i^{j_i} \ldots s_n^{j_n} \tag{3.2}$$

with $\phi_i(j_i, \ldots, j_n)$ being the probability that a single infected host of type i gives rise to j_k secondary infections with $k \in \{i, \ldots, n\}$.

We consider ϕ_i corresponding to a geometric distribution of type-specific infections. In particular, a given host infected with strain i, after causing j_i infections, will either recover or undergo a mutation, in the latter case initiating a new geometric process with strain $i+1$. Assuming independence of infection events, therefore, (3.2) yields the recursive relation

$$\begin{aligned} G_i(s_i,\ldots,s_n) &= \sum_{j_i=0}^{\infty} p_i^{j_i} s_i^{j_i} \left[r_i + m_i G_{i+1}(s_{i+1},\ldots,s_n)\right] \\ &= \frac{r_i + m_i\, G_{i+1}(s_{i+1},\ldots,s_n)}{1-p_i s_i} \end{aligned} \tag{3.3}$$

with $|s_i| < 1/p_i$. The probability of extinction q_i is defined as the probability that a single introduction of strain i causes only a finite outbreak size, ultimately to go extinct. A standard result of multi-type branching processes is that the q_is' constitute fixed points of the generating functions above (Athreya and Ney, 2004), that is

$$G_i(q_i,\ldots,q_n) = q_i \tag{3.4}$$

Substituting into (3.3), and recalling that there exist n progressively transmissible strains with $m_n = 0$, we obtain the system of equations

$$q_i = \begin{cases} \frac{r_i + m_i\, q_{i+1}}{1-p_i q_i} & i < n \\ \frac{1-p_n}{1-p_n q_n} & i = n \end{cases} \tag{3.5}$$

Note that this system has a trivial solution $q_i = 1$, and we seek the non-trivial solution $0 < q_i < 1$. Such a non-trivial solution exists as long as $p_n > \frac{1}{2}$, equivalently as long as the reproductive number of strain n exceeds 1 (i.e. $\alpha_n = p_n/(1-p_n) > 1$).

If all introductions are of wildtype strains, the probability of extinction for each is simply q_1. Solving (3.5) recursively, the probability of emergence per introduction is thus

$$p_{\text{em}} \equiv 1 - q_1 = 1 - \begin{cases} \sum_{x=1}^{n} r_x \prod_{y=2}^{x} \frac{m_{y-1}}{1-p_y q_y} & \text{if } p_1 = \alpha_1 = 0 \\ \frac{1}{2p_1} - \sqrt{\frac{1}{4p_1^2} - \sum_{x=1}^{n} \frac{r_x}{p_1} \prod_{y=2}^{x} \frac{m_{y-1}}{1-p_y q_y}} & \text{otherwise} \end{cases} \tag{3.6}$$

This probability can thus be evaluated explicitly for any given route of adaptation. Figure 3.1 illustrates the probability of emergence per introduction for a fixed number of strains, $n = 3$.

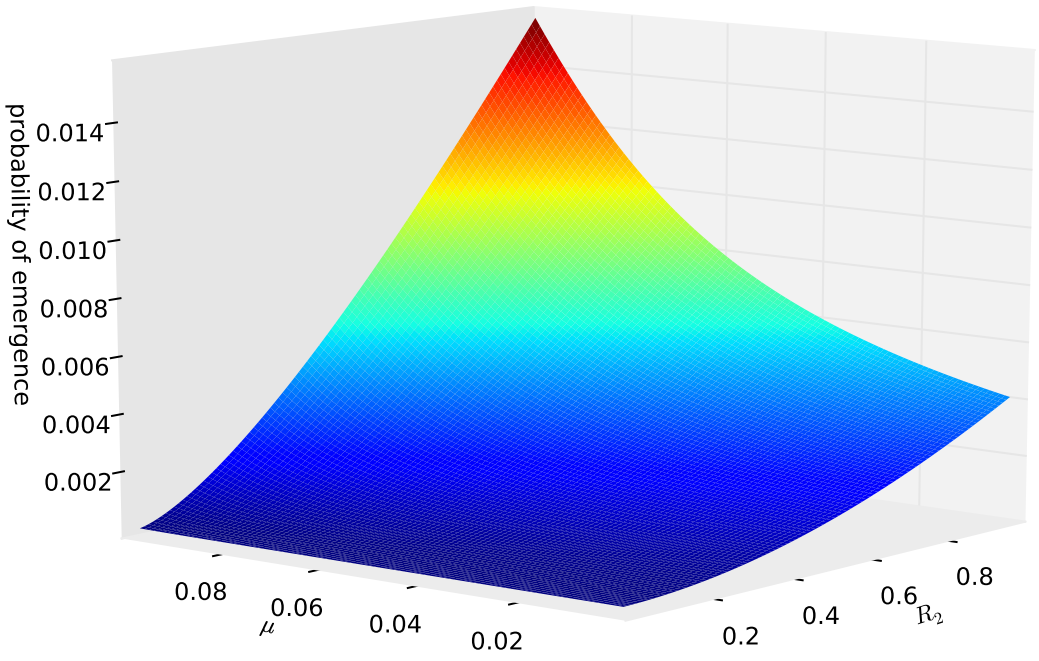

Figure 3.1: Evaluation of the probability of emergence per introduction for a three strain model. Shown is the probability of emergence as a function of the mutation rate μ and the intermediate reproductive number $R_2 = \alpha_2$ as given by equation (3.6). The reproductive numbers of the wildtype strain and fully adapted strain are fixed, $\alpha_1 = 0$ and $\alpha_3 = 2$.

3.3.2 Waiting Time to Emergence

Let the probability of emergence per introduction be p_{em}. Hence the probability of having an emergence after M introductions is

$$\begin{aligned} P_{\text{em}}(M) &= (1 - p_{\text{em}})^M \, p_{\text{em}} \\ &= q_{\text{ext}}^M \, p_{\text{em}} \end{aligned} \tag{3.7}$$

with $q_{\text{ext}} = 1 - p_{\text{em}}$ as probability of extinction per introduction. To derive the mean waiting time and its standard deviation, we can make use of standard properties of pgfs (see section 2.4.1)

$$\begin{aligned} G_M(z) &= \sum_{M=0}^{\infty} P_{\text{em}}(M) z^M \\ &= p_{\text{em}} \sum_{M=0}^{\infty} \left(q_{\text{ext}}\, z\right)^M \\ &= \frac{p_{\text{em}}}{1 - q_{\text{ext}}\, z} \end{aligned} \tag{3.8}$$

with $|z| < 1/q_{\text{ext}}$.

It follows according to equation (2.12) for the average waiting time

$$\begin{aligned} \langle M \rangle = G'_M(1) &= \left[\frac{d}{dM} G_M(z)\right]_{z=1} \\ &= p_{\text{em}} \sum_{M=0}^{\infty} M q_{\text{ext}}^M \\ &= p_{\text{em}} \frac{1 - p_{\text{em}}}{p_{\text{em}}^2} \\ &= \frac{1}{p_{\text{em}}} - 1 \end{aligned} \tag{3.9}$$

Note that this is the average number of introductions before an emergence. The average number of introductions for an emergence to occur is $\langle M \rangle + 1$.

The variance can be obtained in a similar way, following equation (2.13)

$$\begin{aligned} \text{var}(M) =& G''_M(1) + G'_M(1) - \left[G'_M(1)\right]^2 \\ =& \left[p_{\text{em}} \sum_{M=0}^{\infty} M(M-1) q_{\text{ext}}^M z^{M-2}\right]_{z=1} + \left[p_{\text{em}} \sum_{M=0}^{\infty} M q_{\text{ext}}^M z^{M-1}\right]_{z=1} \\ & - \left(\left[p_{\text{em}} \sum_{M=0}^{\infty} M q_{\text{ext}}^M z^{M-1}\right]_{z=1}\right)^2 \\ =& \frac{1}{p_{\text{em}}^2} - \frac{1}{p_{\text{em}}} \end{aligned} \tag{3.10}$$

The standard deviation is thus of the same magnitude as the mean, making the number of actual introductions before emergence highly unpredictable. Figure 3.2 shows the av-

erage waiting time and its standard deviation, illustrating the uncertainty in emergence processes.

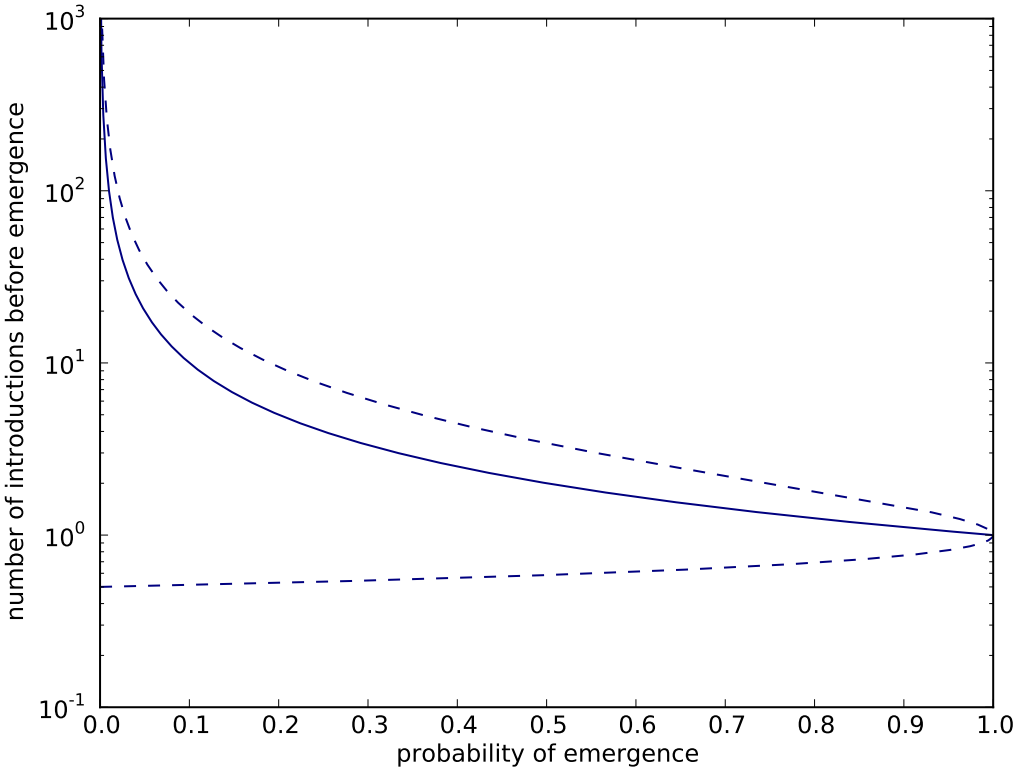

Figure 3.2: Average waiting time to emergence as a function of the probability of emergence. The solid blue line represents the average waiting time before an emergence happens. The dashed blue lines mark the interval of one standard deviation around the average. It is obvious that the waiting time is very unpredictable as the variation around the mean is high, especially for small probabilities of emergence.

This result is remarkable. It is independent of the actual probability of emergence, as well as the kind and number of necessary evolutionary adaptations. It is therefore theoretically impossible to improve our predictions about an emergence event above a certain point even if we had complete knowledge about the necessary adaptations. Zoonotic emergences will always be a highly unpredictable danger.

3.3.3 Comparing Theoretical Expectations with Data

Equations (3.9) and (3.10) allow us to define the expected number of introductions before an emergence finally occurs. But they do not reveal the number of expected emergence events given a fixed number of introductions. Such knowledge is especially useful for comparing data, for example from an experiment with a bacteria-phage model system, with a theoretically predicted $p_{\text{em}}^{(\text{theo})}$. Obviously, such experimental data can also be produced with stochastic simulations.

Imagine a system with n_{ex} introductions of which x_{ex} have successfully lead to an emergence. Using (3.9), one would expect

$$\langle n_{\text{theo}} \rangle = \frac{x_{\text{ex}}}{p_{\text{em}}^{(\text{theo})}} \tag{3.11}$$

introductions to get on average x_{ex} successful emergence events.

Because emergence events are independent and identically distributed random events, the process itself is a Bernoulli process. This knowledge, and the use of equation (3.10), leads to the following standard deviation for the number of introductions

$$\sigma(\langle n_{\text{theo}} \rangle) = \sqrt{x_{\text{ex}} \frac{1 - p_{\text{em}}^{(\text{theo})}}{p_{\text{em}}^{(\text{theo})\,2}}} \tag{3.12}$$

This expected deviation in the number of introductions allows for comparison between experimental data and our theoretical model. Furthermore, it highlights the previously mentioned result of the large variance in possible outcomes, making the behaviour of an EID effectively unpredictable.

3.3.4 Outbreak Size Distribution

An important, simplifying feature of our branching-process approach is that susceptible hosts are essentially unlimited: we can thus treat transmission chains as developing independently of one another in time. In particular, suppose that during an outbreak there are $I_{i,0}$ introductions of strain i, whether exogenously (e.g. as zoonoses) or through mutation

of strain $i-1$. For a specified $I_{i,0}$, the probability distribution of cumulative cases of strain i is equivalent to that of a stochastic 'birth/death' process, having $I_{i,0}$ initial cases. In our model, the birth process is equivalent to a transmission with probability p_i, and a death process is equivalent to an effective recovery with probability $r_i + m_i$. Note this latter term represents the fact that exit from infected status with strain i occurs either through actual recovery, or through mutation. Ultimately, however, we must also account for the fact that $I_{i,0}$ is a random variable, itself conditional on the cumulative cases of strain $i-1$. In so doing we arrive ultimately at a probability distribution for cumulative cases of strain i, conditional on those for cumulative cases of strain $i-1$, and so on. Given the initial condition of the introduction of w zoonotic infections with the wildtype pathogen, this approach ultimately yields a well-defined probability distribution for the overall outbreak size.

Adopting the simplified, equivalent process described above for strain i, and assuming $I_{i,0}$ to be specified, suppose that extinction occurs following precisely x_i 'birth' events (equivalently, the total incidence of strain i). It follows that there must also have been

$$N_i = I_{i,0} + x_i \tag{3.13}$$

'death' events (equivalently, the cumulative number of cases of strain i). Thus we have, for the conditional probability distribution of N_i,

$$g_i(I_{i,0}, x_i) = T(I_{i,0}, x_i)\, p_i^{x_i}\, (r_i + m_i)^{N_i} \tag{3.14}$$

where $T(I_{i,0}, x_i)$ enumerates all possible epidemic trajectories involving $I_{i,0}$ zoonoses, and a cumulative incidence of x_i. To be precise, consider representing such 'trajectories' on a two-dimensional lattice with co-ordinates (a, b), such that a northward step represents infection, a, and an eastward step represents recovery, b. In our case, $T(I_{i,0}, x_i)$ essentially counts all lattice paths connecting $(I_{i,0}, 0)$ to $(I_{i,0} + x_i, I_{i,0} + x_i - 1)$, and never touching the line $a = b$ (since an epidemic is over if, and only if, this line is reached). $T(I_{i,0}, x_i)$ is thus

given by the ballot theorem (Renault, 2008), yielding

$$T(I_{i,0}, x) = \binom{2x_i + I_{i,0} - 1}{x_i + I_{i,0} - 1} - \binom{2x_i + I_{i,0} - 1}{x_i + I_{i,0}} \tag{3.15}$$

which may now be used in the evaluation of (3.14).

We now incorporate the distribution of $I_{i,0}$ as a random variable, returning to the context of the full model with mutations. Assuming that all introductions into the host population are of wildtype pathogens, then for $i \geq 2$, $I_{i,0}$ consists entirely of mutations from cases of strain $i-1$. As such it is drawn from the N_{i-1} cumulative cases of strain $i-1$, each individual case contributing a mutation with probability $\frac{m_{i-1}}{m_{i-1}+r_{i-1}}$. In other words $I_{i,0}$ with $i \geq 2$ follows a binomial distribution

$$\zeta_i(I_{i,0}) = \binom{N_{i-1}}{I_{i,0}} \left(\frac{m_{i-1}}{m_{i-1}+r_{i-1}}\right)^{I_{i,0}} \left(\frac{r_{i-1}}{m_{i-1}+r_{i-1}}\right)^{N_{i-1}-I_{i,0}} \text{ with } 0 \leq I_{i,0} \leq N_{i-1} \tag{3.16}$$

The fixed boundary condition $I_{1,0} = k$ is expressed with

$$\zeta_1(I_{1,0}) = \begin{cases} 1 & \text{if } I_{1,0} = k \\ 0 & \text{otherwise} \end{cases} \tag{3.17}$$

Now, an outbreak of total size X may be realised by any sequence of values $\{x_1, \cdots x_n\}$ summing to $X-k$. Combining (3.16) and (3.17) with (3.14), we thus find for the probability distribution of the outbreak size X

$$p_{\text{out}}(X) = \sum_{x_i, I_{i,0}} \prod_{i=1}^{n} \zeta_i(I_{i,0})\, g_i(I_{i,0}, x_i) \tag{3.18}$$

with k prescribed through (3.17) as a boundary condition and the summation taken over all $I_{i,0} \leq N_{i-1}$ with $i \geq 2$ and x_i's such that $\sum_{i=1}^{n} x_i + k = X$. As an aside, the probability of emergence follows simply as

$$p_{\text{em}} = 1 - \sum_{X \geq I_{1,0}} p_{\text{out}}(X) \tag{3.19}$$

which equals equation (3.6).

3.4 A Model System with Three Adaptive Stages

In what follows we restrict our attention to the case of $n = 3$, $\mu_{1,2} = \mu$, and $\mu_3 = 0$, allowing us to model two routes of adaptation with opposite and distinct characteristics, while minimising the overall complexity in number of required strains. For both routes of adaptation, the first, wildtype strain has very low transmissibility, the third has pandemic potential, and the second strain has intermediate transmissibility. This intermediate transmissibility is not enough to sustain the novel pathogen within the human host population, but secondary infections by humans are possible. Thus we have $\alpha_1 < \alpha_2 < 1$. Finally, the human adapted strain has a reproductive number $\alpha_3 > 1$. Further, we assume an identical mean infectious period for all strains.

Following Arinaminpathy and McLean (2009), we first distinguish two routes to adaptation: the 'punctuated' scenario has an evolutionary course $\alpha_i = [0, 0.1, 2]$, while the 'gradual' scenario has $\alpha_i = [0, 0.9, 2]$, the only difference being α_2, the fitness of the intermediate stage. Table 3.1 gives an overview of the adaptive routes.

evolutionary route	$\alpha = R_0$	μ
(I) punctuated adaptation	$[0, 0.1, 2]$	$[\mu, \mu, 0]$
(II) gradual adaptation	$[0, 0.9, 2]$	$[\mu, \mu, 0]$

Table 3.1: Parameters of used evolutionary routes.

This leads to the following SIR-model, normalised with respect to the mean infectious period,

$$\begin{aligned} \frac{dS}{dt} &= -\sum_{i=1}^{3} \alpha_i I_i \frac{S}{N} \\ \frac{dI_i}{dt} &= \alpha_i I_i \frac{S}{N} + \mu_{i-1} I_{i-1} - (\mu_i + 1)\, I_i \\ \frac{dR}{dt} &= \sum_{i=1}^{3} I_i \end{aligned} \tag{3.20}$$

where S is the number susceptible, I_i is the number infected with strain i, and R the num-

ber of recovered or removed. We do not include births and deaths as we expect a zoonotic emergence, or extinction, to happen on a much shorter timescale than the human lifespan. We translate this model to a stochastic simulation of a multi-type branching process, using the Gillespie algorithm (Gillespie, 1977). The infection is seeded in a single random, susceptible host with the wildtype strain.

Computationally we use a threshold of $I_3 = 100$ infectious hosts with the fully adapted strain to distinguish between emergence and extinction. This threshold ensures a probability of extinction less than $(1/\alpha_3)^{I_3} = 7.9 * 10^{-31}$ (May et al., 2001). Therefore, the number of falsely identified emergences, which would truly be extinctions, is negligibly small. Moreover, these arbitrarily small probabilities ensure that our simulation results are insensitive to the precise choice of threshold used. In situations where the host population is very small we relax our emergence threshold to smaller numbers of infectious hosts as some of the population sizes we consider are smaller than 100.

3.5 Incorporating Spatial Heterogeneity

3.5.1 Spatial Heterogeneity on Short Timescales

We use a meta-population model to explore the effects of spatial host heterogeneity, effectively dividing the human population into interconnected communities. As an example of the general case of spatially structured communities, we focus on a simple village - city model to approximate the spatial host heterogeneity in rural areas connected, by human mobility, to bigger cities. There are many different types of human mobility between communities such as villages and cities, including short-term commuting and long-term labour migration. Particularly in developing countries, however, information on dominant patterns is sparse. Nonetheless, empirical data from Vietnam (Thanh et al., 2005), for example, suggests that short-term commuting plays an important role: here, a subset of the village residents collects agricultural produce for trading in local markets in the city, and travels to the city on a daily to weekly basis. Accordingly we present a model in which the residence time of villagers in the city is typically less than the infectious period. In the next section, we relax this assumption and present a model which allows for longer commuting

lengths. These two different models illustrate that our results appear qualitatively robust to different types of human movement.

As before, we have a wildtype pathogen capable of acquiring adaptations for human transmission. Assume a finite number of hosts in the village, and an effectively infinite number in the city. To allow for daily commuting, we label individuals in the city according to whether they are commuters from the village or not (neglecting commuters originating from the city and present in the village). The superscripts $(v), (c)$ represent village and city inhabitants respectively, while (vm) denotes villagers in the city. Village members commute to the city at a per capita rate χ_1, and return at per capita rate χ_2. At any one time, a proportion w of villagers, the commuters, are in the city with w being set by χ_1 and χ_2 as described below. Further, we neglect susceptible village commuters acquiring infection in the city (this arises formally from the infinite number of hosts in the city). The number of village residents is fixed at $N = N^{(v)} + N^{(vm)} = 1000$, and we define the average number of commuters as

$$\langle c \rangle = N^{(vm)} = N * w \tag{3.21}$$

Figure 3.3 shows a schematic representation of this commuting model. Normalising time with respect to the mean infectious period, the governing equations are

Village, non-commuting:

$$\begin{aligned} \frac{dS^{(v)}}{dt} &= -\sum_{i=1}^{3} \alpha_i I_i^{(v)} \frac{S^{(v)}}{N^{(v)}} - \chi_1 S^{(v)} + \chi_2 S^{(vm)} \\ \frac{dI_i^{(v)}}{dt} &= \alpha_i I_i^{(v)} \frac{S^{(v)}}{N^{(v)}} + \mu_{i-1} I_{i-1}^{(v)} - (\chi_1 + \mu_i + 1)\, I_i^{(v)} + \chi_2 I_i^{(vm)} \end{aligned} \tag{3.22}$$

Village, commuting:

$$\begin{aligned} \frac{dS^{(vm)}}{dt} &= \chi_1 S^{(v)} - \chi_2 S^{(vm)} \\ \frac{dI_i^{(vm)}}{dt} &= \chi_1 I_i^{(v)} - \chi_2 I_i^{(vm)} + \mu_{i-1} I_{i-1}^{(vm)} - (\mu_i + 1)\, I_i^{(vm)} \end{aligned} \tag{3.23}$$

City:

$$\frac{dI_i^{(c)}}{dt} = \alpha_i \left(I_i^{(c)} + I_i^{(vm)} \right) + \mu_{i-1} I_{i-1}^{(c)} - (\mu_i + 1)\, I_i^{(c)} \tag{3.24}$$

Commuting equilibrium:

$$\begin{gathered} \frac{N^{(vm)}}{N^{(v)}} = \frac{\chi_1}{\chi_2} = \frac{w}{1-w} \\ \text{with } N = N^{(vm)} + N^{(v)} = \text{const.} \end{gathered} \tag{3.25}$$

Note that equation (3.25) arises from the fact that $w = N^{(vm)}/(N^{(v)} + N^{(vm)})$. To represent daily commuting, with an infectious period of 5 days, we set $\chi_2 = 5$ (i.e. average time in the city is one fifth of the infectious period), and choose w to give the required average number of commuters $\langle c \rangle$. To seed a wildtype infection in the village we set $I_1^{(v)} = 1$. In the village-city model, an emergence event is defined as having 100 infectious hosts with the fully adapted strain in the city.

3.5.2 Spatial Heterogeneity on Larger Timescales

As for commuting on a daily basis, we focus on a simple village - city model to approximate the spatial heterogeneity in rural areas only weakly connected to bigger cities. As before, we have a wildtype pathogen capable of acquiring adaptations for human transmission. Assume $N^{(v)} = 1000$ villagers, and an infinite number of hosts in the city. Moreover, assume a commuting rate of χ between city and village. We assume here that commuting is on a timescale longer than the infectious period, but may be temporary. Thus we assume that the return rate is the same as the outgoing rate, in order to maintain a constant number of hosts in the village. Figure 3.4 shows a schematic representation of the model.

For the village, we write: $S^{(v)}$ for the number of susceptibles, and $I_i^{(v)}$ for the numbers of infected with strain i. Correspondingly, $I_i^{(c)}$ is the number infected with type i in the city. Then, assuming that all strains have the same mean infectious period, and normalising continuous time with respect to this mean infectious period, an ODE model for this system

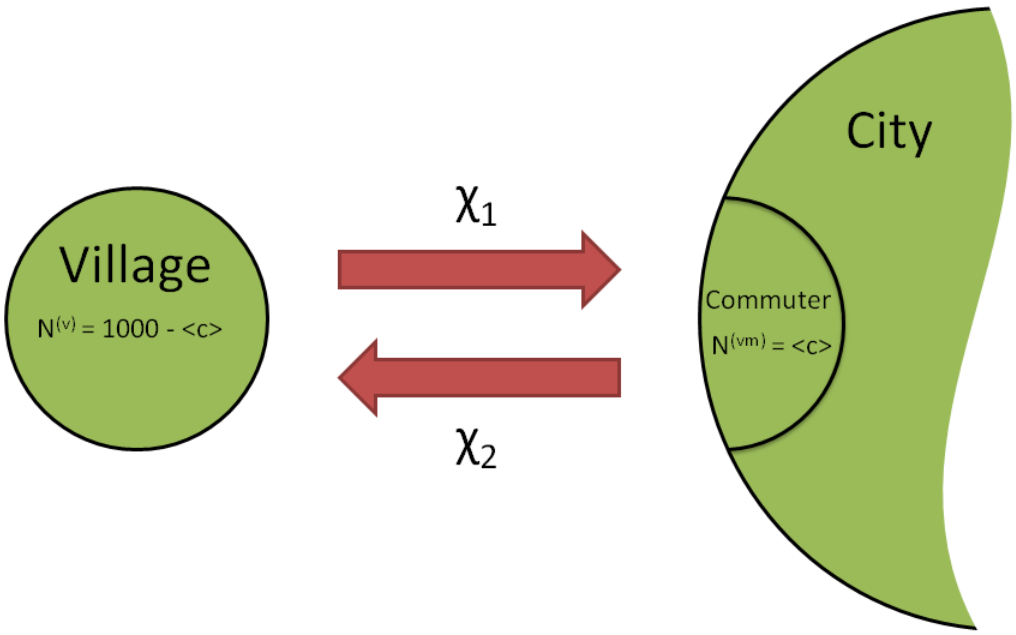

Figure 3.3: Schematic representation of the short-term commuting model. $N^{(v)}$ is the number of residents present in the village and $N^{(vm)} = \langle c \rangle$ the number of commuters in the city. The city has an infinite number of residents. χ_1 is the per capita commuting rate from the village to the city and χ_2 the return rate. Together, both determine the number of commuters $\langle c \rangle = N * w = N(\chi_2/\chi_1 + 1)^{-1}$.

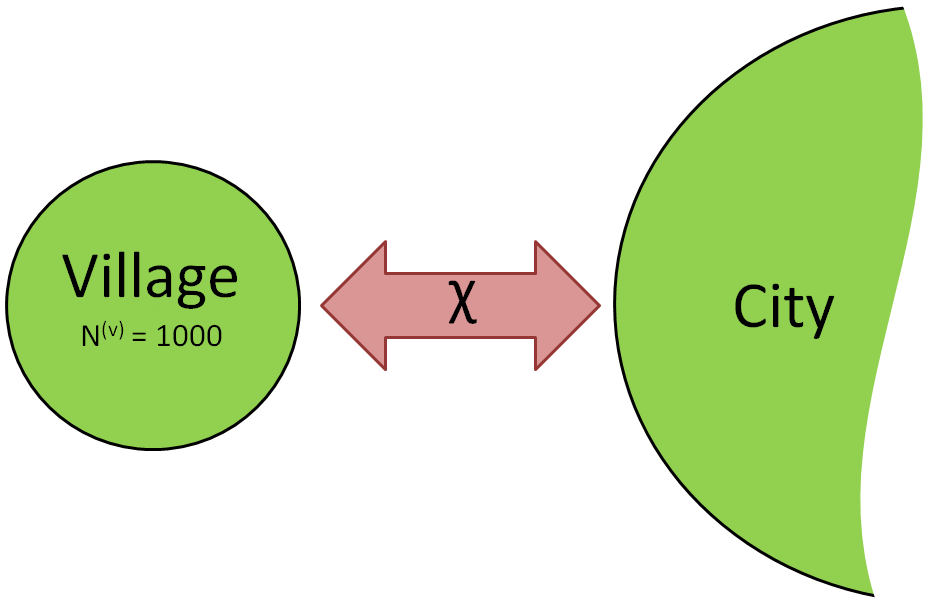

Figure 3.4: Schematic representation of the long-term commuting model. $N^{(v)}$ is the number of hosts in the village and χ the migration rate. Note the number of commuters from the village to the city is a function of the village size and the migration rate. The city has an infinite number of inhabitants.

reads

$$\begin{aligned}
\frac{dS^{(v)}}{dt} &= -\sum_{i=1}^{3} \alpha_i I_i^{(v)} \frac{S^{(v)}}{N^{(v)}} + \chi(N^{(v)} - S^{(v)}) \\
\frac{dI_i^{(v)}}{dt} &= \alpha_i I_i^{(v)} \frac{S^{(v)}}{N^{(v)}} + \mu_{i-1} I_{i-1}^{(v)} - (\chi + \mu_i + 1)\, I_i^{(v)} \\
\frac{dI_i^{(c)}}{dt} &= \alpha_i \left(I_i^{(c)} + I_i^{(v)} \right) + \mu_{i-1} I_{i-1}^{(c)} - (\chi + \mu_i + 1)\, I_i^{(c)}
\end{aligned} \tag{3.26}$$

where μ_i and α_i are, as before, respectively the mutation number and basic reproductive number associated with strain i. Our initial conditions are $I_1^{(v)} = 1$, and all other infected numbers zero, corresponding to seeding of a single wildtype infection in the village. An emergence is defined as before (total of 100 infected with the fully adapted strain in the city at one time).

3.6 Results

3.6.1 Validation of Analytical Results

We use the three strain model described before to study the impact of the mutation rate, μ, and the basic reproductive number of the intermediate strain, α_2, on the probability of emergence per introduction in a single, infinite population. We assume $0 < \mu \leq 10^{-1}$ as an illustrative spectrum of possible mutation rates. Figure 3.5 shows the probability of emergence for different mutation rates and basic reproductive numbers of the intermediate strain. Not surprisingly, the probability of emergence grows non-linearly with α_2 and μ. The probability of having no mutation in the second strain is $(1+\mu)^{-x}$ where x is the number of infected hosts with strain 2. While the intermediate reproductive number affects the exponent, the mutation rate has a direct influence on the base. We validate our analytical results by comparing them with the average probability of emergence of 10^3 simulated emergence processes, using one homogeneous population as described in (3.20). Figure 3.5 reveals an excellent agreement between the analytical results and simulations.

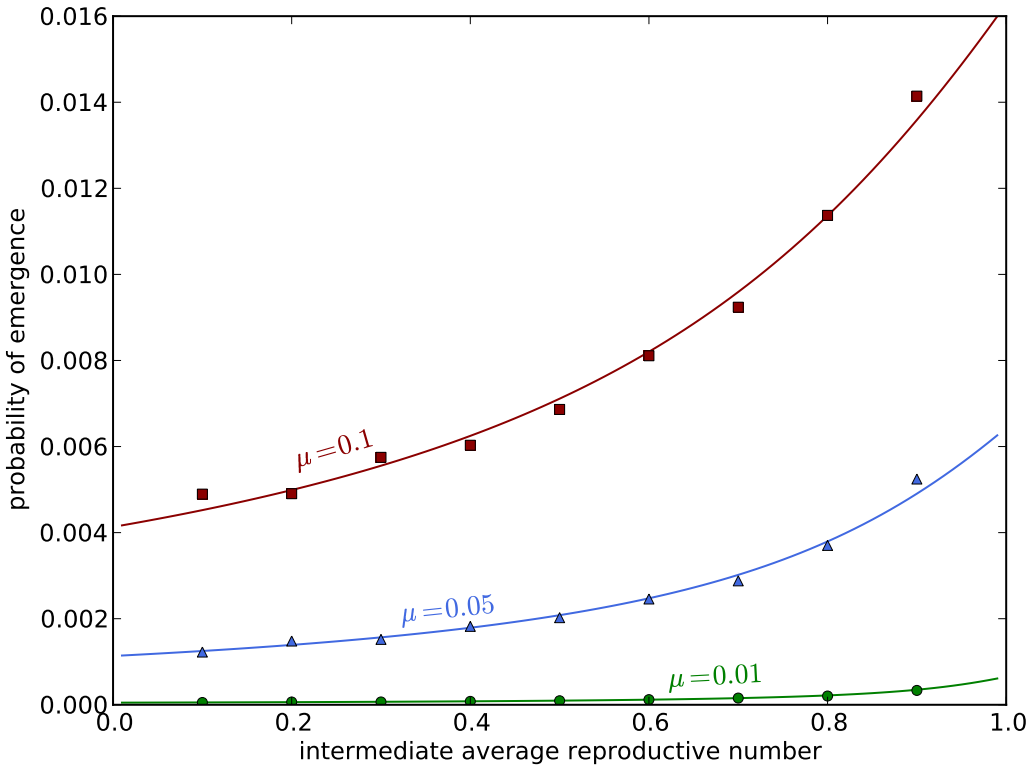

Figure 3.5: Comparison of the probability of emergence per introduction for the given routes of adaptation. The probability of emergence is a function of the mutation rate μ and the intermediate reproductive number α_2. The basic reproductive number for each strain is $\alpha_i = [0, \alpha_2, 2]$ and the mutation rate $\mu_i = [\mu, \mu, 0]$, $i \in \{1, 2, 3\}$. The solid lines are analytical results. The symbols represent the average probability of emergence over 1000 simulated emergences with a host population size of $N = 1000$. The agreement between analytical calculations and simulations is excellent. Furthermore, the probability of emergence grows non-linearly with α_2 and μ.

3.6.2 Effect of Limited Host Population Size

For small host communities, the depletion of susceptible hosts can play a significant role in limiting an ongoing outbreak. What is the effect of a finite population size on these analytic results which assume an infinite host population? Figure 3.6 compares the simulated outbreak size distribution of different sized populations with our analytical predictions. Note that, for populations greater than 500, there is close agreement between numerical and analytical results. When considering populations of size $1,000$ or more, we do not expect population size dependence to have a substantial effect.

3.6.3 Effect of Spatial Heterogeneity

The number of residents in the village ($N^{(v)} = 1000$) in our commuting model is sufficient to avoid finite size effects on the outbreak size. Furthermore, it is independent of any spatial heterogeneity. The number of residents only has a limiting effect on the outbreak size distribution. Figure 3.7 shows the outbreak size distribution for our short-term commuting model. The average number of commuters ranges from 1 to 100. No significant effect on the outbreak size distribution can be seen, even for $\langle c \rangle = 1$. Figure 3.7 validates our assumption of the independence of infectious hosts, necessary for a branching-process formulation, as the simulations closely match the analytical predictions. It is noteworthy here that only the biology of the novel pathogen determines the emergence process, as outbreak sizes group according to the intermediate basic reproductive number α_2. The minimal deviation for $\langle c \rangle = 100$ commuters is based on the fact that the effective village size is only $N^{(v)} = 900$ due to the absent commuters.

In Figure 3.8, we extract the probability of emergence per introduction given a certain number already infected. These data are easily calculated using the outbreak size distribution and the probability of emergence per introduction. Assume an introduction has caused x infectious hosts already. The probability of extinction is the cumulative probability of getting an outbreak size equal to or larger than x, renormalized by all possible outcomes (extinctions and emergences) once x hosts are infectious. This yields the probability of emergence per number infected.

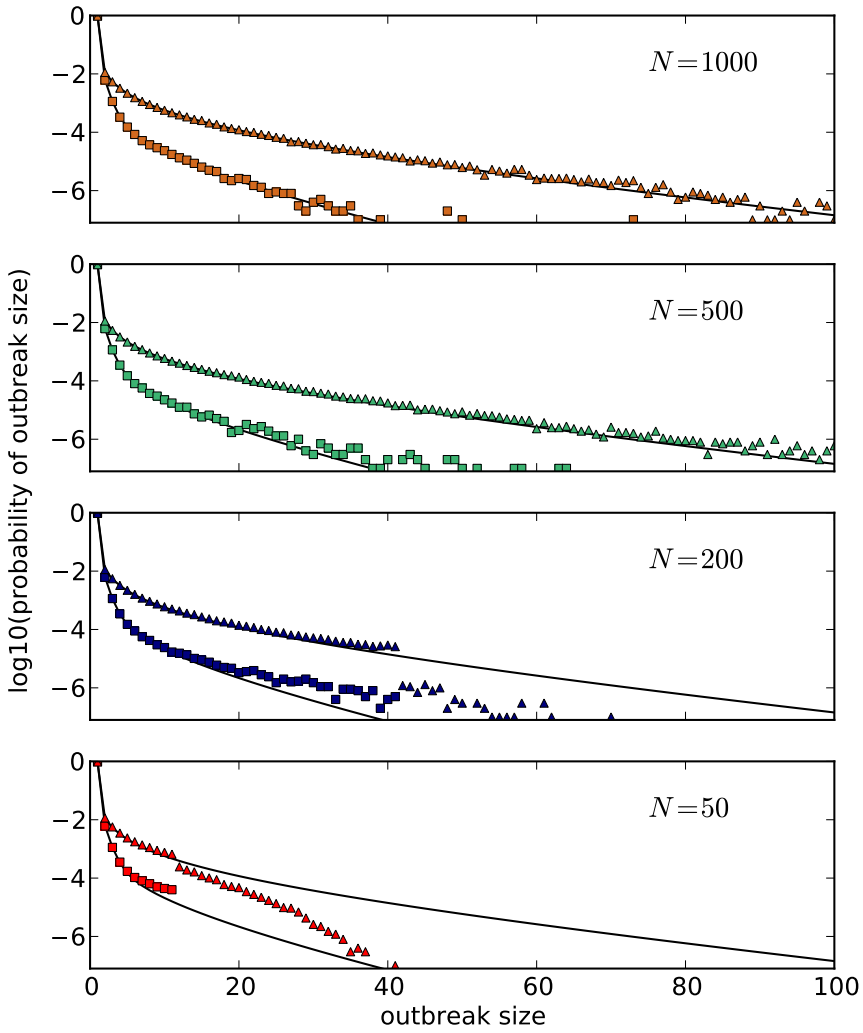

Figure 3.6: Outbreak size distribution for one single population. This figure shows the outbreak size distribution for outbreaks within one single population, i.e. the village within the meta-population model and zero connectivity. The solid black line represents the analytical result for an infinite population. Data points are the average probability of 10^7 simulations. Squares represent the punctuated route (I), and triangles the gradual route (II), both with the mutation probability $\mu = 0.1$. Outbreak size distribution as a function of host population size. For a population size, N, of 50, 200 and 500, an emergence event is defined as having at least 10% of the population size infected with the fully adapted strain, while the number was fixed at 100 infected with the fully adapted strain for $N = 1000$. While host populations with $N < 500$ clearly show finite size effects, even small host populations with $N \approx 500$ can be effectively treated as infinite as the outbreak size is small compared to the population size. The figure reveals the effect of the pathogen's evolution on the outbreak size distribution as the distributions group according to the route of adaptation.

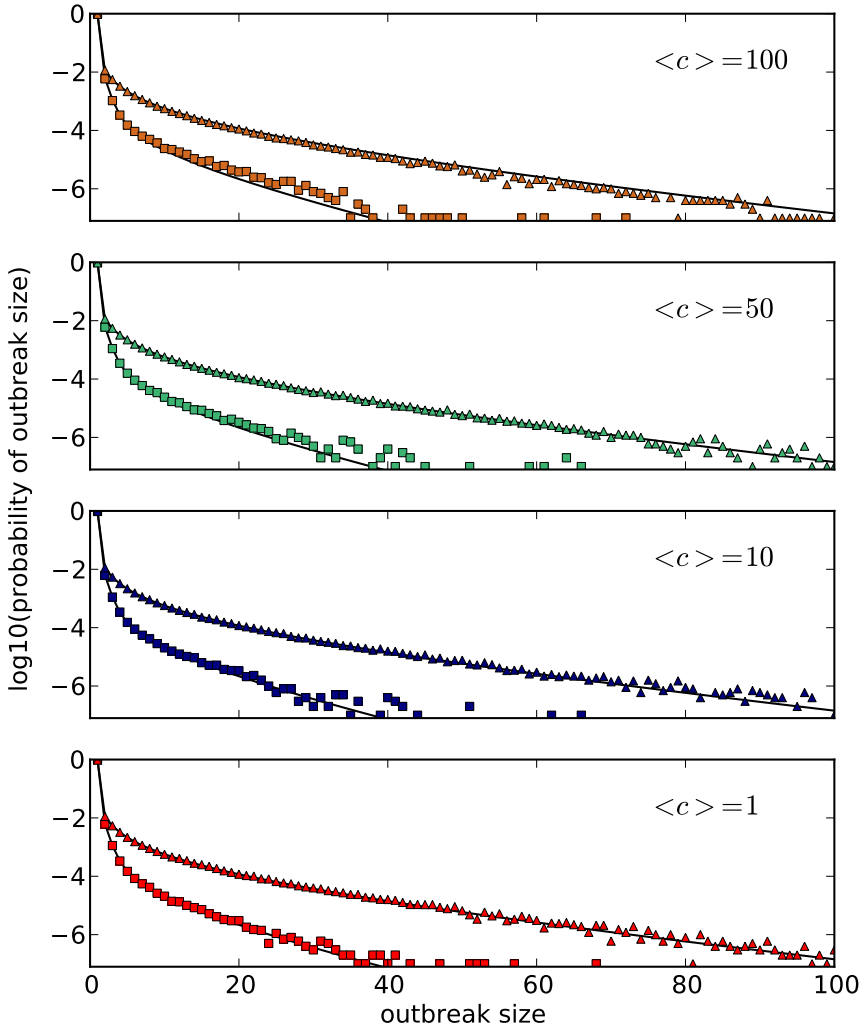

Figure 3.7: Outbreak size distribution of the village-city model with short-term commuting. The solid black line represents the analytical result for an infinite population. Data points are the average probability of 10^7 simulations. Squares represent the punctuated route (I), and triangles the gradual route (II), both with the mutation probability $\mu = 0.1$ and the basic reproductive numbers of Table 3.1. The probability is a function of the overall number of infectious hosts in the village-city model with $N = 1000$. The red color represents $\langle c \rangle = 1$, blue $\langle c \rangle = 10$, green $\langle c \rangle = 50$ and gold $\langle c \rangle = 100$ commuters. As for homogeneous populations, the outbreak size distributions group according to the evolutionary route of adaptation. Spatial heterogeneity does not have an influence as all simulations show no significant deviation from the analytical results which assume a single homogeneously mixed population.

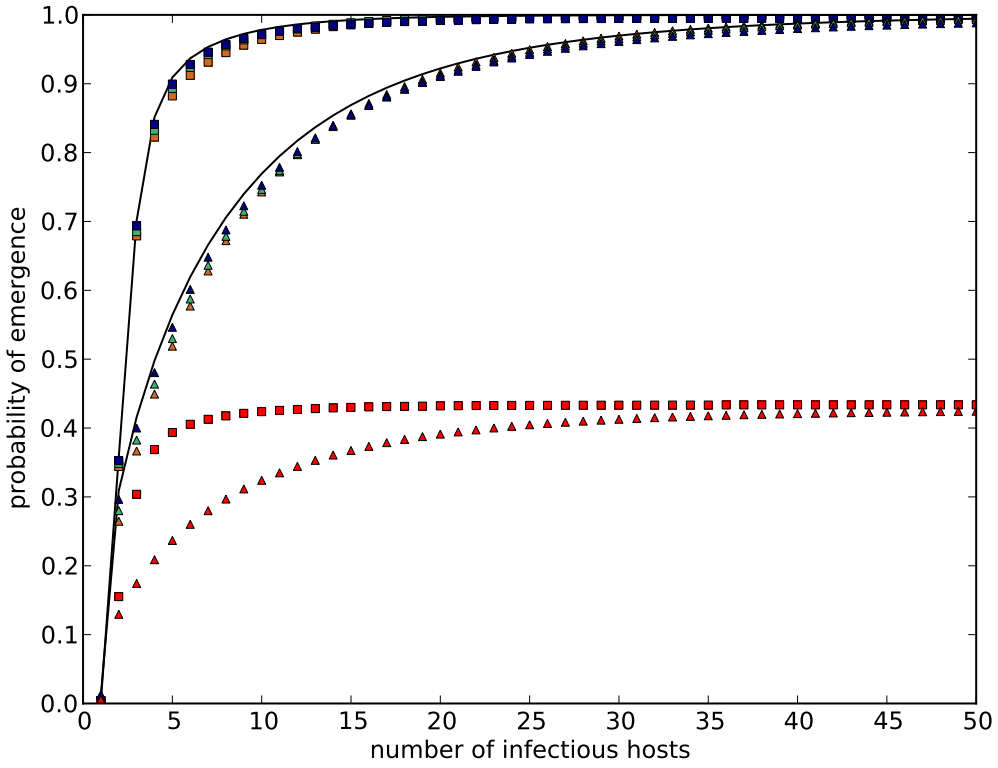

Figure 3.8: Probability of emergence in the city with short-term commuting. The probability is a function of the overall number of infectious hosts in the village-city model. The solid black lines represent the analytical results for an infinite population without spatial heterogeneous structure. Data points are the average probability of 10^7 simulations. Squares represent the punctuated route (I), and triangles the gradual route (II), both with the mutation probability $\mu = 0.1$ and the basic reproductive numbers of Table 3.1. The colour code of the data follows the same scheme as in Figure 3.7. As with the outbreak size distribution, probabilities group according to the route of adaptation instead of connection strength in number of commuters. However, the probability of emergence does not converge to 1 for $\langle c \rangle = 1$ commuters, revealing the effect of spatial heterogeneity when the number of commuters is small relative to the basic reproductive number for the well-adapted strain.

In Figure 3.8, the effect of spatial heterogeneity can be seen directly. For $\langle c \rangle \geq 10$, the village-city simulations agree very well with the analytical results assuming a single, infinite population. But for $\langle c \rangle = 1$, the probability of emergence converges to approximately 0.42. This effect reoccurs if we look at the probability of emergence as a function of total infectious hosts numbers in the village-city system, Figure 3.9. For $\langle c \rangle = 1$, the probability of emergence does not converge to 1 because of insufficient connectivity.

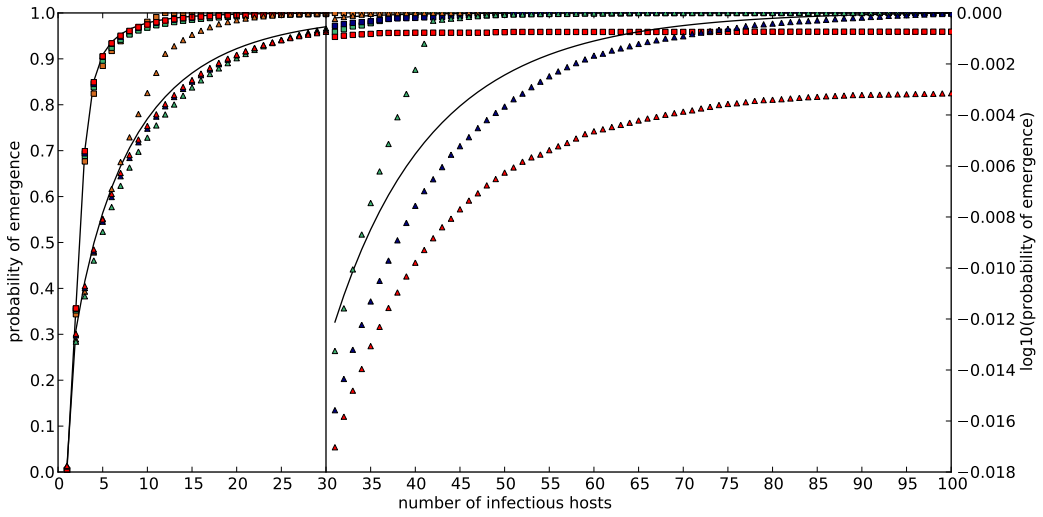

Figure 3.9: Probability of emergence for one single population. This Figure illustrates the probability of emergence as a function of the overall number of infectious hosts within one single population, i.e. the village within the meta-population model and zero connectivity. The solid black lines represent the analytical results for an infinite population without spatial heterogeneous structure. Data points are the average probability of 10^7 simulations. Squares represent the punctuated route (I), and triangles the gradual route (II), both with the mutation probability $\mu = 0.1$ and the basic reproductive numbers of Table 3.1. The colour code of the data follows the same scheme as in Figure 3.6. As with the outbreak size distribution in Figure 3.6, probabilities group according to the route of adaptation.

While Figure 3.8 shows the probability of emergence as a function of the number infected, the actual outcome is highly unpredictable even if the probability of an event is known as the average waiting time to an emergence shows (see equations (3.9) and (3.10), and Figure 3.2). It can be generalized for the probability of emergence given x infectious hosts. For example, the probability of emergence given five infectious hosts in the gradual route (II) is $p_{\text{em}}(x = 5) = 0.567$. It follows on average every 1.736 times this happens an

emergence will happen. The standard deviation is ± 1.161, which leads to the conclusion that even if the probability is known, it is inherently unpredictable when this will actually lead to an emergence.

Figure 3.10 shows the outbreak size distribution of our long-term migration model, compared with the analytical results for one infinite population. The number of commuters varies between 10 and 100. Even for the smallest average number of commuters the simulations agree very well with the analytical derivations, a result consistent with that for the short term commuting model discussed previously.

Figure 3.11 shows the probability of emergence as a function of the infectious host number. Again, the analytical derivations show very good agreement with the simulations, neglecting an influence of spatial heterogeneity, and revealing a grouping according to the biological adaptation process of the novel pathogen. Note that Figure 3.11 also illustrates the possibility of an emergence in the village without further sustained transmission in the city for $\langle c \rangle = 10$ as the probability of emergence does not reach 1. Nevertheless, this is a very unlikely scenario with a probability of less than 0.02 at 100 infected hosts.

Interestingly, five infectious hosts or more are needed in the gradual route (II) before emergence becomes more likely than extinction (see Figure 3.11). In the punctuated route (I), this becomes three or more infected, caused by the very low probability of generating secondary infected with a basic reproductive number of $\alpha_2 = 0.1$. Secondary infections are most likely to occur with the fully adapted strain only.

3.6.4 Spatial Homogeneity Coefficient

This confirms that a pathogen needs a sufficient connection between communities to emerge, despite its ability to always cause outbreaks of limited size. Hence we expect the existence of a threshold where spatial heterogeneity effectively does not matter any more. Previous research has shown that the effect of heterogeneity in spatially structured population models depends on the interconnectivity with a threshold effectively allowing the pathogen to spread between communities (Watts et al., 2005; Vazquez, 2007; Colizza and Vespignani, 2008). Our approach allows new insights, as we do not need to specify the actual num-

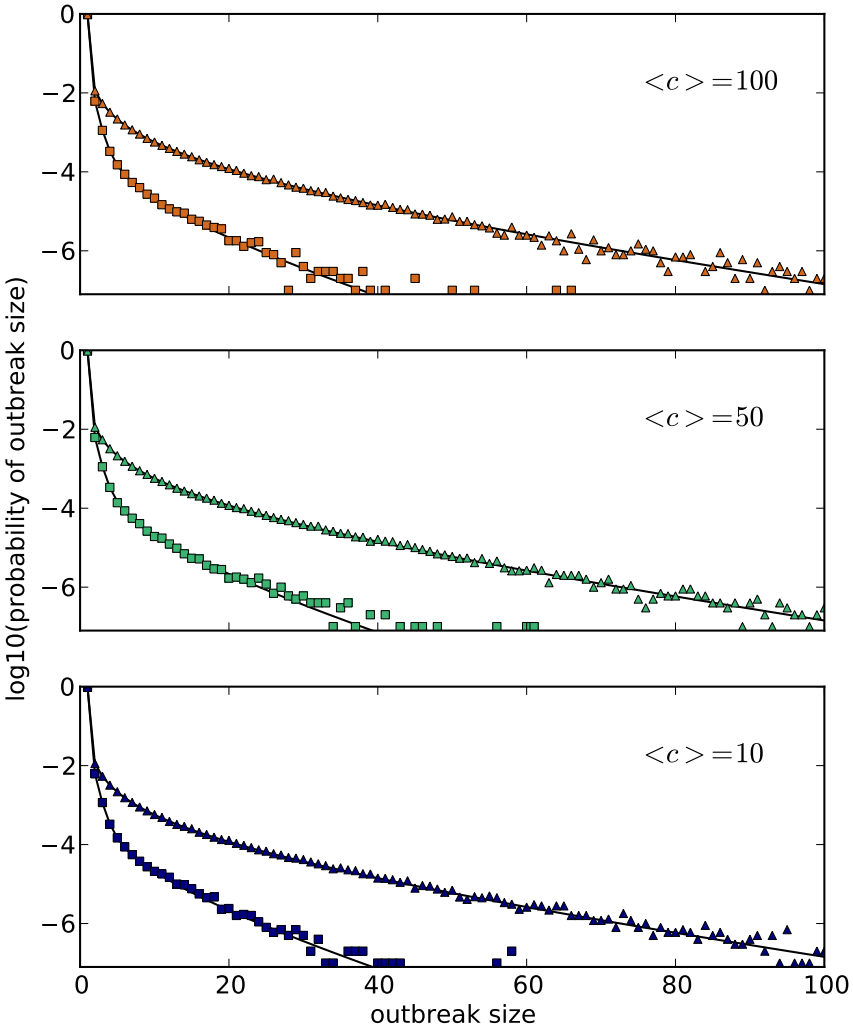

Figure 3.10: Outbreak size distribution of the village-city model with long-term commuting. The solid black lines represent the analytical results for an infinite population without spatial heterogeneous structure. Data points are the average probability of 10^7 simulations. Squares represent the punctuated route (I), and triangles the gradual route (II), both with the mutation probability $\mu = 0.1$ and the basic reproductive numbers of Table 3.1. As for homogeneous populations, the outbreak size distributions group according to the evolutionary route of adaptation. The average number of commuters, $\langle c \rangle$, does not have an influence as all simulations do not show a significant variation from the analytical results.

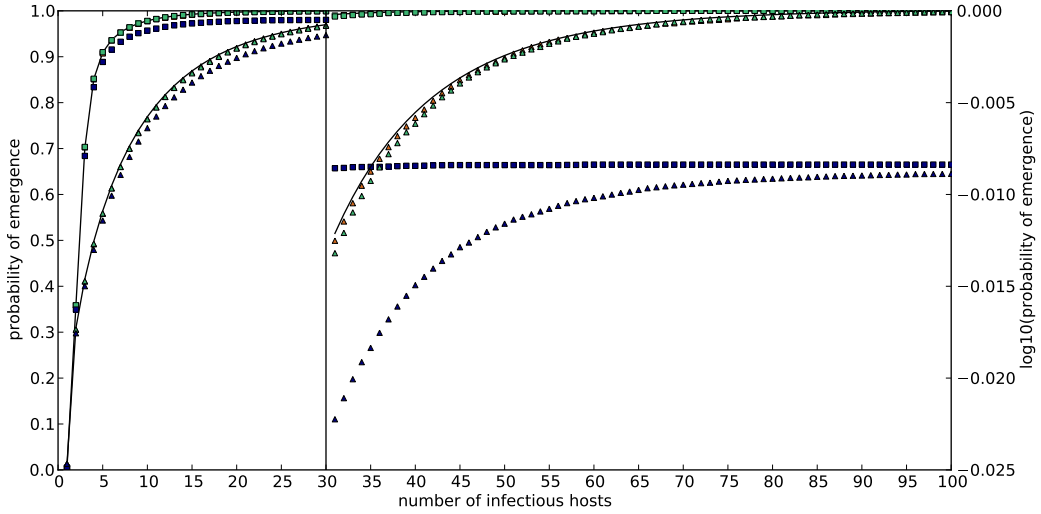

Figure 3.11: Probability of emergence in the city with long-term commuting. The probability is a function of the overall number of infectious hosts in the village-city model. The solid black lines represent the analytical results for an infinite population without spatial heterogeneous structure. Data points are the average probability of 10^7 simulations. Squares represent the punctuated route (I), and triangles the gradual route (II), both with the mutation probability $\mu = 0.1$ and the basic reproductive numbers of Table 3.1. The orange colour represent an average number of commuters $\langle c \rangle = 10$, the blue colour $\langle c \rangle = 50$, and the red colour $\langle c \rangle = 100$. The plot is dived at 30 infectious hosts into a linear probability scale (left), and a log_{10} probability scale (right). As with the outbreak size distribution, probabilities group according to the route of adaptation instead of connection strength in number of commuters. However, the probability of emergence does not converge to 1 for $\langle c \rangle = 10$ commuters. Here the effect of spatial heterogeneity is very small but visible as a small number of commuters can lead to an emergence in the village without further, sustained transmission to and within the city.

ber of infectious hosts migrating to a new community. We measure connectivity between communities in terms of the average number of commuters $\langle c \rangle$ for which rich empirical datasets can be found. Figure 3.12 presents illustrative examples of empirical data of commuting patterns in different parts of the world. Most data has been collected by Offices of Statistics of five countries on three continents (Instituto Brasileiro de Geografia e Estatistica, 2000; US Census Bureau, 2000; Office for National Statistics, 2001; Census and Statistics Department, 2001; Statistics Bureau Japan, 2005). A further, two independent studies have been used to estimate commuting patterns of towns in Indonesia (Leinbach, 1983) and China (Xu, 2001).

We next attempt to quantify the regimes in $\langle c \rangle$ for which spatial heterogeneity may be neglected. We approach this question using a simple analytical derivation for the effect of spatial heterogeneity, which considers only the fully adapted strain. Assume a connected community such as the village in our village-city model, with a fully adapted pathogen introduced into the village population. Given an emergence and epidemic in the village, the probability that this causes an emergence and epidemic in the city is

$$H_{\text{spatial}} = 1 - \left(\frac{1}{R_0}\right)^{\langle c \rangle f} \tag{3.27}$$

Therefore, H_{spatial} is a spatial homogeneity coefficient, measuring the impact of spatial heterogeneity on an emergence process. It ranges from 0, leading to two isolated communities, to 1, effectively removing any spatial heterogeneity and forming one homogeneous population. f is the fraction of the village residents eventually becoming infectious. It can be derived using (Anderson and May, 1992)

$$f = 1 - e^{-R_0 f} \tag{3.28}$$

The spatial homogeneity coefficient depends only on the connection strength expressed in commuters $\langle c \rangle$ and the basic reproductive number $R_0 = \alpha_3$ of the fully adapted strain. Though it only considers the fully adapted strain, we expect this coefficient to be a good approximation for a multi-strain model as the vast majority of transmissions will be from

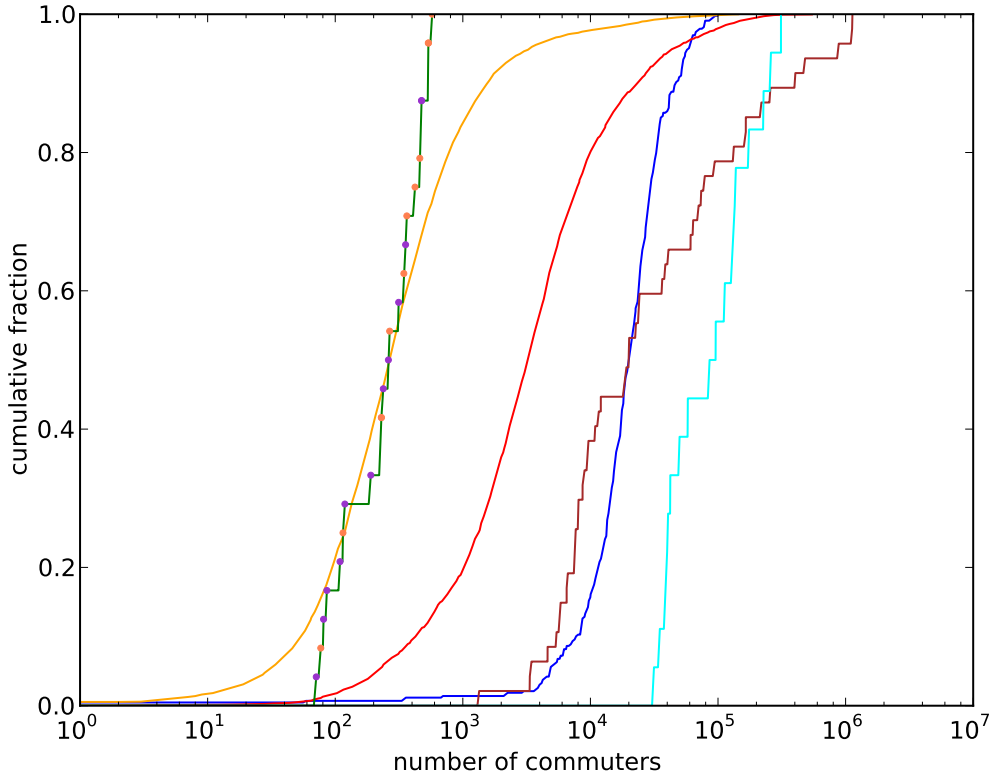

Figure 3.12: Data of commuting patterns in different parts of the world. Shown is the cumulative fraction of all communities with equal or less than the specified number of commuters. The data was mostly collected by the National Statistical Offices of the respective countries. The gold line represents commuting data from Brazil (Instituto Brasileiro de Geografia e Estatistica, 2000), the red line data from the USA (US Census Bureau, 2000), the blue line data from the UK (Office for National Statistics, 2001), the brown line data from Japan (Statistics Bureau Japan, 2005), the cyan line data from Hong Kong (Census and Statistics Department, 2001), and the green line data from two independent sources. The green line has orange (Xu, 2001) and pink data points (Leinbach, 1983), corresponding to its data sources. Our data represents the commuting flows between administrative units. The definition of administrative units varies highly between countries. For example, the US data is on a granularity of 3,141 counties, while the data from Japan is based on its 47 prefectures. However, heterogeneity can also be found within countries datasets. The Brazilian data is on a level of 5,507 municipalities with resident sizes ranging from 1,166 to 10,435,548.

those infected with the fully adapted strain in the case of an emergence.

Figure 3.13 gives an overview of the influence of spatial heterogeneity as a function of $\langle c \rangle$. Effectively, spatial heterogeneity is negligible once a critical number of ten commuters connect the two communities. This is a very low threshold, and empirical data shows that the probability of having a community with less than the critical number of ten commuters is approximately 1% for all the reviewed data combined.

As illustrated by the close fit between analytical and numerical results in figure 3.13, there is only a small error in the analytical expression arising from neglecting infections with mal-adapted strains. This error is greatest for the gradually adapting pathogen, because an intermediate strain with $\alpha_2 = 0.9$ tends to cause larger outbreaks than one with $\alpha_2 = 0.1$. Nevertheless, the deviation remains small.

In light of this agreement, how does the critical average number of commuters vary with the adapted reproductive number? While for $R_0 = 2$ ten average commuters are sufficient to dissolve spatial heterogeneity, this changes dramatically for smaller basic reproductive numbers (see Figures 3.14 and 3.15). If a well-adapted strain is only just pandemic capable (i.e. R_0 just above 1) villages with only ten commuters are only 50% likely to seed an epidemic in their local city, and spatial structure becomes important again. For example, the critical number of commuters is close to 100 for $R_0 = 1.2$. For $R_0 = 2$ the spatial homogeneity coefficient is approximately 0.42 for one commuter. This agrees with what we find in Figure 3.8 using simulated results.

3.7 Discussion

In this chapter, we first present analytical results to calculate epidemiological parameters of a novel disease, adapting to humans. We explore the influence of spatial host contact structure, and validate our result with stochastic simulations of simple village-city models as an example of interconnected communities within a spatially structured population. Our research reveals that for plausible parameter ranges, spatial heterogeneity only has very limited impact on the probability of emergence, as well as the outbreak size distribution. Neither a change in strength of spatial heterogeneity (e.g. number of commuters),

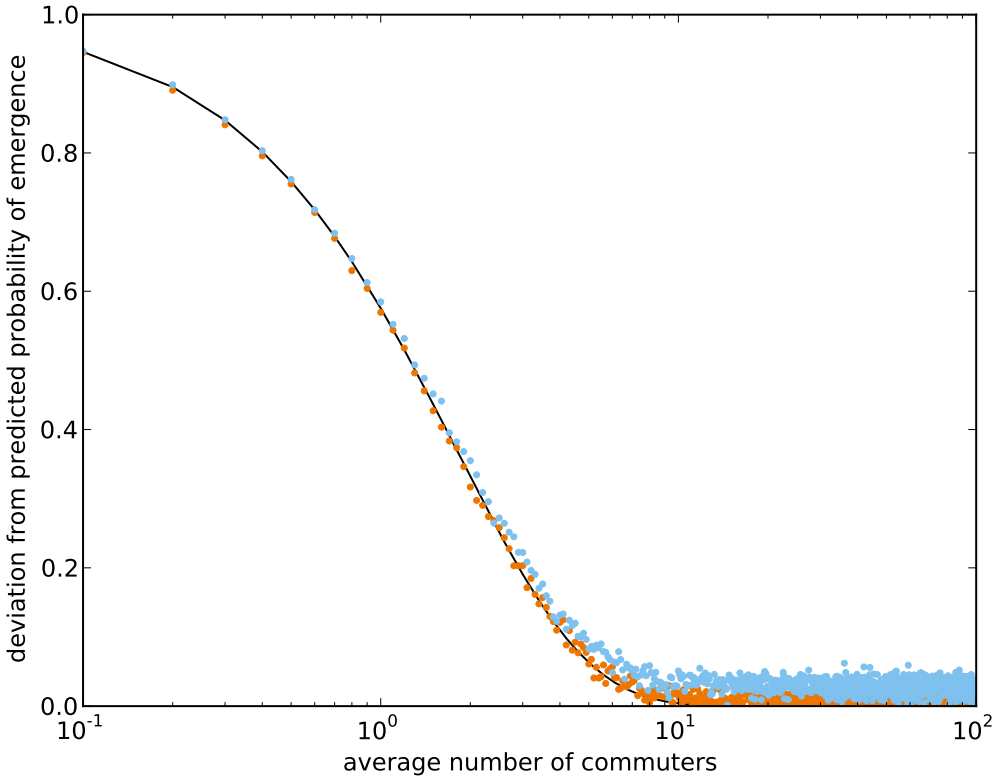

Figure 3.13: Deviation between simulated and analytical predicted probability of emergence. The deviation is a function of the average number of commuters $\langle c \rangle$. The deviation is defined as $|p_{\text{em}}^{(\text{theo})} - p_{\text{em}}^{(\text{ex})}| / p_{\text{em}}^{(\text{theo})}$. The analytical probability of emergence $p_{\text{em}}^{(\text{theo})}$ is for an infinite population without spatial structure. The simulated probability of emergence $p_{\text{em}}^{(\text{ex})}$ is for short-term commuting with the blue data points representing the gradual route (II) of adaptation, and the orange data points representing the punctuated route (I). The solid black line is $1 - H_{\text{spatial}}$, with the spatial homogeneity coefficient defined in equation (3.27). It is the analytical expected deviation for spatial heterogeneity as a function of the spatial homogeneity coefficient. The simulations agree very well with the analytical expected deviation. The gradual route (II) is slightly more off from the theoretical prediction as a result of the small but significant number of infected with the intermediate strain. Effectively, it lowers the number of commuters infected with the fully adapted strain and therefore the probability of transmission from the village to the city. Nevertheless, the analytical prediction as well as the simulations show no significant impact of spatial heterogeneity once the number of commuters exceeds $\langle c \rangle = 10$.

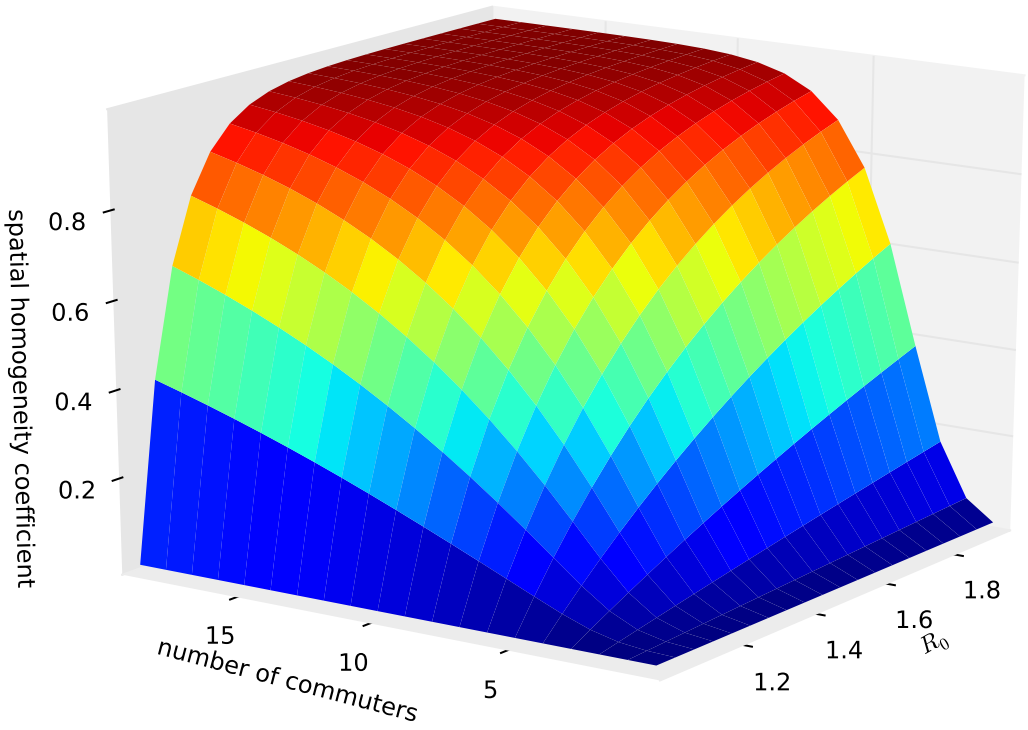

Figure 3.14: Impact of spatial heterogeneity on disease transmission between communities (I). The impact is measured with the spatial homogeneity coefficient with $0 \leq H_{\text{spatial}} \leq 1$. Given $H_{\text{spatial}} = 1$ every emergence in the village automatically leads to an emergence in the city, and $H_{\text{spatial}} = 0$ represents no chance of successfully transmitting the pathogen into the city. The figure reveals that spatial structure becomes especially important for small basic reproductive numbers. In addition, the average number of commuters needed to show an effect of spatial heterogeneity is surprisingly small.

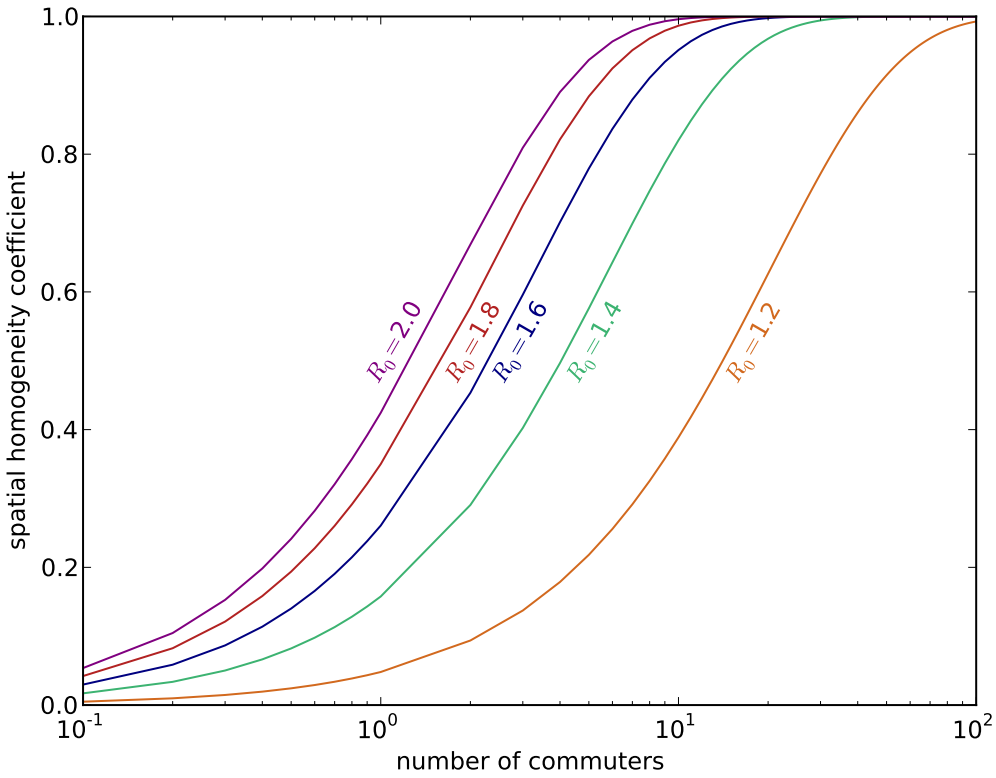

Figure 3.15: Impact of spatial heterogeneity on disease transmission between communities (II). This Figure illustrates the spatial homogeneity coefficient as presented in equation (3.27) for five different basic reproductive numbers. Therefore, it presents five 'slices' through Figure 3.14.

nor in its quality (e.g. short term versus long term commuting) shows a significant influence. Our results suggest that only the most remote rural communities would be subject to epidemiological isolation. In particular, the available empirical data suggests that communities tend to be highly interconnected with relatively high connection strengths. Of course, it is the most remote communities of the world for whom we have the least relevant data. More empirical research on spatial heterogeneity is needed to form a better understanding of its effect, and this need is greatest in developing parts of the world.

In addition, population size becomes an important factor only when that population is relatively small − fewer than $N \approx 500$ individuals. Only a small number of infectious hosts are actually involved in the emergence process, which relates to the small reservoir of susceptibles needed for a successful emergence. Moreover, biological processes such as the speed of evolution and the adaptive route show a strong influence on the overall emergence process. We show that epidemiological parameters such as the outbreak size group according to the evolutionary route. Previous research has shown the effect of the pathogen's route of evolutionary adaptation and mutation rate on the probability of emergence per introduction (Antia et al., 2003; Arinaminpathy and McLean, 2009). Our theoretical derivation of the probability of emergence extends this and offers the benefit of being analytically solvable for any possible route of adaptation and any mutation probabilities. We note here that previous work has highlighted the role of other, significant types of heterogeneity in the emergence of a novel infection. For example, André and Day (2005) describe the effect of the pathogen's life history, such as the length of infection, on the emergence of a novel pathogen. Further, heterogeneity in human-to-human transmission within a population may have an influence on the course and probability of emergence and outbreaks (Lloyd-Smith et al., 2005; Yates et al., 2006), usually lowering the probability of emergence. While we have concentrated here on simple types of spatial heterogeneity, a significant question for future research is the role of mixed heterogeneities, for example spatially structured populations with additional heterogeneity in the human-to-human transmission.

We also find that the waiting time for an outcome, that is emergence or extinction, of a novel pathogen's introduction is highly unpredictable, even if the probability for such an

event is known. Conversely, this means that an estimate of the underlying epidemiological parameters from observed data will be highly uncertain. Unfortunately, a large number of observations will be necessary to achieve confidence in the parameters, and even a large number of introductions gone extinct do not rule out the possibility of emergence for a pathogen. Arinaminpathy and McLean (2009) come to a similar result using a measurement on the upper bound of the probability of emergence (see Figure 2.3). Moreover, the probability of emergence given a certain number of infectious hosts can be surprisingly low. Even a comparatively large number of infectious hosts can end in extinction, especially for low mutation rates and intermediate-stage basic reproductive numbers just below one.

Our work has relevance for important public health issues: if a novel disease is detected in a rural setting, and it appears to be spreading, how feasible is it to contain infection by restricting movements to and from the village? Our results suggest that first, an infeasibly tight level of quarantine would be required for any chance of containment, corresponding to enforcing a low level of $\langle c \rangle$ in Figure 3.14. To all intents and purposes isolation would have to be absolute to be effective. In most circumstances such extreme intervention would not be acceptable. Second, given typical mobility patterns, it is likely that once there is a detectable number of cases in the village, there may already be a significant number of cases outside of it. Therefore quarantining interventions are likely to come too late.

Our work raises important questions for future research: where should surveillance be focused to detect an emergence as early as possible, especially if resources are limited? Given emergence of a novel infection in a rural setting, how much time can we buy through limiting travel to and from major urban centres? These and other questions will undoubtedly benefit from more systematic studies of emergence in the context of population distributions. Nonetheless, theoretical models such as those presented here can offer useful, fundamental insights to guide such studies.

CHAPTER 4

THE INFLUENCE OF THE ENVIRONMENT ON PATHOGEN SPECIALISATION

4.1 Introduction

In the literature review, I introduced the environment as the 'totality of everything excluding the pathogen' (see section 1.3). Undoubtedly, there is extensive evidence in the literature that the environment plays a substantial role in infectious disease transmission (Holt et al., 2003; Coleman and Welburn, 2004; Cauchemez et al., 2009). Here, I focus on a specific part of these environmental factors: the interaction between host species and the ensuing impact upon their specialist and generalist pathogens. A pathogen might be specialised to its donor species, get transmitted into a recipient species, but then need evolutionary adaptations to emerge into the novel species (see chapter 3 for an example of such a process). On the other hand, an infectious agent might cross into the novel, recipient species and already be well suited for sustained transmission within this species. In the first case, pathogens are bound to their respective host species (ignoring possible evolutionary adaptations). These pathogens are called 'specialists' in ecology (Woolhouse et al., 2001). In the other case, the pathogen is a 'generalist' and able to propagate in both species. Obviously, it is of great interest in the context of zoonotic EIDs if an infectious agent is a specialist or generalist, and hence to what extent evolutionary adaptations are necessary for emergence.

This leads to the following naïve question: why are some pathogens ecological specialists only able to infect a single hosts species? It has been suggested that specialists tend to have a higher fitness than generalists within their target host species, presumably because generalists cannot adapt to a single host species as well as specialists without loosing their ability to infect multiple species (Futuyma and Moreno, 1988; Van Tienderen, 1991; Woolhouse et al., 2001; McMichael, 2001; Coleman and Welburn, 2004; Wiklund and Friberg, 2009; Benmayor et al., 2009). Hence, a generalist might only out-compete a specialist pathogen if its fitness disadvantage within one single host species is balanced through spill-over infections from other host species.

In general, three scenarios are possible in the order of an increasing connectivity between the species:

I. specialist pathogens are the dominant pathogens in all host species,

II. a specialist and a generalist pathogen are each the dominant pathogen in at least one host species,

III. or a generalist pathogen is the dominant pathogen in all host species.

Here, I will focus on the establishment and dominance of generalists or specialists based on the connectivity between their host species. My research questions are: what is the basic reproductive number of a generalist pathogen infecting multiple species? Which connectivity between host species facilitates which of the three scenarios listed above? What happens with the non-predominant pathogen in each scenario? What is the influence of different population sizes? This research is also motivated by the prominent notion that 'change' seems to facilitate emergence events (Cohen, 2000). A change in the connectivity between host species might change the dominance of one pathogen in an animal species with possible implications for the likelihood of future emergence events into humans. In the following, I will use a simple theoretical model of two species with variable inter-connectedness to address these research questions.

4.2 Theoretical Analysis

4.2.1 Defining a Model System

Pathogen fitness is measured with the basic reproductive number as it directly describes the transmission process - the bottleneck of reproduction (see section 2.2). Here, the defining characteristic for a specialist is its higher fitness within a single host species, while a generalist can infect multiple host species but has a lower fitness in each one of them. This is a common assumption in the literature (Futuyma and Moreno, 1988; Woolhouse et al., 2001), and results of experimental studies seem to agree with it (Benmayor et al., 2009). I define the specialist as a 'pure' specialist without any between-species transmission capabilities. Only the generalist pathogen can successfully transmit between different host species.

The theoretical model consist of two separate host species, A and B. The reproductive numbers of an infectious specialist and generalist of species A within their own species

are defined as $\alpha_{AA}^{(s)}$ and $\alpha_{AA}^{(g)}$ respectively. The number of secondary infections in the other species are expressed with $\alpha_{AB}^{(s)}$ and $\alpha_{AB}^{(g)}$. In an analogous manner, the reproductive numbers of an infectious specialist and generalist of species B within their own species are defined as $\alpha_{BB}^{(s)}$ and $\alpha_{BB}^{(g)}$ respectively, and the number of secondary infections in the other species are expressed with $\alpha_{BA}^{(s)}$ and $\alpha_{BA}^{(g)}$. However, these reproductive numbers are limited in that they only express the secondary infections in a single species, while the basic reproductive number is the overall number of secondary infections in both species.

Using the above definition of generalist and specialist strains, a formalised set of reproductive numbers reads

$$\begin{aligned} \alpha_{AA}^{(s)} > \alpha_{AA}^{(g)} > 0 \quad \text{and} \quad \alpha_{AA}^{(s)} > \alpha_{BA}^{(g)} > 0 \quad \text{and} \quad \alpha_{AB}^{(s)} = 0 \\ \alpha_{BB}^{(s)} > \alpha_{BB}^{(g)} > 0 \quad \text{and} \quad \alpha_{BB}^{(s)} > \alpha_{AB}^{(g)} > 0 \quad \text{and} \quad \alpha_{BA}^{(s)} = 0 \end{aligned} \tag{4.1}$$

That is to say: each specialist outcompetes any generalist in its own host species, each specialist cannot infect the other host species, and the generalist can infect either host species. I assume that the infectious period is the same for both pathogen types, because my aim is to compare two similar pathogens which are only different in their transmission potential within and between two different host species.

The transmission model, as schematically shown in Figure 4.1, is formulated as a simple frequency-dependent SIS model (see section 2.3). It contains three types of pathogens. One specialised on host species A, one specialised on host species B - which both have the same epidemiological parameters -, and a generalist which can jump between the host species and infect both. Normalising time in respect to the infectious period, my model

reads

$$\begin{aligned}
\frac{dS_A}{dt} &= -\left(\lambda_A^{(s)} + \lambda_A^{(g)}\right)\frac{S_A}{N_A} + I_A^{(s)} + I_A^{(g)} \\
\frac{dI_A^{(s)}}{dt} &= \lambda_A^{(s)}\frac{S_A}{N_A} - I_A^{(s)} \\
\frac{dI_A^{(g)}}{dt} &= \lambda_A^{(g)}\frac{S_A}{N_A} - I_A^{(g)} \\
&\text{and} \\
\frac{dS_B}{dt} &= -\left(\lambda_B^{(s)} + \lambda_B^{(g)}\right)\frac{S_B}{N_B} + I_B^{(s)} + I_B^{(g)} \\
\frac{dI_B^{(s)}}{dt} &= \lambda_B^{(s)}\frac{S_B}{N_B} - I_B^{(s)} \\
\frac{dI_B^{(g)}}{dt} &= \lambda_B^{(g)}\frac{S_B}{N_B} - I_B^{(g)}
\end{aligned} \tag{4.2}$$

with the 'force of infection' as the effective transmission according to definition (4.1) (Lloyd and May, 1996)

$$\begin{aligned}
\lambda_A^{(s)} &= \alpha_{AA}^{(s)} I_A^{(s)} \quad \text{and} \quad \lambda_A^{(g)} = \alpha_{AA}^{(g)} I_A^{(g)} + c\,\alpha_{AB}^{(g)} I_B^{(g)} \\
\lambda_B^{(s)} &= \alpha_{BB}^{(s)} I_B^{(s)} \quad \text{and} \quad \lambda_B^{(g)} = \alpha_{BB}^{(g)} I_B^{(g)} + c\,\alpha_{BA}^{(g)} I_A^{(g)}
\end{aligned} \tag{4.3}$$

The connectivity c defines the degree of inter-species mixing, and therefore the likelihood an infected host of species B meets a host of species A and vice versa.

4.2.2 The Basic Reproductive Number

The basic reproductive number is a crucial piece of information in epidemiology (see section 2.2). It is the secondary number of infections in all host species combined. Because specialists can only transmit within one species, their basic reproductive numbers reduce to the standard case with just one species. This is also expressed in the force of infection (4.3) with the reproductive numbers $\alpha_{AA}^{(s)}$ and $\alpha_{BB}^{(s)}$ being the basic reproductive numbers.

The basic reproductive number of the generalist pathogen has a more complex expression, because it includes infections in both species. It therefore depends on the reproductive number within each single species and the connectivity between the species (see (4.3)).

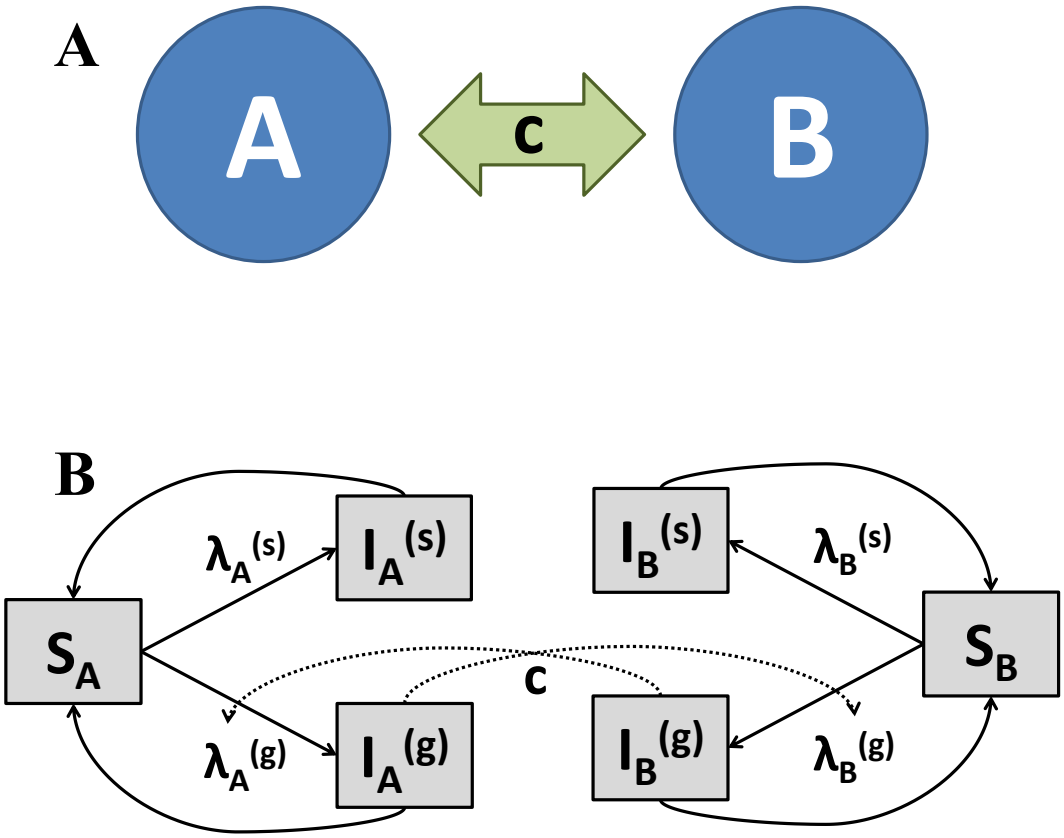

Figure 4.1: Schematic representation of the generalist-specialist model (see equation (4.2)). **A** shows the interaction between the two species with the connectivity parameter c. **B** illustrates the transmissions between the susceptible and infectious states in both species.

It is possible to derive the total number of transmission events with the next-generation-matrix $\mathcal{G}$ (Diekmann et al., 1990, 2010)

$$\mathcal{G}^{(g)} = \begin{pmatrix} \alpha_{AA}^{(g)} & c\,\alpha_{AB}^{(g)} \\ c\,\alpha_{BA}^{(g)} & \alpha_{BB}^{(g)} \end{pmatrix} \tag{4.4}$$

The basic reproductive number $\alpha^{(g)}$ for the generalist pathogen is the dominant eigenvalue of the next-generation-matrix. For a $2x2$ matrix, this eigenvalue can be obtained using the trace and the determinant of the matrix (Diekmann et al., 2010)

$$\begin{aligned} \alpha^{(g)} &= \frac{1}{2}\left(\text{trace}\left(\mathcal{G}^{(g)}\right) + \sqrt{\text{trace}\left(\mathcal{G}^{(g)}\right)^2 - 4\det\left(\mathcal{G}^{(g)}\right)}\right) \\ &= \frac{\alpha_{AA}^{(g)} + \alpha_{BB}^{(g)}}{2} + \sqrt{\left(\frac{\alpha_{AA}^{(g)} + \alpha_{BB}^{(g)}}{2}\right)^2 - \alpha_{AA}^{(g)}\,\alpha_{BB}^{(g)} + c^2\,\alpha_{AB}^{(g)}\,\alpha_{BA}^{(g)}} \end{aligned} \tag{4.5}$$

Under the assumption that the generalist reproductive numbers are symmetric, i.e. $\alpha_{AA}^{(g)} = \alpha_{BB}^{(g)}$ and $\alpha_{AB}^{(g)} = \alpha_{BA}^{(g)}$, (4.5) reduces to

$$\alpha^{(g)} = \alpha_{AA}^{(g)} + c\,\alpha_{AB}^{(g)} = \alpha_{BB}^{(g)} + c\,\alpha_{BA}^{(g)} \tag{4.6}$$

In an analogous manner, the specialists' basic reproductive number reads in a symmetric setting

$$\alpha^{(s)} = \alpha_{AA}^{(s)} = \alpha_{BB}^{(s)} \tag{4.7}$$

Although these are very clear, in what follows I explore the general, asymmetric case.

4.2.3 Pathogen Predominance and Co-Existence

The analytical derivation of $\alpha^{(g)}$ reveals the importance of the connectivity c, which allows between-species transmission. The definition of the reproductive numbers in (4.1) already states that the specialist pathogen will always out-compete the generalist in single, isolated populations. Hence the generalist can only win over the specialist if its effective number of secondary infections is higher than that of the specialist. Therefore, the transmission events

induced by infectious hosts from the other species need to be above a certain threshold to generate enough momentum for the generalist pathogen to out-compete the specialist. The generalist's ability to access more resources (in terms of the larger number of susceptible hosts) can, if c is large enough, compensate for the generalist's lower fitness in each individual type of host (Futuyma and Moreno, 1988; Woolhouse et al., 2001).

This leads to the question: what is the threshold connectivity necessary for the generalist pathogen to become the dominant pathogen, that is for the system to go from scenario I to II, or from scenario II to III? There is one threshold parameter, $\hat{c}_1$, above which the generalist becomes the dominant pathogen within one single host species (transition from scenario I to scenario II), and one threshold parameter, $\hat{c}_2$, above which the generalist becomes the dominant pathogen within both host species (transition from scenario II to scenario III). Figure 4.2 illustrates the connection between thresholds and scenarios. By definition, when the entire system is at equilibrium and $c = \hat{c}_1$, the number of hosts infected with the specialist pathogen is equal to the number of hosts infected with the generalist pathogen within one species, while in the other species the number of hosts infected with the specialist pathogen is larger than the number infected with the generalist pathogen. Moreover, at $c = \hat{c}_2$, the number of hosts infected with the specialist pathogen is equal to the number infected with the generalist pathogen within one species, while in the other species the number of hosts infected with the generalist pathogen is larger than the number infected with the specialist pathogen.

The condition for the endemic equilibrium is that the flux of new infections equals the flux of recoveries. That is, equation (4.2) reads at the endemic equilibrium

$$\frac{dS_A}{dt} = \frac{dS_B}{dt} = \frac{dI_A^{(s)}}{dt} = \frac{dI_B^{(s)}}{dt} = \frac{dI_A^{(g)}}{dt} = \frac{dI_B^{(g)}}{dt} = 0 \tag{4.8}$$

Endemic states are constant in value and marked with an asterisk, i.e. $I_A^{(g,*)} = \text{const}$. The thresholds between the scenarios are $\hat{c}_1$ and $\hat{c}_2$. If the connectivity equals one of these threshold values, i.e. $c = \hat{c}_1$ or $c = \hat{c}_2$, the number of infected hosts with the specialist pathogen and the number of infected hosts with the generalist pathogen are of equal size

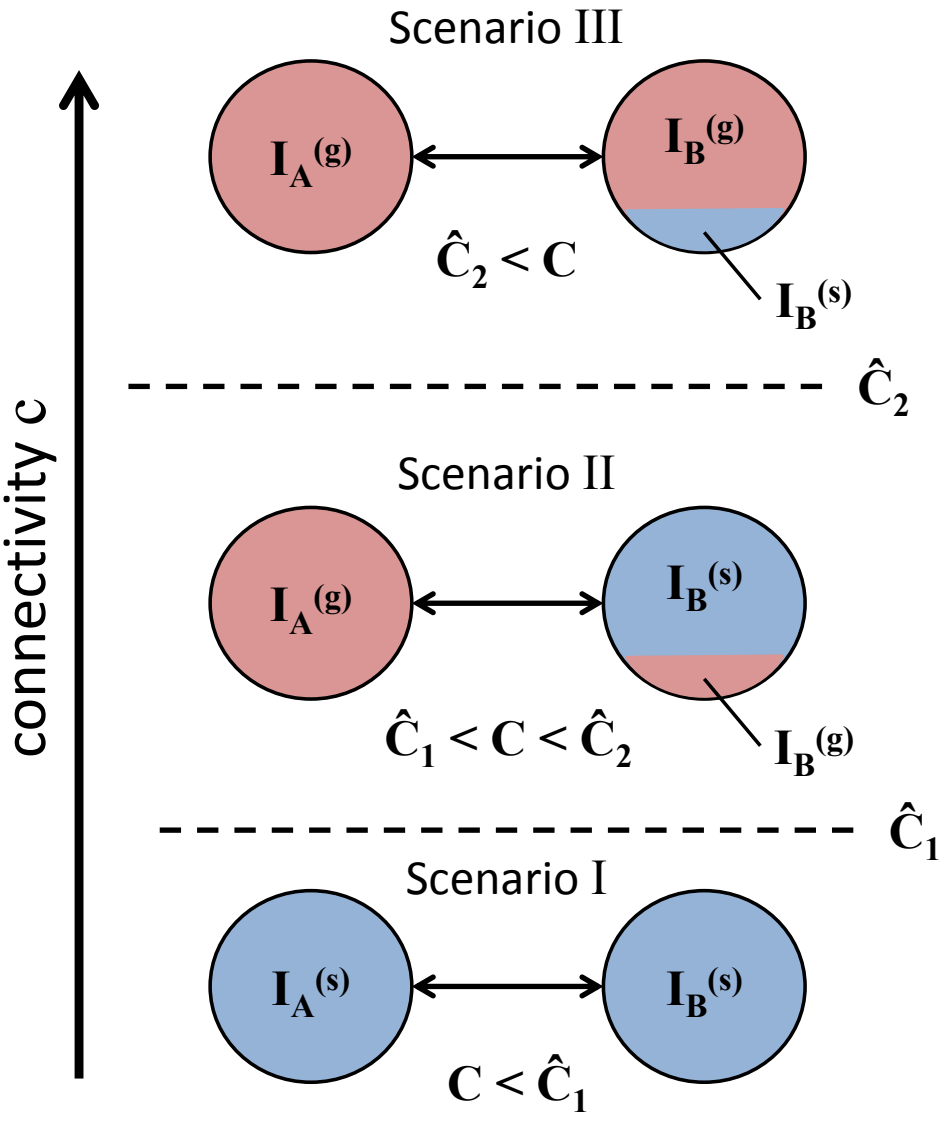

Figure 4.2: Schematic overview of the three different scenarios. The host species, A and B, are represented with a circle. The number of infected hosts with a specialist pathogen is marked in blue, and the number of infected hosts with the generalist pathogen is marked in red. The system starts in scenario I with $c < \hat{c}_1$. Only the generalist pathogen can transmit between the species. Once c increases above $\hat{c}_1$, the system reaches scenario II in which one species is dominated by the generalist pathogen with the specialist extinct (left circle). In the other species, the number infected with the specialist pathogen is still larger than the number infected with the generalist pathogen (right circle). Scenario III describes the case for $c > \hat{c}_2$. Here, the generalist pathogen is dominant in both species, and the specialist exists only within one species (right circle).

in one of the host species, which I arbitrarily chose to be A for demonstration purposes

$$\begin{aligned} I_A^{(*)} &= I_A^{(s,*)} = I_A^{(g,*)} \\ \lambda_A^{(s,*)} \frac{S_A}{N_A} &= \lambda_A^{(g,*)} \frac{S_A}{N_A} \\ \alpha_{AA}^{(s)} I_A^{(s,*)} &= \alpha_{AA}^{(g)} I_A^{(g,*)} + c\, \alpha_{AB}^{(g)} I_B^{(g,*)} \end{aligned} \tag{4.9}$$

Any susceptible host in A is equally likely to be infected with the generalist or specialist pathogen. Solving equation (4.9) for the endemic number of infectious hosts in species A yields

$$I_A^{(*)} = \frac{c\, \alpha_{AB}^{(g)}}{\alpha_{AA}^{(s)} - \alpha_{AA}^{(g)}} I_B^{(g,*)} \tag{4.10}$$

Equation (4.10) demonstrates the inter-dependence of the endemic states. The endemic number of infectious hosts in species A depends on the endemic number of infectious hosts with the generalist pathogen in species B. The endemic equilibrium condition (4.8) can be used to gain the endemic states in species B. It is

$$\frac{dI_B^{(g)}}{dt} = 0 = \lambda_B^{(g)} \frac{S_B^{(*)}}{N_B} - I_B^{(g,*)} \tag{4.11}$$

which leads to

$$I_B^{(g,*)} = \left(\alpha_{BB}^{(g)} I_B^{(g,*)} + c\, \alpha_{BA}^{(g)} I_A^{(*)} \right) \frac{S_B^{(*)}}{N_B} \tag{4.12}$$

Here, $I_B^{(g,*)}$ cannot be solved without knowledge of the endemic number of susceptible hosts in species B. It is possible to determine $S_B^{(*)}$ by looking at the endemic condition for the specialist pathogen in B

$$\frac{dI_B^{(s)}}{dt} = 0 = \alpha_{BB}^{(s)} I_B^{(s,*)} \frac{S_B^{(*)}}{N_B} - I_B^{(s,*)} \tag{4.13}$$

The number of new infections can only equal the number of recoveries if

$$\text{(a)}\ \ S_B^{(*)} = \frac{N_B}{\alpha_{BB}^{(s)}} \quad \text{or} \quad \text{(b)}\ \ I_B^{(s,*)} = 0 \tag{4.14}$$

There are two excluding conditions in species B. (a) The number of susceptible hosts is

of a fixed value, $N_B/\alpha_{BB}^{(s)}$, with possibly both types of pathogens present. This is the case at the transition from scenario I to scenario II. Furthermore, the specialist pathogen must be the predominant one or both, specialist and generalist, must be of equal disease prevalence as in species A. The endemic number of susceptible hosts in a standard SIS model with only one species and a single pathogen is $S^{(*)} = N/R_0$ (Keeling and Rohani, 2007). Therefore, the specialist's reproductive capability must be larger than the generalist's one as it determines the endemic number of susceptible hosts, or equal to the generalist's one in which case both pathogens infect the same number of hosts, explaining the character of co-existence. On the other hand, condition (b) requires the specialist strain to be extinct in species B. This is the case in scenario II, where one species does not have any infected hosts with the specialist pathogen. (b) describes therefore the transition from scenario II to scenario III. There, the endemic number of susceptible hosts is $S_B^{(*)} = N_B - I_B^{(g,*)}$. In addition, (4.14) shows that generalist predominance with specialist co-existence cannot be achieved in both species.

4.2.3.1 Threshold for Partial Generalist Predominance - From Scenario I to II

The conditions (a) and (b) are a mathematical formulation of the three different scenarios, because they describe the scenario-defining thresholds. Furthermore, they allow me to solve equations (4.10) and (4.12): (a) relates to the threshold $\hat{c}_1$ between scenario I and scenario II, with which I will start.

Substituting the endemic number of susceptible hosts, $N_B/\alpha_{BB}^{(s)}$, into equation (4.12) and solving for the endemic equilibrium number of infectious hosts yields

$$I_B^{(g,*)} = \frac{\hat{c}_1 \, \alpha_{BA}^{(g)}}{\alpha_{BB}^{(s)} - \alpha_{BB}^{(g)}} I_A^{(*)} \tag{4.15}$$

$\hat{c}_1$ can be derived by solving equation (4.10) using (4.15) which yields

$$\hat{c}_1 = \sqrt{\frac{\left(\alpha_{AA}^{(s)} - \alpha_{AA}^{(g)}\right)\left(\alpha_{BB}^{(s)} - \alpha_{BB}^{(g)}\right)}{\alpha_{AB}^{(g)} \, \alpha_{BA}^{(g)}}} \tag{4.16}$$

Note that the connectivity threshold is species size independent. Only the reproductive numbers define the threshold. Both pathogen types, the specialist and generalist, have equal reproductive capabilities within at least one of both species if, and only if, the connectivity equals the threshold. That is, the critical threshold (4.16) yields a connectivity for which both pathogens are of equal disease prevalence at the endemic equilibrium within one species. However, reaching the endemic equilibrium might take a long time.

4.2.3.2 Threshold for Full Generalist Predominance - From Scenario II to III

Condition (b) in (4.14) describes the transition from scenario II to scenario III, which are separated by the threshold $\hat{c}_2$. As expressed in (b), the generalist is already the only pathogen in one host species (which I arbitrarily chose to be B). Hence, the number of susceptible hosts at the endemic equilibrium is $S_B^{(*)} = N_B - I_B^{(g,*)}$ (see (4.14)). Substituted into (4.12), this yields

$$\begin{aligned} I_B^{(g,*)} = & \frac{1}{2\,\alpha_{BB}^{(g)}} \Big(\alpha_{BB}^{(g)}\, N_B - \hat{c}_2\, \alpha_{BA}^{(g)}\, I_A^{(*)} - N_B \\ & + \sqrt{\left(\alpha_{BB}^{(g)}\, N_B - \hat{c}_2\, \alpha_{BA}^{(g)}\, I_A^{(*)} - N_B \right)^2 + 4\, \alpha_{BB}^{(g)}\, \alpha_{BA}^{(g)}\, \hat{c}_2\, N_B\, I_A^{(*)}} \Big) \end{aligned} \tag{4.17}$$

In addition, (4.14) is a species-independent condition which allows me to re-write the endemic infectious population size

$$I_A^{(*)} = \frac{N_A - S_A^{(*)}}{2} = N_A \left(\frac{1}{2} - \frac{1}{2\,\alpha_{AA}^{(s)}} \right) \tag{4.18}$$

which, substituted together with (4.17) and $p = N_A/(N_A + N_B)$ into (4.10), yields

$$0 = \frac{\hat{c}_2\,\alpha_{AB}^{(g)}}{2\,\alpha_{BB}^{(g)}\left(\alpha_{AA}^{(s)} - \alpha_{AA}^{(g)}\right)}\left(\alpha_{BB}^{(g)}\,(1-p) - \hat{c}_2\,\alpha_{BA}^{(g)}\,p\left(\frac{1}{2} - \frac{1}{2\,\alpha_{AA}^{(s)}}\right) - (1-p)\right.$$
$$\left. + \sqrt{\left(\alpha_{BB}^{(g)}\,(1-p) - \hat{c}_2\,\alpha_{BA}^{(g)}\,p\left(\frac{1}{2} - \frac{1}{2\,\alpha_{AA}^{(s)}}\right) - (1-p)\right)^2 + 4\,\alpha_{BB}^{(g)}\,\alpha_{BA}^{(g)}\,\hat{c}_2\left(\frac{1}{2} - \frac{1}{2\,\alpha_{AA}^{(s)}}\right)}\right)$$
$$- p\left(\frac{1}{2} - \frac{1}{2\,\alpha_{AA}^{(s)}}\right) \tag{4.19}$$

It is possible to repeat these calculations for the threshold $\hat{c}_2$ in species B in an analogous way. This yields equation (4.19) with permuted species indices. However, it is a demanding task to find an algebraic, closed form for the threshold connectivity $\hat{c}_2$ from (4.19). Therefore, I will use numerical root finding methods to accurately estimate the $\hat{c}_2$s for each species, A and B.

4.3 Results

The theoretical analysis yields answers to my research questions. First, I asked for the basic reproductive number of a generalist pathogen in a multi-species system. I derived the expression for the basic reproductive number using the next-generation-matrix (see equation (4.5)). The result incorporates infections in both species. Furthermore, it allows for conclusions about the possible range of the connectivity c. Obviously, the connectivity cannot be smaller than 0 as this describes the case of two isolated species. In addition, an infectious host is equally likely to encounter a host from its own species or the other species for $c = 1$. A connectivity larger than 1 leads to a basic reproductive number which is larger than the reproductive numbers within each species and between the species combined. While such a large connectivity does not invalidate any of the previous derivations, it might not make sense when relating the theoretical model to an actual infectious disease. Hence, I assume $0 \leq c \leq 1$ as a realistic connectivity range for between-species interactions.

Second, I asked for the scenarios' dependency on the between-species connectivity, for

which I derived thresholds. Logical constraints of the theoretical analysis reveal further information on the three scenarios: (4.14) shows two conditions which state that the specialist must be extinct in one host species if the generalist is predominant in both species (see Figure 4.2 for a schematic representation). Conversely, a non-zero endemic number of hosts infected with the specialist in both species directly shows that the system must be in scenario I, and $c < \hat{c}_1$. The generalist can only sustain within both species if it is predominant in at least one of them (scenario II), which fuels the generalist prevalence in the other species with spill-over infections. Otherwise, it will go extinct due to the superior fitness of the specialist. Hence, I can re-write the three possible scenarios from the introduction as follows and answer my third research question:

I. only specialist pathogens exist in all both species,

II. a specialist and a generalist pathogen are each predominant in one host species with the specialist extinct where the generalist is predominant and co-existing in the other species,

III. or the generalist is the dominant pathogen in both species with possible co-existence in only one species.

Note that these scenarios ignore the possibility of mutations or any other kind of stochastic effects. They describe the final states at the endemic equilibrium. These scenarios do not describe the special cases where the connectivity equals a threshold value, $c = \hat{c}_1$ or $c = \hat{c}_2$. In each of these cases, both pathogen types can co-exist within a species as the reproductive capabilities are equal. Figure 4.3 shows the time series of the number of infected hosts in scenarios I and II, as well as the threshold between them at which $c = \hat{c}_1$. Figure 4.4 shows the time series of the number of infected hosts in scenarios II and III, and the threshold between them at which $c = \hat{c}_2$.

In a non-symmetric setting, there are overall three thresholds: one single $\hat{c}_1$ for both species, and two $\hat{c}_2$s, one for species A and one for species B. One of the $\hat{c}_2$s is larger than $\hat{c}_1$ and one is smaller than $\hat{c}_1$. Usually, $\hat{c}_1$ cannot be larger than both $\hat{c}_2$s, because the transition from scenario I to scenario II happens before the transition from scenario II to scenario III

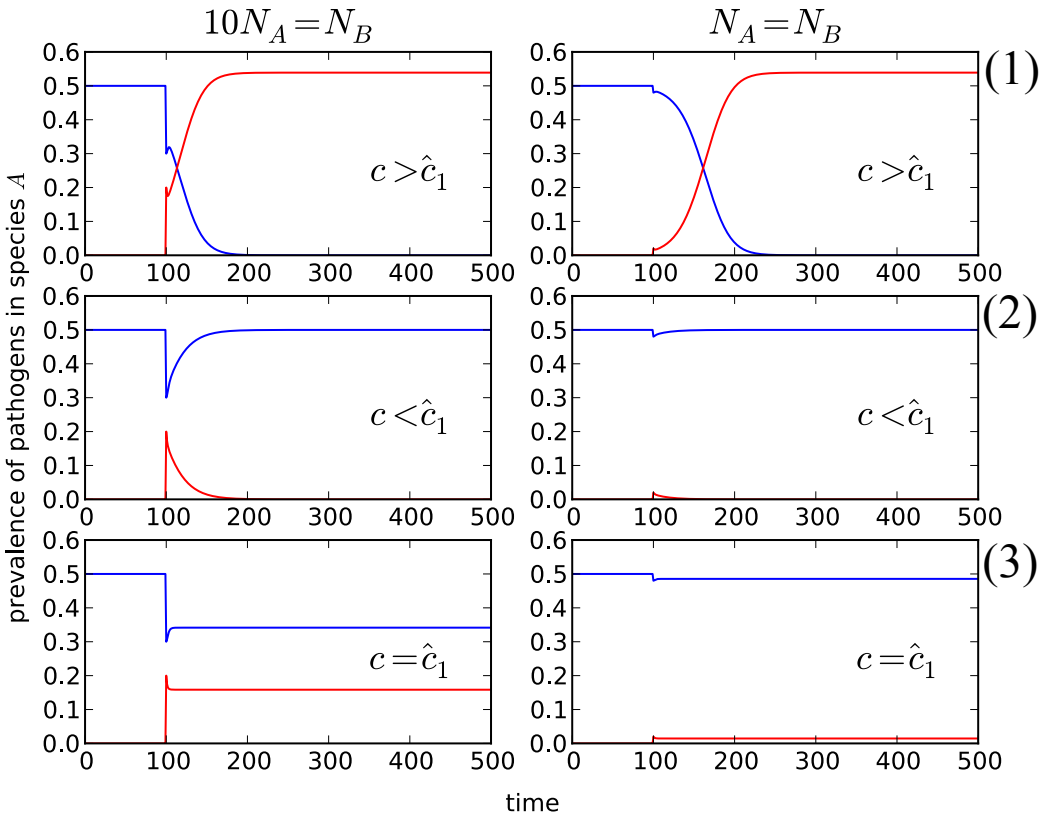

Figure 4.3: Exploring the threshold $\hat{c}_1$. Time series of the prevalence in species A for different connectivities and species sizes. The left column of plots is from a system with species B ten times larger than species A. The right column shows plots from a system with equally large species A and B. The systems is started in scenario I, i.e. both species only contain specialist pathogens represented with blue lines. Ten infectious hosts are converted to generalist infections at the time point $t = 100$ within species A, represented by the red lines. Plots in line **(1)** are the cases for a connectivity slightly above the threshold $\hat{c}_1$. The generalist pathogen becomes the predominant strain and the specialist goes extinct in A. This resembles scenario II. Plots in line **(2)** are the cases for a connectivity slightly below the threshold $\hat{c}_1$. The generalist pathogen gets introduced but fails to sustain within the species. At the end of the time series, only the specialist pathogen remains in species A. The system stays in scenario I. Plots in line **(3)** is the case for a connectivity equalling the threshold $\hat{c}_1$. Here, the generalist pathogen's prevalence first reduces after introduction, but stabilises at a non-zero value. The small dip in the number of generalist infections after introduction is explained by the need to establish generalist infections in species B. Once the number of generalist infections is sufficient in B, the number of generalist infections stabilises in both species. However, the time interval shown is not sufficient to reach the endemic equilibrium. The plots in line **(3)** reveal that a population size difference between the host species does not have an influence on the threshold $\hat{c}_1$.

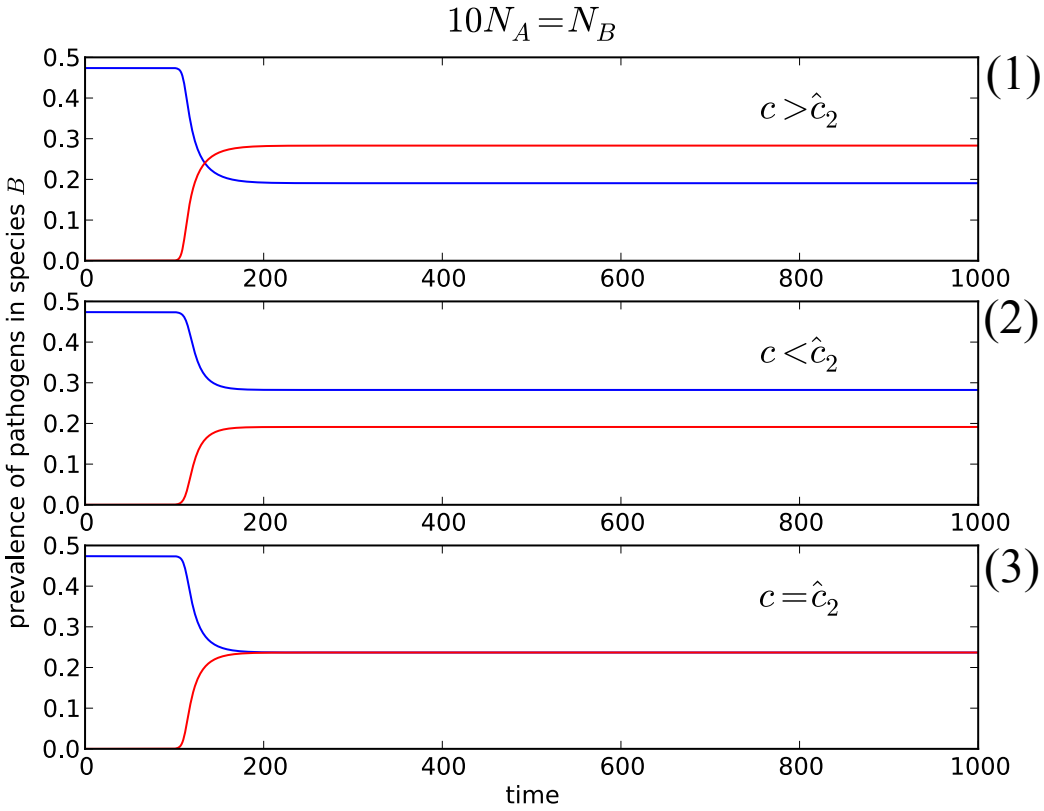

Figure 4.4: Exploring the threshold $\hat{c}_2$. Time series of the prevalence in species B for different connectivities. Here, species B is ten times larger than species A. The systems is started in scenario I, i.e. both species only contain specialist pathogens represented with blue lines. Ten infectious hosts are converted to generalist infections at the time point $t = 100$ within species A, represented by the red lines. **(1)** is the case for a connectivity slightly above the threshold $\hat{c}_2$. The generalist pathogen becomes the predominant strain but is in co-existence with the specialist pathogen. At the endemic equilibrium, the system is in scenario III. **(2)** is the case for a connectivity slightly below the threshold $\hat{c}_2$. The generalist pathogen gets introduced but fails to become predominant. It is in co-existence with the specialist, because of spill-over infections from A. The system is now in scenario II. **(3)** is the case for a connectivity equalling the threshold $\hat{c}_2$. The generalist becomes the predominant pathogen in A, and leads to co-existence in B with equal prevalence of the specialist and generalist.

for an increasing connectivity c. However, the derivation of $\hat{c}_2$ assumes that the system starts in scenario II. This condition is only fulfilled for one of the species' $\hat{c}_2$s. Figure 4.2 illustrates this: in scenario II, the generalist is the only pathogen in the left species. Hence, only $\hat{c}_2$ for the right species fulfils the condition and becomes the applicable threshold. $\hat{c}_2$ evaluated for the left species will yield a threshold connectivity smaller than $\hat{c}_1$, because it assumes that the generalist is the only pathogen in the right species - which is not the case.

But the non-applicable $\hat{c}_2$, corresponding to the $\hat{c}_2$ from the left species in Figure 4.2, offers further insights. The species for which $\hat{c}_2$ is smaller than $\hat{c}_1$ will be the species in which the specialist will go extinct at the transition from scenario I to scenario II. If $\hat{c}_2(A) < \hat{c}_1 < \hat{c}_2(B)$ the situation in Figure 4.2 pertains, i.e. the specialist pathogen goes extinct in species A but stays the dominant pathogen in B. Conversely, if $\hat{c}_2(A) > \hat{c}_1 > \hat{c}_2(B)$ the specialist pathogen will go extinct in species B as c becomes larger than $\hat{c}_1$.

There is a third possibility in which the system goes from scenario I to scenario III directly. Here, both species' $\hat{c}_2$ will be smaller than $\hat{c}_1$. The generalist will become predominant in both species at the same connectivity threshold $\hat{c}_1$.

The population size independence of $\hat{c}_1$ is a surprising result, answering my fourth research question. The threshold connectivity $\hat{c}_1$ only depends on the respective reproductive numbers. Figure 4.3 illustrates this fact by comparing the prevalence of the generalist and specialist for different host species' sizes. Figure 4.5 shows $\hat{c}_1$ in a symmetrical system. In such a system, the reproductive numbers are the same for both species and equation (4.16) reduces to $\hat{c}_1 = (\alpha_{AA}^{(s)} - \alpha_{AA}^{(g)})/\alpha_{AB}^{(g)}$. Figure 4.5 is diagonally separated into two distinctive sections: the coloured, upper left half represents the area for which generalist predominance in at least one species is possible. The white, lower right half is the area for which specialist pathogens will always be predominant because $\hat{c}_1 > 1$.

On the other hand, the threshold separating scenario II and scenario III, $\hat{c}_2$, is population size dependent (see Figure 4.3). To be exact, it depends on the fraction of the species' host population size in the total number of hosts $p = N_A/(N_A + N_B)$ (see equation (4.19)). Not the actual number of hosts but the relative size difference matters.

Figure 4.6 gives an example of connectivity thresholds for a sample of population size differences and between-species reproductive numbers. The chosen parameters yield a

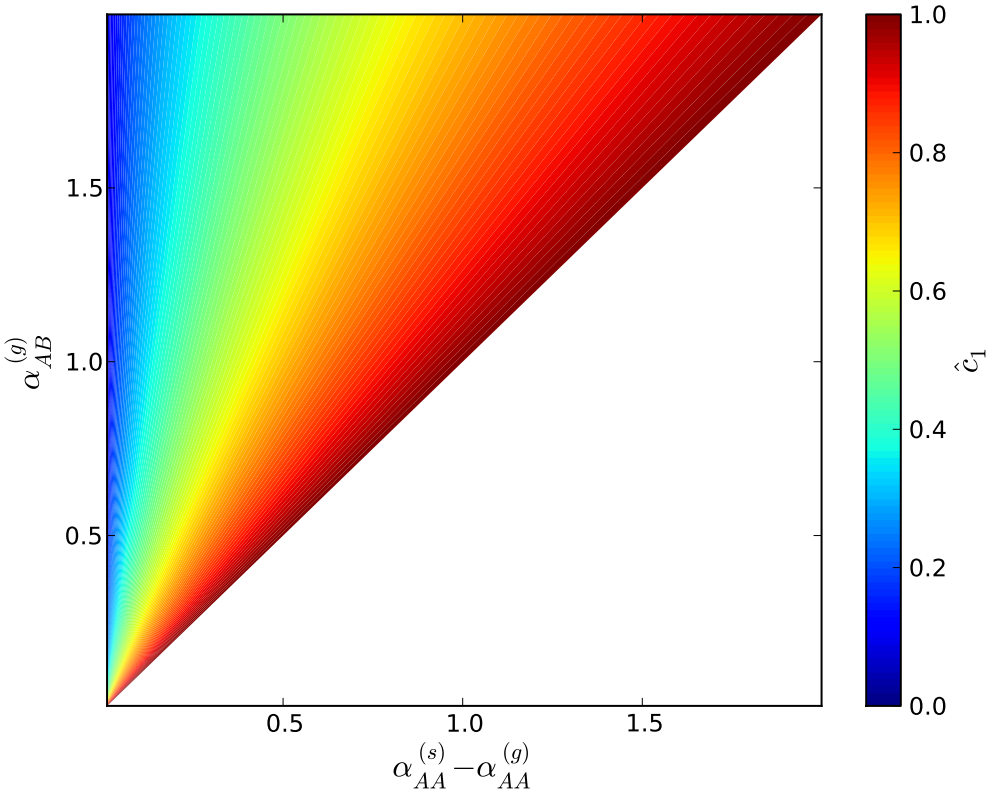

Figure 4.5: Connectivity threshold $\hat{c}_1$ dependence on reproductive numbers (see equation (4.16)). Shown is the value of $\hat{c}_1$ for a symmetric system with $\alpha_{AA}^{(s/g)} = \alpha_{BB}^{(s/g)}$ and $\alpha_{AB}^{(g)} = \alpha_{BA}^{(g)}$. The white coloured space represents the area for which $\hat{c}_1 > 1$. In this area, no generalist predominance is possible because the connectivity cannot be larger than unity.

symmetric system, i.e. the reproductive numbers are the same in each species. This can also be seen because for $p = 0.5$, $\hat{c}_2$ is the same in both species. Furthermore, the position of $\hat{c}_2$ relative to $\hat{c}_1$ informs us in which species the generalist becomes predominant first. As previously discussed, the condition for this to happen is if $\hat{c}_2$ lies below $\hat{c}_1$. For a range of host species sizes around $p = 0.5$, $\hat{c}_2$ is lower than $\hat{c}_1$ in both species. Hence, a generalist pathogen will become simultaneously predominant upon reaching the connectivity threshold $\hat{c}_1$.

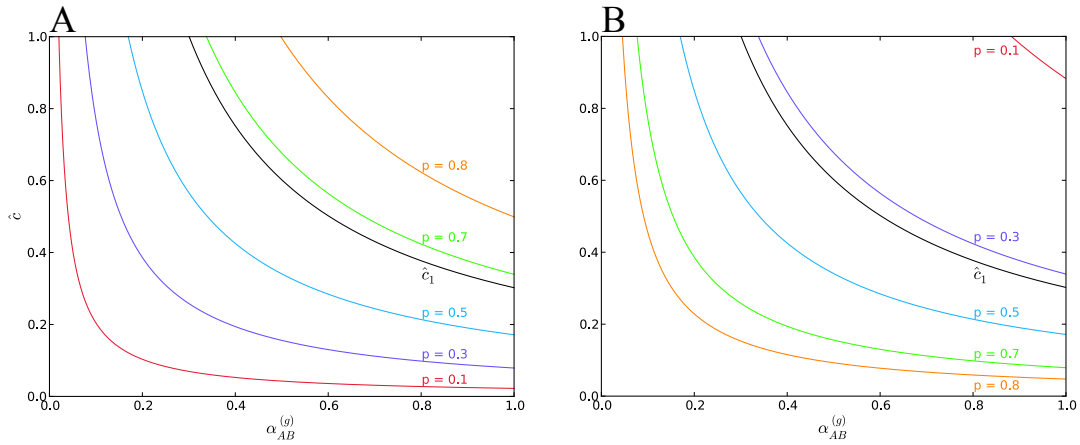

Figure 4.6: Connectivity thresholds $\hat{c}$ as a function of the between-species reproductive number for a symmetric system. Here, $\alpha^{(s)}_{AA} = \alpha^{(s)}_{BB} = 2$ and $\alpha^{(g)}_{AA} = \alpha^{(g)}_{BB} = 1.7$. The black line represents the connectivity threshold $\hat{c}_1$ which is independent of host population sizes. $\hat{c}_2$ is represented with coloured lines, grouped and ordered according to the relative host population size difference with $p = N_A/(N_A + N_B)$. **A** presents $\hat{c}_2$ evaluated for species A, and **B** for species B. The blue line for $\hat{c}_2$ with $p = 0.5$ is the same in both figures, because the reproductive numbers have equal values in both species. If $\hat{c}_2 < \hat{c}_1$, the generalist pathogen will become predominant in this species once the connectivity reaches the threshold $\hat{c}_1$. This is also simultaneously the case in both species for a certain range of p. Therefore, a generalist pathogen will simultaneously become predominant in both species for this species size difference.

Figure 4.7 illustrates the regimes of co-existence and predominance according to the relative population size difference p. As previously discussed, there are three different regimes: I, where the specialist is the exclusive pathogen in both species; II, where the generalist is predominant in one species and the specialist in the other; and III, where the generalist is predominant in both species but co-existence might be possible in one species. Moreover, Figure 4.7 D shows that the generalist might even become predominant in the

larger host species first if the reproductive numbers are much smaller (here factor ten, see Table 4.1). If the reproductive numbers are of comparable size, the generalist pathogen usually becomes predominant in the smaller host species or simultaneously in both (see Figure 4.7 A - C).

Detailed reproductive numbers for Figure 4.7				
A	$\alpha^{(s)}_{AA} = 2$	$\alpha^{(s)}_{AB} = 0$	$\alpha^{(g)}_{AA} = 1.7$	$\alpha^{(g)}_{AB} = 0.8$
	$\alpha^{(s)}_{BA} = 0$	$\alpha^{(s)}_{BB} = 2$	$\alpha^{(g)}_{BA} = 0.8$	$\alpha^{(g)}_{BB} = 1.7$
B	$\alpha^{(s)}_{AA} = 2$	$\alpha^{(s)}_{AB} = 0$	$\alpha^{(g)}_{AA} = 1.9$	$\alpha^{(g)}_{AB} = 0.2$
	$\alpha^{(s)}_{BA} = 0$	$\alpha^{(s)}_{BB} = 2$	$\alpha^{(g)}_{BA} = 0.2$	$\alpha^{(g)}_{BB} = 1.9$
C	$\alpha^{(s)}_{AA} = 2$	$\alpha^{(s)}_{AB} = 0$	$\alpha^{(g)}_{AA} = 1.7$	$\alpha^{(g)}_{AB} = 0.7$
	$\alpha^{(s)}_{BA} = 0$	$\alpha^{(s)}_{BB} = 1.5$	$\alpha^{(g)}_{BA} = 0.7$	$\alpha^{(g)}_{BB} = 1.3$
D	$\alpha^{(s)}_{AA} = 24$	$\alpha^{(s)}_{AB} = 0$	$\alpha^{(g)}_{AA} = 20$	$\alpha^{(g)}_{AB} = 1.9$
	$\alpha^{(s)}_{BA} = 0$	$\alpha^{(s)}_{BB} = 2.4$	$\alpha^{(g)}_{BA} = 1.9$	$\alpha^{(g)}_{BB} = 2$

Table 4.1: Reproductive numbers for Figure 4.7 **A** - **D**. **A** and **B** are symmetrical while **C** and **D** are not. Furthermore, in **D** the within-species reproductive numbers are ten-times higher within species A than within B.

4.4 Discussion

Much ecological literature concerns how infectious diseases influence the co-existence of host species (Holt and Pickering, 1985; Holt and Lawton, 1993). My research in this chapter focuses on the effect of host species' interactions on pathogen specialisation and co-existence. The insights presented here are important for the understanding of infectious agents' cross-species interactions as the interactions between host and pathogen are bi-directional: infectious diseases influence host species and can even drive them to extinction (Holt and Pickering, 1985). But the host species inter-connections also influence the nature of the pathogen. I have shown that a change in the ecological host structure, and therefore the pathogen's environment, can induce dramatic changes in the pathogen's host range. These results improve our understanding of pathogen specialisation, and might have a direct application on the design and evaluation of experimental studies.

My results are in line with previous notions that 'change' seems to facilitate the emergence of novel infectious diseases (Cohen, 2000). A pathogen might be specialised on one

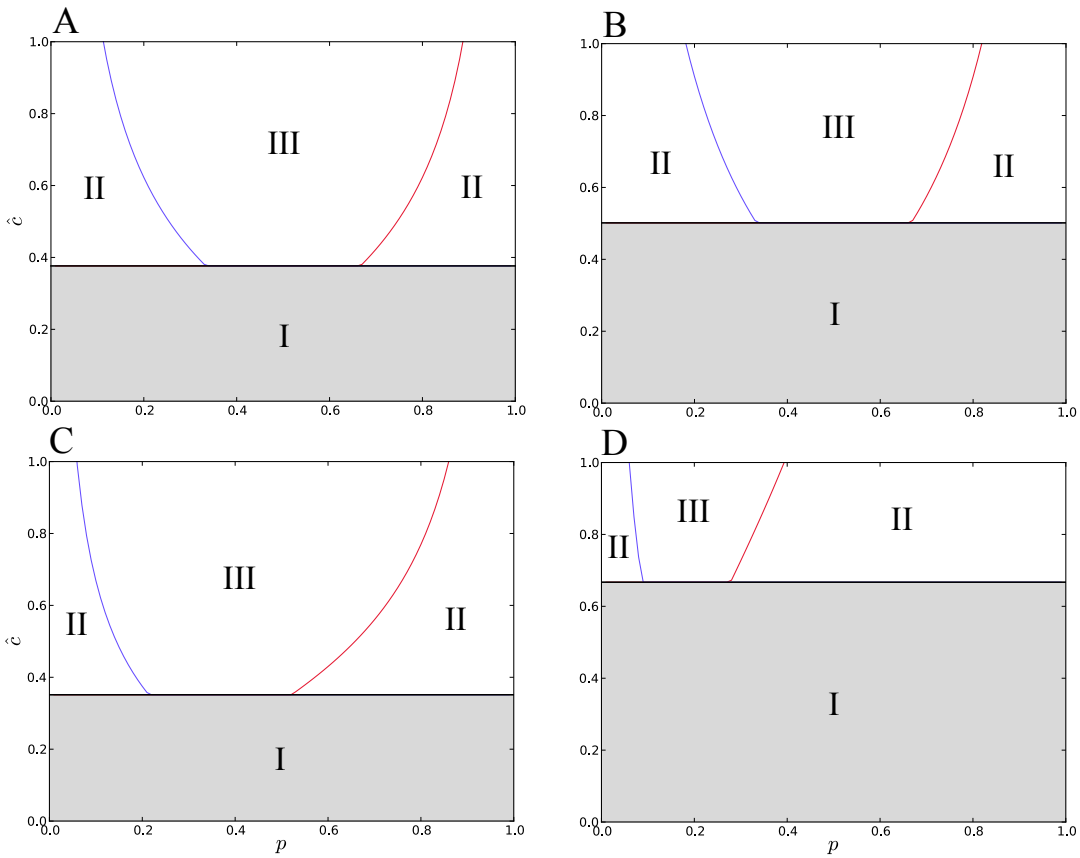

Figure 4.7: Predominance and co-existence regimes for specialist and generalist pathogens according to relative population size differences. Here, $p = N_A/(N_A + N_B)$ is the fraction of one host species from the total number of hosts. Areas marked with III represent regimes with generalist predominance in both species. An II represents generalist predominance in one species, but specialist predominance with generalist co-existence in the other. Both regimes are with white background as they contain some sort of generalist predominance. On the other hand, I marks parameter sets for which specialists are the only pathogen in both species. This area is greyed out. The white area is separated from the grey area through $\hat{c}_1$, the connectivity threshold for generalist predominance. It is represented with a black horizontal line. The blue line represents $\hat{c}_2 \geq \hat{c}_1$ in species A and the red line $\hat{c}_2 \geq \hat{c}_1$ in species B. It therefore separates regime I and II as $\hat{c}_2$ is the threshold for generalist predominance in both species. **A** and **B** show regimes for symmetrical reproductive numbers, while **C** and **D** illustrate non-symmetrical systems (see Table 4.1 for all reproductive numbers in detail). While in **A** - **C** the generalist first becomes predominant in the smaller species A or in both at the same time, **D** illustrates a case for which the generalist can become predominant first in the larger species B.

species which is connected to other ones. A change in this connectivity or a change in relative species size differences might shift the system over a threshold into a new regime of generalist predominance. Hence, a generalist strain might now have the chance to develop by evolution and spread the disease more easily into new species. Experimental data with bacteria and phage in which an intermediate generalist strain evolves to infect new bacteria support this view (Benmayor et al., 2009; Dennehy et al., 2010). Some of the 'ancient' human diseases like measles might even have entered the human population through such a mechanism, facilitated by the previous, significant changes in the human environment (McMichael, 2001, 2004; Weiss and McMichael, 2004).

But my results may also help to answer more fundamental questions about pathogen specialisation. Despite fitness costs within a single species (Woolhouse et al., 2001), generalists might outperform specialists because of the larger pool of susceptible hosts compensating the handicap of lower transmission rates. This becomes directly apparent by looking at the average reproductive number (see equation (4.6)).

However, my model neglects some interactions. The model system presented here assumes a fixed host population size over time. The pathogen's prevalence has no effect on the host population, which neglects any regulatory effects of reduced population sizes on disease transmission. While previous research only focused on the host species effects, I only focus on the pathogen effects. Therefore, this model should be regarded as a first step towards understanding the complex connections between host and pathogen. A next step would be a model incorporating both levels of influence to increase our knowledge on ecological specialisation and its underlying factors.

CHAPTER 5

THE IMPACT OF A CHANGING BIODIVERSITY ON DISEASE PREVALENCE

5.1 Introduction

The importance of 'change' for infectious disease transmission has been one of the main themes in previous chapters. In this chapter, I will focus on a different environmental aspect of 'change': what happens if not only the interactions between species change but the composition of species within a defined environmental niche? This is often referred to as change in biodiversity. It appears that, on a world-wide count, the number of species is decreasing, a phenomenon called loss of biodiversity (Epstein, 2001; Keesing et al., 2006). One assumption is that 'weedy' species are more likely to withstand ecological changes like climate change, and are faster to re-establish in new ecological niches (Stearns, 1992; Pilgrim et al., 2004), but have a higher pathogen prevalence (Keesing et al., 2010). This leads to my research question: do 'weedy' species have a higher prevalence of infection? The answer is important to zoonotic EID research. If higher prevalence facilitates inter-species jumps, and therefore emergence probabilities, the loss of biodiversity might have an overall effect of increasing the occurrence of zoonotic EIDs.

One way to describe 'weedy' species is with r/K-selection theory (Stearns, 1992). In this, r describes the logistic growth rate and K the maximum carrying capacity, representing two kinds of selection paradigms. Moreover, these are thought of to be of opposing kinds. Environmental conditions might favour r-selected, fast growing organisms or K-selected, resource maximising organisms (Pianka, 1970). Therefore, each organism has a defined growth rate and carrying capacity in its ecological niche. One can think of r/K-selection as two opposing endpoints on a continuum.

Totally r-selected organisms are in the quantitative extreme. Imagine an almost empty island without any competition or other density dependent effects. The organisms that will perform best on this island are those whose life history maximises the number of progeny. Hereby, the energy spend on each offspring is minimised in favour of increasing the reproductive capabilities. Hence, r-selection describes 'weedy' species reproduction. On the other hand, totally K-selected organisms are in the opposite extreme. Imagine another island saturated with organisms with maximal density dependent effects. Competition is fierce and an r-selected strategy with many uncompetitive offspring will not be successful.

The optimal strategy is to channel all energy into a few very fit descendants. The number of progeny is small, but the use of resources is optimised to allow the few offspring to successfully compete and survive.

In the following, I will use the r/K-selection theory to derive the prevalence for r- and K-selected host species. I will start with theoretical derivations and explore the impact on infectious disease prevalence.

5.2 Exploration of the Theoretical Model

As mentioned above, r is the logistic growth rate and K the maximum carrying capacity for a species in a defined niche. The population size N can be described with a logistic function

$$\frac{dN}{dt} = r\,N\left(1 - \frac{N}{K}\right) \tag{5.1}$$

The population will grow or decline with the growth rate r until it reaches its carrying capacity K.

Next, I will use the logistic growth function (5.1) to look at the influence of r and K on infectious disease prevalence for different disease transmission types.

5.2.1 Prevalence with Density-Dependent Transmission

5.2.1.1 Susceptible-Only Reproduction

First, I assume that the pathogen gets directly transmitted with a density-dependence of infected and susceptible hosts (see section 2.3 for an overview of different compartmental models). I assume that only susceptible hosts can reproduce and have offspring. This is the case if the infectious disease hinders host reproduction, or if the infectious period is very small. Furthermore, all new hosts are susceptible to the infectious disease when they enter the population. I assume that infectious hosts spread the disease according to the infectious rate β, and are removed from the population with the removal rate γ. A simple

SI model reads

$$\begin{aligned} \frac{dS}{dt} &= r\,S\left(1-\frac{S}{K}\right)-\beta\,S\,I \\ \frac{dI}{dt} &= \beta\,S\,I-\gamma\,I \end{aligned} \tag{5.2}$$

I am interested in the prevalence P which is defined as $P = I^{(*)}/N^{(*)}$. As in previous chapters, the asterisk is a symbol for the endemic equilibrium values and the population size is $N^{(*)} = S^{(*)} + I^{(*)} = \text{const}$. At the endemic equilibrium, the flux of new infections equals the flux of removals. It is synonymous with the derivatives equalling zero in equation (5.2). Solving for the equilibrium states, I obtain

$$\begin{aligned} S^{(*)} &= \frac{\gamma}{\beta} \\ I^{(*)} &= \frac{r}{\beta}\left(1-\frac{S^{(*)}}{K}\right)=\frac{r}{\beta}\left(1-\frac{\gamma}{\beta\,K}\right) \end{aligned} \tag{5.3}$$

This yields a prevalence of

$$P = \frac{r\,(\beta\,K-\gamma)}{\beta\,K\,(\gamma+r)-r\,\gamma} \tag{5.4}$$

The prevalence is a function of r and K, and increases monotonically with the growth rate r.

5.2.1.2 General Reproduction

What happens if infected hosts are able to reproduce offspring? The prevalence as defined in equation (5.4) is only an approximation of this case. But I can derive the exact prevalence by starting with the SI model

$$\begin{aligned} \frac{dS}{dt} &= r\;(S+I)\left(1-\frac{S+I}{K}\right)-\beta\,S\,I \\ \frac{dI}{dt} &= \beta\,S\,I-\gamma\,I \end{aligned} \tag{5.5}$$

This system yields an endemic equilibrium of

$$\begin{aligned} S^{(*)} &= \frac{\gamma}{\beta} \\ I^{(*)} &= \frac{\beta K (r-\gamma) - 2\gamma r}{2\beta r} + \sqrt{\frac{K\left(\beta K (\gamma - r)^2 + 4\gamma^2 r\right)}{4\beta r^2}} \end{aligned} \tag{5.6}$$

Using the endemic states (5.6), I can calculate a prevalence of

$$P = 1 - \frac{2\gamma r}{\beta K (r-\gamma) + \sqrt{\beta K\left(\beta K (\gamma - r)^2 + 4\gamma^2 r\right)}} \tag{5.7}$$

As before in (5.4), the prevalence is monotonically increasing with the growth rate r and decreasing with the carrying capacity K.

5.2.2 Prevalence with Frequency-Dependent Transmission

5.2.2.1 Susceptible-Only Reproduction

In a second step, I assume that pathogen transmission is frequency-dependent. Again as in section 5.2.1.1, the first case describes the situation where only susceptible hosts can reproduce and have offspring. Hence (5.2) becomes

$$\begin{aligned} \frac{dS}{dt} &= r S \left(1 - \frac{S}{K}\right) - \beta \frac{S I}{S + I} \\ \frac{dI}{dt} &= \beta \frac{S I}{S + I} - \gamma I \end{aligned} \tag{5.8}$$

with the endemic states

$$\begin{aligned} S^{(*)} &= \frac{\gamma}{\beta - \gamma} I^{(*)} = \frac{\gamma K (r - \beta + \gamma)}{r\gamma} \\ I^{(*)} &= \frac{K(\beta - \gamma)(r - \beta + \gamma)}{r\gamma} \end{aligned} \tag{5.9}$$

Note that $S^{(*)}$ and $I^{(*)}$ cannot be of negative value. Positive endemic states are only possible for the following condition

$$r > \beta - \gamma \tag{5.10}$$

An infectious disease will not be able to cause sustained transmission chains if the growth rate is below this threshold. Instead, infections with the pathogen will drive the population extinct. Previous research on HIV came to a similar result which showed that the HIV virus might cause a negative growth of a population (Anderson et al., 1988).

The prevalence is comparatively easy to derive because of the interconnection between $S^{(*)}$ and $I^{(*)}$. It is

$$P = \begin{cases} 1 - \frac{\gamma}{\beta} & \text{if condition (5.10) is fulfilled} \\ 0 & \text{otherwise} \end{cases} \tag{5.11}$$

Clearly, P is independent of r and K if it is above the critical threshold for disease establishment. It is a step-function where the location of the step is defined by the epidemiological parameters of the disease, and the host specific growth rate r decides if sustained pathogen transmission is possible or not. Hence, the prevalence is of two fixed values and only depends on the epidemiological parameters of the infectious disease once the disease becomes endemic.

5.2.2.2 General Reproduction

As before, I extend the reproductive capability to include possible progeny of infectious hosts. With the frequency-dependence, (5.5) becomes

$$\begin{aligned} \frac{dS}{dt} &= r\ (S+I)\left(1 - \frac{S+I}{K}\right) - \beta\frac{S\,I}{S+I} \\ \frac{dI}{dt} &= \beta\frac{S\,I}{S+I} - \gamma\,I \end{aligned} \tag{5.12}$$

This differential equations yield endemic equilibrium states of

$$\begin{aligned} S^{(*)} &= \frac{\gamma}{\beta-\gamma}I^{(*)} = -\frac{\gamma}{\beta-\gamma}\frac{K\,(\beta-\gamma)\,\left(\beta\,(\gamma-r)-\gamma^2\right)}{\beta^2\,r} \\ I^{(*)} &= -\frac{K\,(\beta-\gamma)\,\left(\beta\,(\gamma-r)-\gamma^2\right)}{\beta^2\,r} \end{aligned} \tag{5.13}$$

Again, $S^{(*)}$ and $I^{(*)}$ cannot be of negative value. Hence the endemic states (5.13) are only valid for a particular subset of parameters. Indeed, the following condition defines this

subset of parameters for which the pathogen will not drive the population to extinction

$$r > \gamma - \frac{\gamma^2}{\beta} \tag{5.14}$$

Equation (5.14) shows one condition for reaching a non-negative endemic states. The other condition is $\gamma \geq \beta$ which follows from (5.13). However, γ cannot be equal to or larger than β as the basic reproductive number, defined as $R_0 = \beta/\gamma$, must be larger than unity for an infectious disease to establish itself (see section 2.2). This leaves (5.14) as the only valid condition.

As for (5.11), the prevalence follows directly from the values of $S^{(*)}$, $I^{(*)}$, and (5.14). It is

$$P = \begin{cases} 1 - \frac{\gamma}{\beta} & \text{if condition (5.14) is fulfilled} \\ 0 & \text{otherwise} \end{cases} \tag{5.15}$$

The prevalence is again a step-function with two fixed values. It depends on the epidemiological parameters of the infectious disease and the growth rate in that these determine the point of the step from zero to non-zero.

5.2.3 Stochastic Simulation of the Model

Each of these models ((5.2), (5.5), (5.8), and (5.12)) is translated into a stochastic simulation using the Gillespie algorithm (Gillespie, 1977). This approach is aimed to validate the derived prevalences. Furthermore, it gives an insight into the variation of the theoretically derived value, which cannot be assessed through the derivations. Each infection is seeded in a single random, susceptible host. All introductions which lead to immediate extinction of the pathogen are excluded, because I am only interested in the prevalence of successful epidemics rather than the probability of emergence. The prevalence is then derived using the number of susceptible and infectious hosts after the change in these values is only minimal, which corresponds to the endemic equilibrium state. I use $1,000$ individual, successful simulations to derive the average prevalence and compare this value with the analytically derived one (see Figure 5.1).

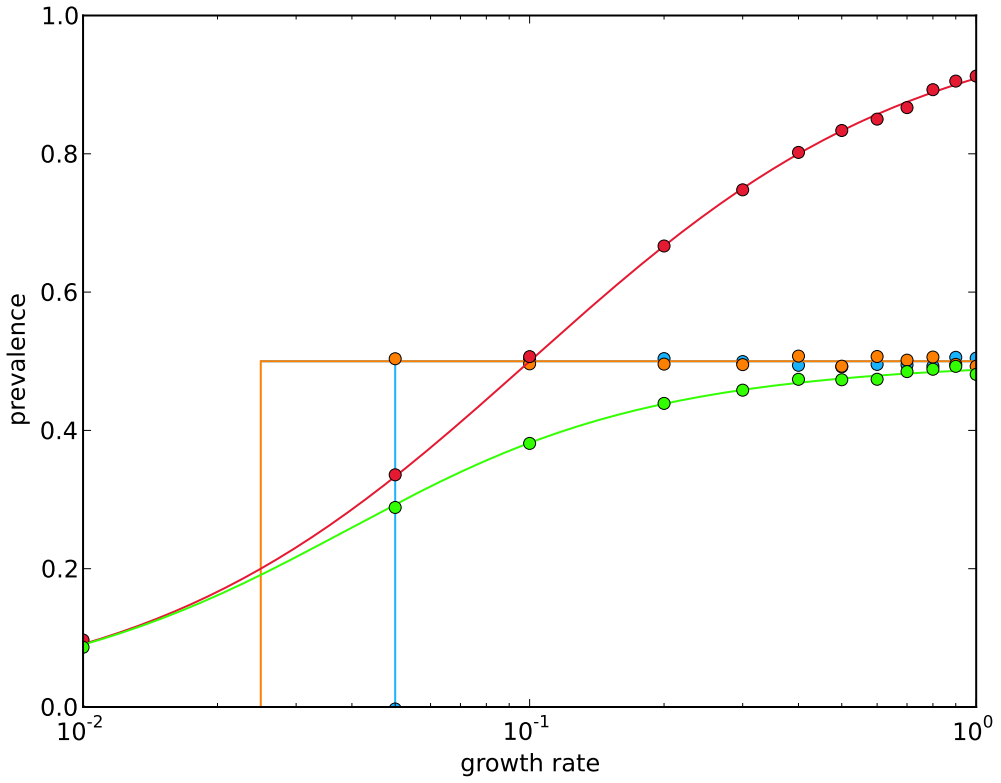

Figure 5.1: Prevalence p as a function of r on a semi-logarithmic scale. I assume $R_0 = 2$, $\gamma = 0.2$ and $K = 10000$. The density-dependent prevalence where only susceptible hosts reproduce (5.4) is shown in red. The density-dependent prevalence where all hosts reproduce (5.7) is shown in green. The frequency-dependent prevalence where only susceptible hosts reproduce (5.11) is shown in blue. The frequency-dependent prevalence where all hosts reproduce (5.15) is shown in orange. The last two prevalences are horizontal lines with vertical lines marking the threshold steps $r = 0.05$ and $r = 0.025$ for positive endemic states in (5.11) and (5.15). The coloured circles are the averaged result of $1,000$ simulations with the respective parameters. The stochastic variation is small, hence the variance in prevalence is small. Furthermore, (5.4) is a good approximation of (5.7) for low growth rates. For larger growth rates, (5.7) goes asymptotically to (5.15), and (5.4) to unity.

5.3 Results

A comparison between density-dependent transmission systems and frequency-dependent ones is best done by using a similar basic reproductive number and infectious period (see section 2.2). For density-dependent systems, it is known that $R_0 = K\beta/\gamma$. Frequency-dependent systems have an basic reproductive number of $R_0 = \beta/\gamma$. Setting a fixed R_0 and infectious period $1/\gamma$ allows for comparison of the prevalence and conditions for the different systems.

My theoretical analysis in section 5.2 illustrates that only systems with density-dependent transmission routes show an influence of r and K on the prevalence of infection. Figure 5.1 illustrates this in a semi-logarithmic plot. It reveals that density-dependent transmission leads, for increasing growth rates, to increased prevalence. For small growth rates, the prevalence is very small and near zero. It tends to unity if infectious hosts do not reproduce, and to the constant prevalence of systems with frequency-dependence if all hosts have progeny. Furthermore, the difference between the two density-dependent prevalences, (5.4) and (5.7), is negligible for small growth rates. The prevalences for systems with frequency-dependent transmission are step-functions between zero and $1 - 1/R_0$. The position of the steps depends solely on the epidemiological parameters of the pathogen. The stochastic simulations reveal that the variations around the theoretically derived prevalences are small.

Figure 5.2 shows the prevalence of a density-dependent model with general reproduction (5.7). It is a function of the basic reproductive number and the growth rate while the carrying capacity and the infectious period are fixed. It illustrates very well the non-linear dependence on the growth rate.

Both Figures 5.1 and 5.2 support the assumption that 'weedy' species show a higher prevalence of infection. However, the quality of the dependence is different for density-dependent and frequency-dependent transmission. Frequency-dependent transmission yields a step-function in prevalence as a function of growth rate. The position of the step depends on the growth rate, but the values of the prevalence before and after the step are independent of the growth rate. The conditions for endemic states, and therefore sustained

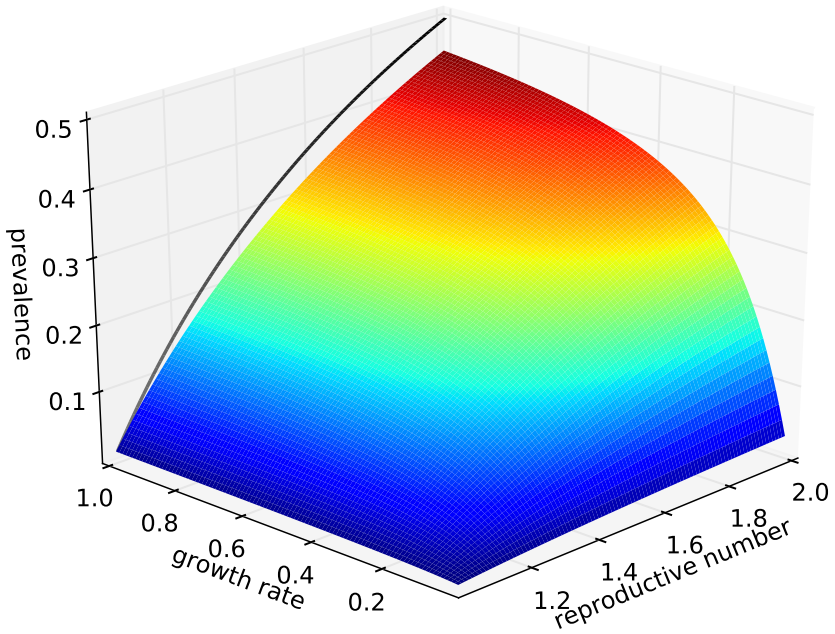

Figure 5.2: Prevalence P for the case of general reproduction as a function of the growth rate r and the basic reproductive number R_0 as presented in equation (5.7). The underlying system includes density-dependent disease transmission for a removal rate of $\gamma = 0.2$ and a carrying capacity of $K = 10000$. The black line shows the prevalence of frequency-dependent systems for comparison.

infectious disease transmission, are outlined in (5.10) and (5.14). For frequency-dependent transmission, it relies on the growth rate r if a disease drives a population extinct. Figures 5.3 and 5.4 illustrate the critical growth rate as a function of the infectious period and basic reproductive number. Both figures have in common that the threshold growth rate increases with the reproductive number, and decreases with the infectious period.

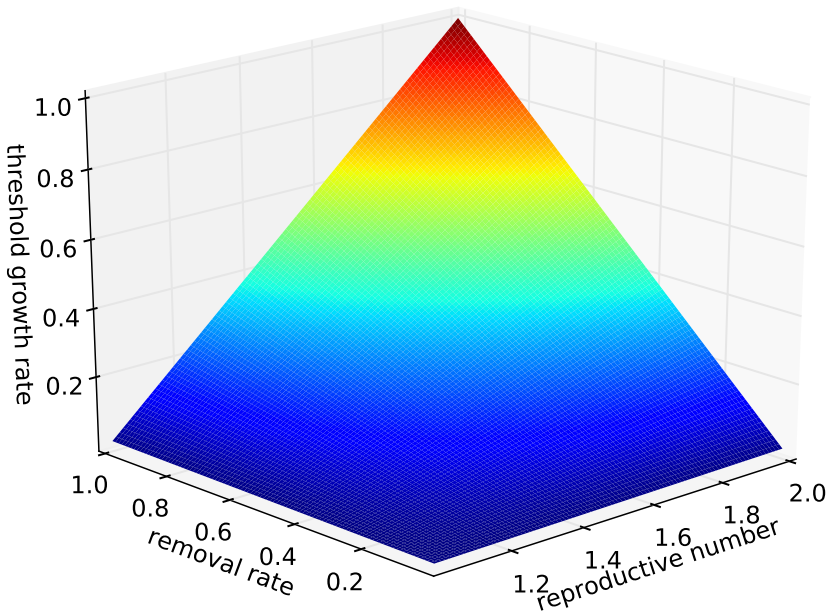

Figure 5.3: Threshold growth rate as defined in condition (5.10). No sustained disease transmission is possible if the growth rate is, for a given removal rate and basic reproductive number, below the critical threshold. The threshold linearly increases with the removal rate γ and the basic reproductive number R_0. Long-lasting infections have a lower growth rate threshold, and are therefore more likely to infect K-selected host species. Furthermore, R_0 has an influence on the threshold, allowing pathogens with small reproductive capabilities to invade K-selected host species, too.

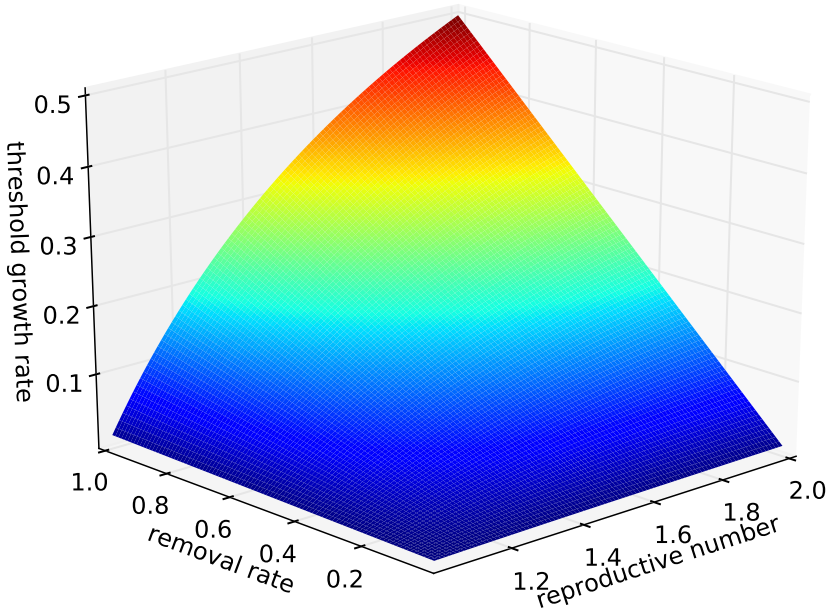

Figure 5.4: Threshold growth rate as defined in condition (5.14). No sustained disease transmission is possible if the growth rate is, for a given removal rate and basic reproductive number, below the critical threshold. The threshold linearly increases with the removal rate γ and non-linearly with the basic reproductive number R_0. It is therefore similar to Figure 5.3, but overall the threshold growth rate is approximately a half smaller for a given removal rate and reproductive number.

5.4 Discussion

I initially asked: do 'weedy' species have a higher prevalence of infection? Overall, I believe this is the case. I used the framework of r/K-selection theory to address the question and derive algebraic expressions describing disease prevalences for different situations. I found that the prevalence depends on the growth rate and carrying capacity for density-dependent disease transmission. On the other hand, in systems with frequency-dependent disease transmission prevalence depends only on the growth rate r. The influence of the disease's transmission type (density- or frequency-dependent) has also been shown by previous research on multi-host systems (Dobson, 2004; Keesing et al., 2006).

Furthermore, the prevalence for frequency-dependent transmission is a step-function between zero and a value depending on the basic reproductive number of the infectious disease. The epidemiological parameters of the infectious disease determine a threshold growth rate above which sustained disease transmission is possible. The possible values of the prevalence are therefore independent of r and K, but the decision which value it is depends on the growth rate. One possibility is to think of it as a filter: the growth rate determines an infectious disease's maximum reproductive number for a given infectious period within this host species. If the basic reproductive number is above this threshold, it will drive the host species extinct.

Previous research supports this connection between growth rates and extinction through pathogen infection. May et al. (1988) show that HIV might drive a human population extinct if its growth rate lies below a specific threshold. They used a frequency-dependent model framework, which in my case also yields a growth rate threshold necessary to sustain the population.

However, it is a non-trivial task to find supporting empirical evidence for these results. Comparing disease prevalence in r-selected and K-selected species might lead to false conclusions, because the epidemiological parameters of the compared infectious diseases, as well as the environmental conditions in which these species live, might be very different and have an unforeseeable impact on diseases prevalence. In an ideal situation, a pathogen would infect two different species, one r-selected and the other one K-selected, in the

same environment with the same epidemiological rate constants. Despite a careful search, I cannot find any such data.

These findings shed new light on the importance of a changing biodiversity. If *r*-selected species are more likely to withstand extinction (Stearns, 1992; Keesing et al., 2010), then a general rise in prevalence is expected as the *K*-selected species with low prevalence disappear. As shown previously in chapter 3, this increases the risk for inter-species jumps and has a direct influence on the probability of establishing zoonotic EIDs.

CHAPTER 6

THE INFLUENZA A(H1N1) PANDEMIC IN ENGLAND

6.1 Introduction

In the introductory chapter, I focus on the underlying factors and influences for a zoonotic emergence. I use VEEV as an example to illustrate the interplay between the different influences, mechanisms, and processes that drive an emergence. Yet, the most prominent example of a recent zoonotic EID is the 'Swine flu' influenza A(H1N1) pandemic of 2009. This chapter examines the epidemiological behaviour of the 'Swine flu' epidemic in England. The material in this chapter is published as Kubiak and McLean (2012).

In the spring of 2009 an influenza pandemic arose in North America. The first influenza A(H1N1) cases were recorded in March 2009 in Mexico (Fraser et al., 2009). Shortly after, infections were reported in the USA (Cauchemez et al., 2009; Lessler et al., 2009) and Canada (Tuite et al., 2010). Infection soon became a global phenomenon (McBryde et al., 2009; de Silva et al., 2009): the first laboratory confirmed cases in the United Kingdom were on 27 April (Health Protection Agency, 2009c), and by the end of May more than 50 countries had confirmed cases. The World Health Organization formally declared the outbreak a pandemic on 11 June 2009 (Donaldson et al., 2009).

As the pandemic spread around the world it was natural to ask "how bad will it be?", and attempts to forecast epidemic impacts were rapidly assembled, broadcast, and then published. Several of these studies focused on estimating the new influenza's basic reproductive number R_0, defined as the number of secondary cases caused by each case as an epidemic spreads into a population with no pre-existing immunity (Anderson and May, 1992). A related parameter, the effective reproductive number, R_{eff}, pertains when an epidemic spreads in a population in which some proportion are already immune (Scherer and McLean, 2002).

Studies interested in the pandemic influenza A(H1N1) genome show that the strain of 2009 is a triple-reassortant virus, consisting of human, swine, and avian influenza parts (Dawood et al., 2009). While it is still unknown which mechanisms led to the genomic mixture (see section 1.4.2), public health authorities were concerned about the possible dangers of the first influenza pandemic for over forty years. Estimates of the basic reproductive number ranged from 1.2 to 2.1 (see Table 6.1). A mid-range estimate of 1.5 implied

that the epidemic would grow until 1/3 of the population were infected or immune, then turn over and come to an end when around 60% of the population were immune. If all infections lead to illness prior to the acquisition of immunity this implied that a very large fraction of the population would become ill during the pandemic's early waves. Some alarming predictions were indeed made, particularly for the United Kingdom.

Reproductive number estimates	
1.22 - 1.58 (CI: 1.05 - 2.04)	(Fraser et al., 2009)
1.3 - 1.7	(Yang et al., 2009)
1.21 - 1.35	(Nishiura et al., 2010)
1.75 (CI: 1.64 - 1.88)	(Balcan et al., 2009)
1.37 (CI: 1.21 - 1.41)†	(Munayco et al., 2009)
1.6 (CI: 1.5- 1.8)†	(McBryde et al., 2009)
1.78 - 2.07 (CI: 1.67 - 2.22)	(de Silva et al., 2009)
1.44 (CI: 1.36 - 1.51)†	(Pourbohloul et al., 2009)
1.31 (CI: 1.25 - 1.38)	(Tuite et al., 2010)

Table 6.1: Reproductive numbers estimates for influenza A(H1N1) from several independent studies. A range of possible reproductive numbers is presented when multiple estimations are made. Estimates marked with † are the effective reproductive number R_{eff}, while unmarked ones are the basic reproductive number R_0. The confidence interval (CI) is given wherever possible.

In England, the Health Protection Agency (HPA) monitored case estimates over the whole period of the pandemic, and published its findings online in its 'Weekly Pandemic Flu Media Updates' and its 'Weekly National Influenza Reports' (Health Protection Agency, 2009d,c). The last pandemic flu media update was published on 14 January 2010, the official end of the pandemic in England (Health Protection Agency, 2009c,d). By this time HPA's estimates totalled 910,000 cases over the course of the pandemic in England (Health Protection Agency, 2009c). This amounts to less than 2% of the English population (Office for National Statistics, 2009), a much lower figure than expected even with a basic reproductive number from the lower end of the published estimates.

This large discrepancy between the predicted and observed epidemics could have several causes. The basic reproductive number might have been very much smaller than estimated, a large part of the population might have been immune before the pandemic arrived, or many susceptible individuals might have been infected and become immune

without being ill enough to register in the case estimates.

In what follows we use simple mathematical models to reproduce the HPA case estimates, and try to learn how each of these three factors contributed to the surprisingly small epidemic. We then ask at what point in time during 2009 it should have become apparent that the English epidemic would be so small.

6.2 Data and Theoretical Model

6.2.1 Phases of the Epidemic

We use the HPA case estimates which were made available to the public via the internet (Health Protection Agency, 2009c) as a data source for estimated numbers of symptomatic cases with influenza A(H1N1) in England. On 27 April 2009, the HPA announced the first laboratory confirmed infected case in the United Kingdom. It updated the number of new laboratory confirmed cases in England on a daily basis until 2 July 2009 when it switched to weekly updates of estimated numbers of new cases in the previous week. The last pandemic media update was published on 14 January 2010, when the influenza epidemic in England was officially declared to be over (Health Protection Agency, 2009d).

HPA's case estimates were an estimate of the number of people with symptomatic infections. The number of cases was estimated as a function of four factors: the number of GP visits with influenza-like-illness, the proportion of diagnosed cases that were confirmed as infected with A(H1N1), the proportion of symptomatic cases likely to contact the health services, and the likely impact of the National Pandemic Flu Service when it became available (Health Protection Agency, 2009b,a).

The data fall into three distinct phases:

Phase I: The first phase is the initial growth phase. It starts in week 19 with the first infected cases in the United Kingdom, and ends in week 30. From week 19 to week 27, the HPA released case reports based on laboratory confirmed cases. It changed its method to estimate cases using surveillance measurements in week 28 (Health Protection Agency, 2009c), explaining the jump in new cases between week 27 and 28. However, we expect that the laboratory confirmed case numbers released before week 28 are under-estimates

of actual case numbers, because of limited laboratory facilities and the time delay to test for influenza A(H1N1).

Phase II: From week 30 to week 36 the number of cases declined. This is the period of school summer holidays in England. State schools were closed from 22 July to 3 September 2009 (Directgov, 2009), exactly corresponding to phase II and marking the borders of the three phases. It is very likely that school closures during the summer break caused the decline in transmission events as children and young adults have been identified as the main age groups that were infected (see Figure 6.1 and (Miller et al., 2010; Girard et al., 2010)). Most universities were also shut during this phase, reducing mixing amongst young adults.

Phase III: The last phase is the second epidemic wave after the school holiday. Schools reopened in week 36, and the number of new infections increased until week 44 after which numbers of new cases decreased again. The third phase ends in week 1 of 2010 with no more published HPA updates on estimated A(H1N1) cases.

6.2.2 Theoretical Model

We developed a simple version of the SIR mathematical model that allows for two routes to immunity; one with disease and one without. As we are interested in short-term epidemic behaviour it is a model for a closed population without any births. There are four types of individual: susceptibles, S, catch the infection at a rate that is proportional to the fraction of the population that is infectious. Upon infection they become infectious, either with symptoms, I_s, or without, I_a. The proportion that develop symptoms on infection is p. After a period of infectiousness they recover to the immune class, R. The rate of transmission of infection from those infected to the susceptibles is reduced by some fixed factor r_h during phase II, the school holidays. At the start of the epidemic in England most people are susceptible so in class S, but some may already be immune so are already in class R. This model has five parameters: the proportion initially immune, $R(t = 0)/(R(t = 0) + S(t = 0))$, the transmission rate, β, the reduction in transmission during the holidays, r_h , the duration of infectiousness, $1/\gamma$, and the proportion of the

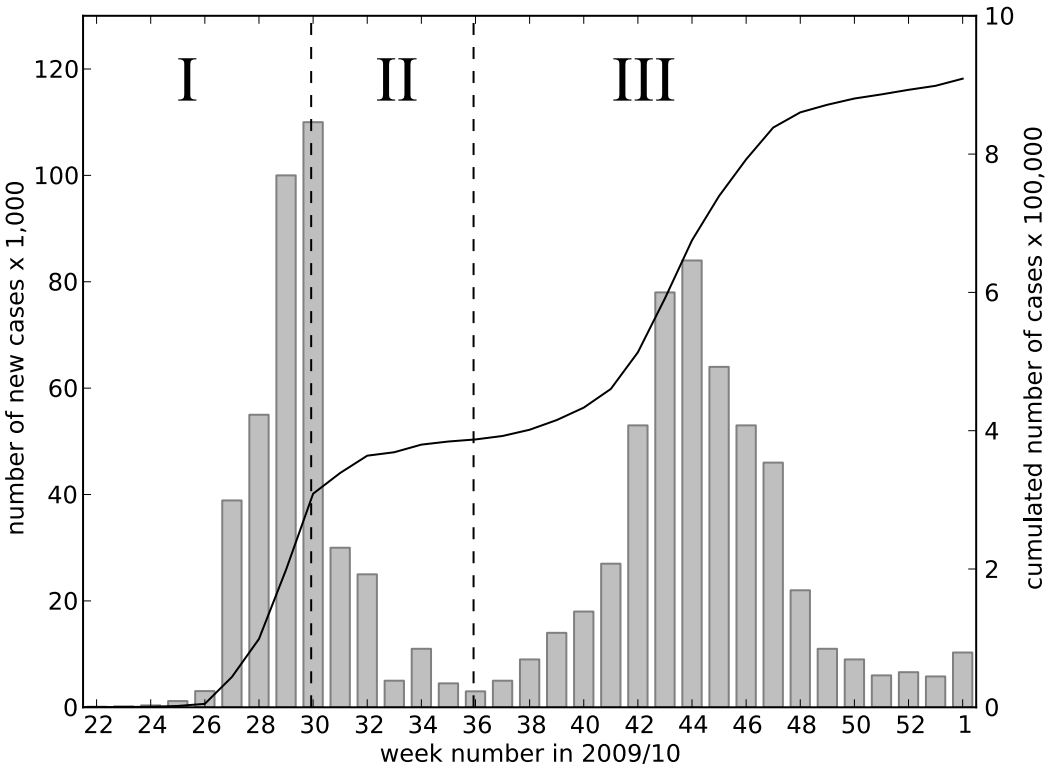

Figure 6.1: HPA case data and cumulative number of cases. The bars show the number of estimated new cases per week. The cumulative number of cases per week is presented as the black line on a separate axis. The plot is separated into three phases, I - III, divided by dashed lines in week 30 and 36 between which state schools in England were closed (Directgov, 2009). Phase I is the initial growth phase before 22 July 2009, phase II is the school holiday phase until 3 September 2009, and at this date phase III with another growth and decline begins.

population who become symptomatic if infected, p:

$$\begin{aligned}
\frac{dS}{dt} &= -r_h(t)\,\beta\,\left(I_s + I_a\right)\,\frac{S}{N} \\
\frac{dI_s}{dt} &= p\,r_h(t)\,\beta\,\left(I_s + I_a\right)\,\frac{S}{N} - \gamma\,I_s \\
\frac{dI_a}{dt} &= (1-p)\,r_h(t)\,\beta\,\left(I_s + I_a\right)\,\frac{S}{N} - \gamma\,I_a \\
\frac{dR_s}{dt} &= \gamma\,I_s \\
\frac{dR_a}{dt} &= \gamma\,I_a \\
N &= S + I_s + I_a + R_s + R_a
\end{aligned} \tag{6.1}$$

with $R(t) = R_s(t) + R_a(t)$. The reduction in transmission is a time dependent function:

$$r_h(t) = \begin{cases} r_h & \text{if within school holiday} \\ 1 & \text{otherwise} \end{cases} \tag{6.2}$$

Furthermore, the basic reproductive number can be defined as $R_0 = \beta/\gamma$, and the effective reproductive number is just $R_{\text{eff}} = R_0\,S/N$. In the special case $p = r_h = 1$ our model reduces to the standard, textbook epidemiological model called the "SIR model" (Anderson and May, 1992; Keeling and Rohani, 2007).

6.2.3 Serological Data

We used serological data based on samples drawn in England between July and September 2009. These were published in 2010 (Miller et al., 2010). They yield estimates of the proportion of the population immune before the pandemic arrived and the proportion immune at the end of the first wave of the pandemic. The simple model that we use to understand these data is age-independent, and we therefore need to calculate proportions of the total population immune without age-stratification. To do that from age-stratified data we used demographic data on the current age and regional distribution of the English population from the Office of National Statistics (Office for National Statistics, 2009).

6.2.4 Qualitative Analysis

Our Method proceeds in four steps:

First we calculate how many people were immune to A(H1N1) before the pandemic and in September 2009. The difference between these two numbers yields the total number of people who were infected during the English epidemic up to the time of sampling. This is a conservative estimate as people might have been infected between their sampling, which started in mid-July, and September 2009. Our method therefore underestimates the actual number of infected cases.

Second we compare the number of seroconverters with the cumulative number of estimated symptomatic cases by the beginning of September to get an estimate of the proportion of all infections that manifest themselves in disease (parameter p in the mathematical model). This parameter is calculated as a central value with upper and lower confidence bounds.

We then use these two pieces of information (proportion immune before the pandemic arrived and proportion of all infections manifesting as reported cases) and fit our model to the case report data. In fitting the model to the data we infer values for the three free model parameters: the basic reproductive number (R_0), the infectious period ($1/\gamma$) and the reduction in transmission during the summer holidays (r_h).

In a fourth step we use our model to ask if it could have been obvious from case-reports alone that the English epidemic would be so small, and if so, when. If we assume $S/N \approx$ const., the number of symptomatic infected can be expressed through an exponential growth:

$$I_s(t) \approx I_0 \, e^{(p\, R_{\text{eff}} - \gamma)t} \tag{6.3}$$

The HPA data on symptomatic cases is based on weekly updates. This weekly sampling rate might be too broad to estimate the correct number of symptomatic cases at any time point, especially if the generation time is smaller than a week. But it allows us to evaluate the accumulated sum of symptomatic cases (new infections plus recovered ones), which grows with:

$$I_s(t) + R_s(t) \approx \frac{\beta \, I_0}{R_{\text{eff}} - \gamma} \, e^{(p\, R_{\text{eff}} - \gamma)t} + \text{const.} \tag{6.4}$$

where the last term expresses the normalisation constant. As before, the growth is described with an exponential function. We can determine the effect of the depletion of susceptible hosts, or in other words the difference in S/N between the two waves, by taking the natural logarithm of the accumulated sum of HPA's estimates. The slope of the logarithmic growth periods is ($p\,R_{\text{eff}}$ - γ), so differences in the slope of the two epidemic waves reflect the drop in R_{eff} caused by a depletion of susceptible hosts from one wave to the next.

6.3 Results

The English epidemic unfolded in three phases (Figure 6.1). The first phase of rapid growth was followed by a second phase when the number of cases fell as schools closed for the summer holiday. The third phase saw the start of the second wave as schools reopened in the autumn. After a short period of exponential growth, the second wave peaked in mid-November and then tailed off. This was surprisingly early, given the predicted attack rates.

These data raise four questions:

1. How transmissible was this strain of influenza amongst the population of England (i.e. how big was R_0)?
2. By how much did transmission decline during the summer holidays (i.e. what was r_h)?
3. For each case detected how many seroconverters were there (i.e. what was p)?
4. How soon could the failure to develop into a large epidemic have been recognised?

These questions can be addressed by combining the case report data with serological data (Figure 6.2). Those serological data show (red bars) that a substantial number of people were immune before the epidemic arrived and that this was particularly true for those over the age of 45. Correspondingly there were many seroconversions by early September (blue bars) in the young and very few in those over 45. Finally by early September a large part of the English population was already immune (green bars), regardless of age.

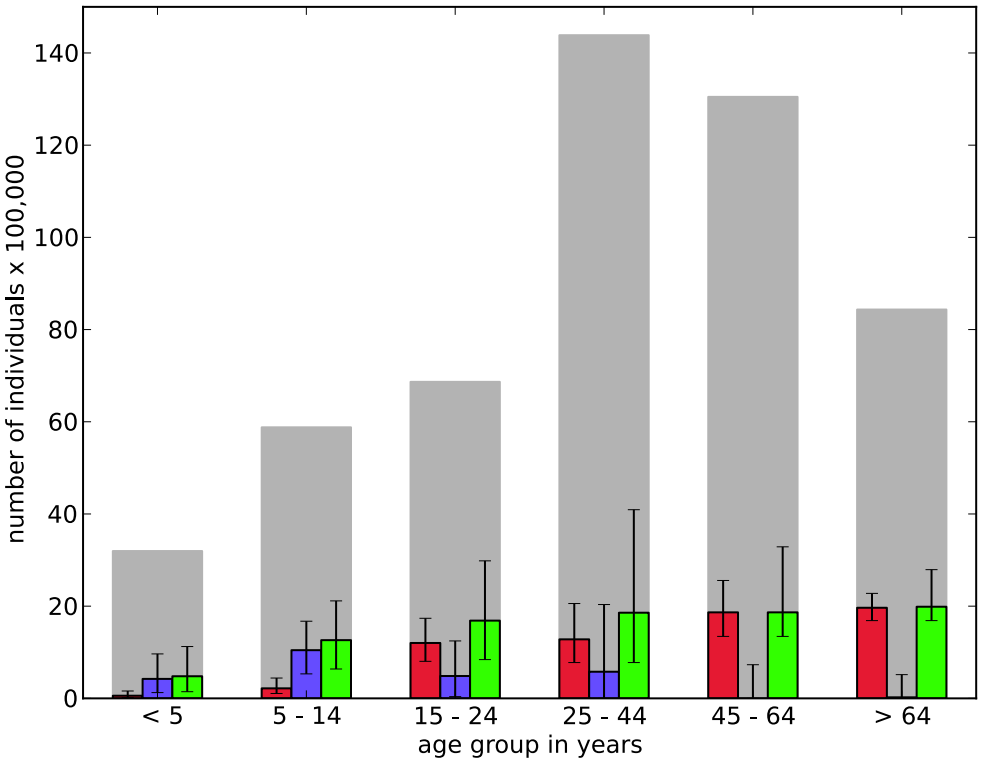

Figure 6.2: Comparison of serological data by age group. The number of individuals in each age group is shown with large, grey bars. Red bars show the number of individuals immune before the pandemic. Blue bars represent numbers who seroconverted between April and September 2009. Green bars show the number immune by early September 2009. All numbers are calculated using serological data and demographic data from England (Office for National Statistics, 2009; Miller et al., 2010).

Adding up the number of serconverters across ages gives an estimate of 2,552,000 infections in England by the beginning of September 2009 (range 696,000 - 7,186,000 based on 95% confidence intervals for the published serological studies). Yet, by that time, only 385,368 symptomatic cases had occurred according to HPA case estimates. We can take the ratio of symptomatic cases to all infections as an estimate of the proportion of cases that are symptomatic, parameter p in our model. The serological data imply that $p = 0.15$ (range 0.05 - 0.55). This is more easily interpreted as the inverse of p; the number of seroconverters for each recorded case. This has a mean of 6.62 with range 1.81- 18.62. In what follows we use four different values of $1/p$: $1/p = 1$, every seroconverter is symptomatic; $1/p = 1.81$, the lower bound estimate from the serological data; $1/p = 6.62$, the mean from the serological data; and $1/p = 18.62$, the upper bound from the serological data.

The serological data also yield an estimate for the proportion of the population who were immune to the pandemic influenza strain before it arrived. The estimate is 12.7% and is used to define initial conditions for the model in equation (6.1).

Having made estimates for the initial conditions and the ratio of seroconverters to cases, we proceed to fit the model to the case data. This amounts to finding a set of parameter values that are consistent both with the case reports and with the serological data. Figure 6.3 presents the best fits of our mathematical model to the case reports for four different values of $1/p$ ranging from the situation where all seroconverters present as cases ($1/p = 1$) to the upper limit of estimated values in which for each reported case there are more than 18 seroconverters ($1/p = 18.62$) (Miller et al., 2010). All fits assume that 12.7% of the population are immune before the pandemic arrives, reflecting pre-existing immunity in the population. Our fitting procedure is to fix p to one of our four values, fix the percentage immune before the pandemic to 12.7% and then estimate the three parameters R_0, $1/\gamma$ and r_h by making a least square fit of the solution of the model (equations (6.1) and (6.2)) to the HPA's case report data. Table 6.4 reports best fit parameter estimates for each value of $1/p$. Overall, the best fit is for $1/p = 18.62$. The parametric fit yields $R_0 = 1.59$, very much in agreement with previous estimates (see Table 6.1). In addition, an infectious period of 2.42 days lies well within previous independent estimates for influenza A(H1N1) (see Table 6.2 and Table 6.3). The best-estimate reduction factor, $r_h = 0.65$, reveals that

transmission is 35% lower over the period of school holidays. While there is no empirical data to validate this result, it seems plausible as the number of infected cases was highest in the age group of 5 to 14 years old (see Figure 6.2). This reconciliation of the case report data and the serological data concludes that the 2009 English epidemic of symptomatic cases was small for the following reasons:

1. a large fraction (12.7%) of the population were immune before the pandemic arrived in England,

2. for every reported case there were many (approx. 18) seroconversions that were not reported as cases so the susceptible pool was very quickly diminished,

3. initial estimates for R_0 of around 1.5 were roughly right.

Infectious period	
2.5 (CI: 1.1 - 4.0)	(Balcan et al., 2009)
3.38 (CI: 2.06 - 4.69)	(Tuite et al., 2010)

Table 6.2: Infectious period estimates for influenza A(H1N1) from two independent studies with confidence interval (CI). The infectious period is presented in days.

Generation time	
6 - 8	(Dawood et al., 2009)
1.91 (CI: 1.30 - 2.71)	(Fraser et al., 2009)
2.6 (CI: 2.2 - 3.5)	(Cauchemez et al., 2009)
2.6 - 3.2	(Yang et al., 2009)
2.7 (CI: 2.0 - 3.5)	(Lessler et al., 2009)
2.9	(McBryde et al., 2009)
1.9 - 2.6 (CI: 1.3 - 3.0)	(de Silva et al., 2009)
6 (CI: 5 - 7)	(Pourbohloul et al., 2009)

Table 6.3: Generation time estimates for influenza A(H1N1) from several independent studies. The generation time is presented in days. A range of possible generation times is shown when multiple estimations were made. The confidence interval (CI) is given wherever possible. A median latent period of 1.4 days (CI: 1.0 - 1.8) (Lessler et al., 2009) and 2.62 days (CI: 2.28 - 3.12) (Tuite et al., 2010) has been reported. The infectious period can be estimated using the generation time as latent period and half the infectious period are on average the generation time (Anderson and May, 1992).

This raises the following question: how early would it have been possible to detect that there were so many seroconverters for every reported case? Was the epidemic already

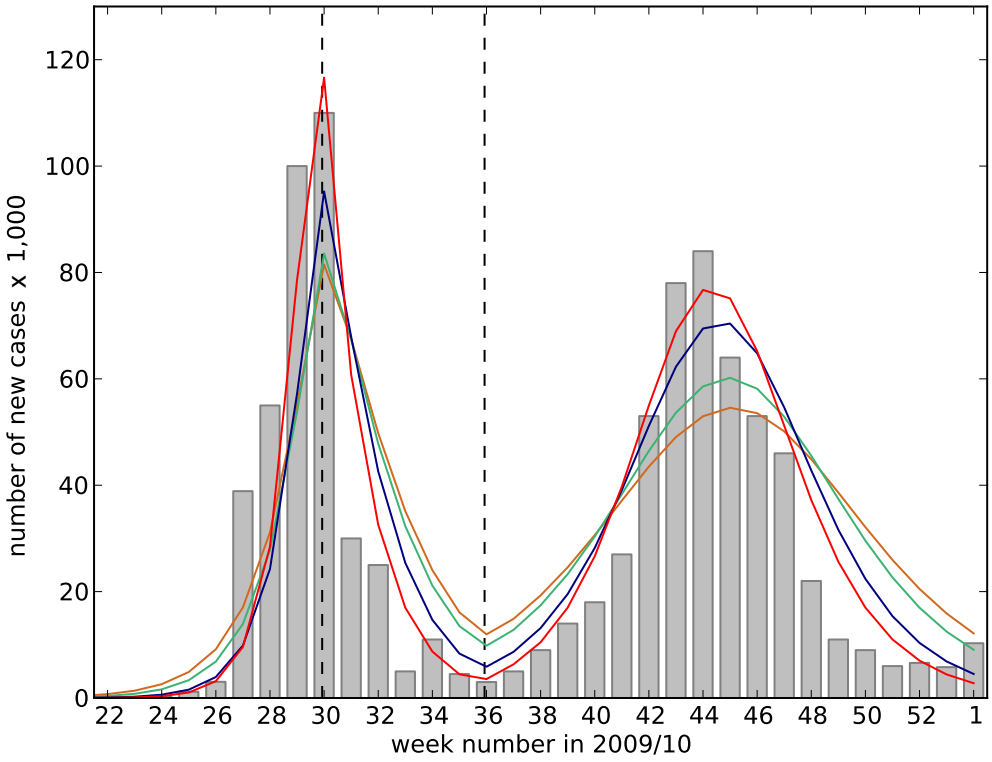

Figure 6.3: Comparison of HPA case estimates with theoretical predictions of new cases. The HPA case estimates per week are represented in a bar plot. The model in equation (6.1) was fitted to these data with state variable I_s representing reported cases. The initial conditions were set at the observed value of 12.7% of the population immune before week 22. Four different values of parameter p were assumed corresponding to four different fits: $1/p = 1$ - brown line, $1/p = 1.81$ - green line, $1/p = 6.62$ - blue line and $1/p = 18.62$ - red line. The figure illustrates that $1/p = 18.62$ gives the best agreement between model and data with the smallest least square error, and allows our model to reproduce the case estimates. The fitting procedure yields estimates for the model's three parameters, shown in Table 6.4. Only when $1/p = 18.62$ do these parameter estimates agree with previously published values, compare Table 6.4 with Tables 6.1 - 6.3.

Estimated parameters			
$1/p$	R_0	$1/\gamma$	r_h
1.00	1.16	0.18	0.98
1.81	1.18	0.28	0.97
6.62	1.27	0.82	0.89
18.62	1.59	2.42	0.65

Table 6.4: Estimated parameters using a least square fit. All parameters were estimated assuming an immune fraction of 12.7% at the beginning of the pandemic. A least square fit was used to minimise the error between our model and HPA case estimates.

starting to slow in late July, before schools broke up? Figure 6.4 A answers this question with a clear "no". If the schools had not broken up, using our best-fit parameters, we believe that the epidemic would have continued to grow rapidly for several more weeks. It is therefore not surprising that a logarithmic plot of cumulated cases during the weeks before schools closed (Figure 6.4 B) shows that the epidemic was clearly still in its exponential growth phase with no indication of a falling off in the growth rate. However, when schools reconvened in September, the second pandemic wave clearly grew at a much slower rate than before the holiday. Figure 6.4 C plots cumulated case numbers in a logarithmic plot that compares the rising phases of the first and second waves. The second wave is clearly growing very much more slowly than the first wave, which is consistent with the large fraction of the population that had already seroconverted by the end of the first wave. In short, the very slow growth of the second wave of the pandemic was a clear indicator that many people had already seroconverted during the first wave.

6.4 Discussion

The HPA published their case estimates on influenza A(H1N1) infections over the period of the pandemic in England. Overall, the HPA estimated that approximately $910,000$ individuals were ill with the pandemic influenza virus during 2009, amounting to less than 2% of the English population (Health Protection Agency, 2009c). This surprisingly small number of symptomatic cases only makes sense in the light of the HPA's serological survey which found that around 17.7% of the English population were already immune by the

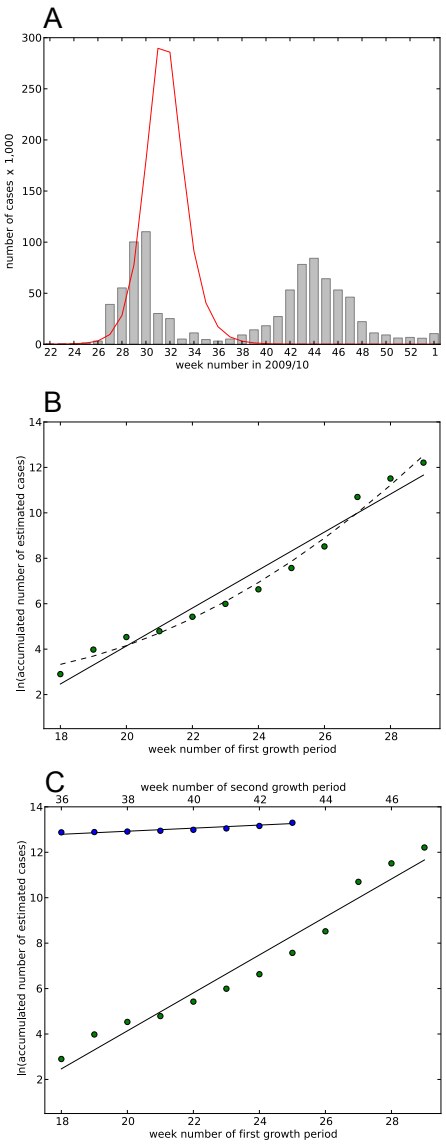

Figure 6.4: When did growth data reveal that the English epidemic would be small? **A** shows that it was unlikely that the first wave would indicate that susceptible hosts were becoming exhausted. If there had been no school break we expect that the epidemic would have gone on growing for several weeks after week 30 when they actually shut. The red line shows predicted dynamics of the English epidemic if the schools had not closed (i.e. if $r_h = 1$ and other parameters are the best fits for the model of equation (6.1)). **B** This is confirmed in regression analyses of the natural logarithm of the cumulative number of cases. Data are filled circles, linear fit is the solid line and quadratic fit is the dashed line. Linear fit: ln(cumulative cases) $= -13.5 + 0.84 * \text{week}$. Quadratic fit: ln(cumulative cases) $= -17.16 + 0.22 * \text{week} + 0.05 * (\text{week})^2$. The quadratic term is significant ($p = 0.002$) but small and positive, signifying that, if anything, the epidemic was accelerating just before the schools broke up for summer. **C** compares growth rates of the first wave (green dots) and second wave (blue dots). An analysis of covariance reveals that the two slopes are significantly different ($p < 0.001$). Linear fit for second wave: ln(cumulative cases) $= 10.58 + 0.06 * \text{week}$.

end of the first wave of the pandemic in late September 2009. This observation, combined with knowledge of the level of immunity prior to the pandemics' arrival and the dates of the English summer holidays, combine together to tell a coherent story about the 2009 influenza pandemic in England. The pandemic had a comparatively small impact, because for every case that was reported there were around 18 seroconversions with no symptoms at all.

In retrospect, it was clear from the very slow growth of the second epidemic wave that something was different as the known cases up until September were too few to explain such slow growth. Depletion of susceptibles through asymptomatic cases would always have been a strong contender as a mechanism for this slow observed growth.

The HPA's serology study was organised, collected, analysed, and published in a very short space of time. No other country has, to our knowledge, published such extensive data (Chen, 2009; Prachayangprecha et al., 2010; Delangue et al., 2011). In principle, the same pattern that was picked up in these data in September would have been seen in samples drawn even earlier. This modelling study is a ringing endorsement of the explanatory power of serological surveys in a pandemic setting. However, it is not obvious that a serological survey could have been organised and executed any earlier in a pandemic than this one was.

The distinction between infection and disease is an early lesson in the teaching of infectious disease biology (Mims et al., 2000). Predictions of the likely impact of the 2009 influenza pandemic were made using models that did not distinguish between cases of infection and cases of disease. A very simple model that allows asymptomatic seroconversion is presented here and yields a coherent reconciliation of the estimated numbers of disease versus the estimated number of infections through the course of the 2009 pandemic in England. Once it is clear how many asymptomatic cases there were, the epidemiology of the 2009 influenza pandemic in England makes sense.

CHAPTER 7

CONCLUSION

7.1 Introduction

The overall focus of this thesis is on species-crossing, emerging infectious diseases. Zoonotic EIDs are one of today's great dangers to human health (Cohen, 2000), and, not surprisingly, they are one of the important areas in infectious disease research. However, while much research has been focused on its extent, the biological mechanisms, and its impact on public health (Taylor et al., 2001; Woolhouse, 2002; Antia et al., 2003; André and Day, 2005; Jones et al., 2008; Arinaminpathy and McLean, 2009), zoonotic EIDs are far from being completely understood. This thesis presents new advances by looking at species-crossing EIDs from different angles. While it might not be possible to ever fully model and explain every aspect of EIDs, this thesis contains the complete and exact explanation of some aspects, and first insights into others. I look at different aspects of zoonotic EIDs through four main questions, which I previously outlined in section 2.8:

(i) How can zoonotic EIDs come 'out of nowhere'?

(ii) What environmental conditions can facilitate inter-species jumps of an infectious disease?

(iii) Does a change in the species mixture facilitate zoonotic EIDs?

(iv) What was the impact of an exemplar novel infectious disease in a human population?

Note that 'environment' is very broadly defined and used in this thesis (see section 1.3 for a comprehensive introduction of this term). I answer each of these questions in at least one of my research chapters 3 - 6. While each research chapter focuses on a particular problem, the aim of this conclusion is to set the knowledge gained into context, link results, and reveal the 'bigger picture'.

7.2 How Can Zoonotic EIDs Come 'Out of Nowhere'?

The emergence of novel, infectious diseases is a natural process. Novel strains arise by recombination, mutation, and reassortment, creating vast number of variants that can chal-

lenge the host in very different ways (Scholtissek, 1995; Duffy et al., 2008). This evolutionary 'trial and error' process has the side effect that most new variants perform worse than the original in the host (Gago et al., 2009; Holmes, 2009b). But it opens the opportunity for generating a small number of variants that might perform better. This process is the reason why viruses like HIV have the potential to develop drug resistance. Paired with an environment that allows for cross-species transmission, such genomic changes, either happening during the emergence process or already existing ('off the shelf'), are the cause of novel infectious diseases arising in new host populations. The interaction of this process with environmental conditions is what I am most interested in.

The advent and emergence of a novel pathogen is therefore the result of two fundamental drivers - environment and evolution. The environment determines the possible route of transmission, and evolution determines the potential to use this route. It is the complex interplay between these two concepts which decides the fate of a novel pathogen. The quality and quantity of the between-species contact, and therefore the environmental conditions, are often the deciding point for a successful or failed cross-species transmission. However, the pathogen may require evolutionary adaptations to sustain transmission in the novel host species, which can happen as pre-jump adaptations in the old, donor host species, or as post-jump adaptations in the novel, recipient host species. I previously introduced these as the two main factors that influence a zoonotic emergence (see section 1.2.2).

These factors are of constant interest to infectious disease researchers: several studies have made contributions towards a better understanding of the mechanisms allowing zoonotic infectious diseases to emerge, concentrating on necessary evolutionary adaptations (Antia et al., 2003), epidemiological properties of the pathogen (André and Day, 2005), the implications for public health (Arinaminpathy and McLean, 2009), or environmental aspects within and between the host species (McMichael, 2001, 2004; Weiss and McMichael, 2004; Yates et al., 2006; Reluga et al., 2007; Alexander and Day, 2010). All these studies have in common that they assume an emergence 'out of nowhere'. That is as proposed in question (i), a novel pathogen arises in the human population without any or only little prior warning in the form of outbreak clusters with a less than fully-adapted pathogen.

The probability of a successful emergence is of crucial interest for public health as a measure of the actual danger a novel pathogen might possess. In chapter 3, I present the first algebraic solution to the probability of such an emergence per introduction. Previous research relied on numerical methods to approximate this critical information (Antia et al., 2003; André and Day, 2005). My approach uses a branching process formulation to include the stochastic nature of emergence events, and only depends on the pathogen's biology. Its usefulness becomes directly apparent as it allows for calculating the average waiting time distribution and its standard deviation in terms of introductions, which reveal that the actual time point of an emergence becomes effectively unpredictable for low probabilities of emergences. Conversely, it reveals the average value for the probability of emergence if the number of introductions leading to a successful emergence is known. It is therefore similar to Arinaminpathy and McLean (2009)'s estimate of the upper bound on the probability of emergence per introduction, but has the benefit that it is a more accurate, algebraic solution with the standard deviation as an estimate of confidence for the derived value.

These introductions are thought to be of small size, because larger outbreak sizes with thousands of cases would warn public health authorities and not allow for unexpected emergences as seen with SARS (World Health Organization, 2006). I investigate these cluster sizes in chapter 3, and derive an algebraic formulation of the outbreak size distribution for a given evolutionary pathway of adaptation. I show that, even for pathogens near the necessary threshold for sustained transmission, the probability for large outbreaks is small. This is in line with results previously gained by simulations (Arinaminpathy and McLean, 2009). Additionally, outbreak sizes tend to be smaller for evolutionary routes of adaptation which lead to smaller probabilities of emergence per introduction. Hence, these routes have the largest uncertainty in the number of failed introductions before a successful emergence, and each failed introduction has on average the smallest outbreak size which could offer warning. Pathogens following these routes are of greatest danger for public health.

But my method has further possible applications. It can be incorporated into a maximum-likelihood framework to examine empirical outbreak data. A similar approach has been

used to estimate the effective reproductive number of measles in the United Kingdom (Jansen et al., 2003). The benefit of my approach is that it works with several intermediate, not fully-adapted strains. It should therefore be able to estimate the actual number and transmission potential of all intermediate steps for a sufficiently large database. To my knowledge, no database of such quality exists yet, but the trend towards better worldwide reporting might change this in the future.

7.3 What Environmental Conditions Can Facilitate Inter-Species Jumps of an Infectious Disease?

I address one part of question (ii) in chapter 3. The spatial homogeneity coefficient shows in a simple but accurate way if communities are sufficiently connected to support an emergence and sustained transmission of an infectious disease. This is an important result in two respects: first, previous research only evaluated host heterogeneity on an individual level (Yates et al., 2006; Alexander and Day, 2010). I extend the host heterogeneity to a community such as villages and cities. Second, the spatial homogeneity coefficient works independently of the disease model. It can be used to predict the impact of spatial heterogeneity on zoonotic EIDs, but also to describe the transmission-behaviour of a well established infectious disease. In both cases, the spatial homogeneity coefficient and its underlying concept increase our understanding of infectious disease spread between communities, and allows us to determine connectivity ranges for which infectious disease transmission is affected. As an example, I can apply this new knowledge to determine success rates of preventive measures targeting disease transmission such as a quarantine.

However, these are only the environmental conditions within a host species. Chapter 4 examines environmental conditions between the host species, therefore covering another part of question (ii). The questions of a specialist or generalist pathogen and of between-species connections allowing pathogen jumps share obviously many aspects. A generalist pathogen must, by definition, be able to infect more than one host species. Nevertheless, it is fundamentally different to 'out of nowhere' emergences of zoonotic pathogens. As I have shown in chapter 4, there are three different classes of outcome. First, the case where

specialists are the only pathogens in both species. Second, a region of co-existence where the specialist and the generalist each are pre-dominant in one species, with the specialist being only endemic in its dominated host species. And third, an area where a generalist strain will out-compete a specialist one in both species and a specialist might co-exist in one host species. Chapter 4 illustrates that threshold connectivities exist, which separate these regions of parameter space. They only depend on the epidemiological parameters of the pathogens, that is the reproductive numbers, and the difference in the host species sizes. This result has a similarity to chapter 3's observation that it is mainly the pathogen's biology which influences the outcome of an introduction.

Does this mean that a disease introduction by a generalist will never come 'out of nowhere'? No, but it is unlikely. Assume two host species connected below the critical threshold, and a specialised infectious disease being exclusively endemic in one species and no disease in the other. If I assume that possible generalists are generated from time to time, which is supported by the large mutation rate of some pathogens (Gago et al., 2009), the connectivity must be too low to support generalist co-existence. Hence, the connectivity must drastically enhance in a short time to see a generalist strain arise in both populations without previous, sustained circulation in any of the host species. Nonetheless, this is not to say that cross-species jumps of pathogens are impossible below the threshold connectivity. A generalist strain might very well arise by mutation and cause a limited outbreak. With some chance, it might also get transmitted into the other host species and cause further infections there. This allows the pathogen to mutate even further, and adapt to its novel host species. In fact, I describe such a scenario in chapter 3, starting with the introduction into the novel host species. In this case, the novel strain would arise 'out of nowhere', but it would not be a generalist invasion. Experimental studies with bacteria and phage show such a transition where a generalist strain is only an intermediate step through the fitness valley separating both host species (Benmayor et al., 2009).

7.4 Does a Change in the Species Mixture Facilitate Zoonotic EIDs?

Having considered the impact of connectivity rates between species in chapter 4, I go on to consider the impact of species' growth rates in chapter 5. It is generally assumed that *r*-selected species, also often called 'weedy' species, have higher prevalence of infection (May et al., 1988; Keesing et al., 2010). I use models that combine logistic growth and infection dynamics to make a formal study of the impact of host demographic strategy on infection prevalence. Chapter 5 yields some insights pertinent to changes in biodiversity. In general, a change in biodiversity should have an overall neutral effect on the probability of zoonotic EIDs if the lost or replaced species, and their respective pathogens, are selected at random with a fixed connectivity to the remaining species. However, it is believed that species loss, and therefore a change in biodiversity, is not a random process (Keesing et al., 2010). This becomes plausible if the reason for a change in biodiversity is taken into account. A stable species mixture within an ecological niche is unlikely to change if surrounding factors remain constant. It needs a changing factor to drive the stable equilibrium of species into an unstable region, facilitating the increase or decrease of species sizes. For example, a change in the available resources might increase competition between species, resulting in the extinction of some. Therefore, the type of factor driving the change in species mixtures also determines which species are better adapted and more likely to survive the change.

Interestingly, the literature suggests that fast growing species with many weakly competitive progeny ('weedy' species) seem to withstand ecological changes best (Stearns, 1992; Keesing et al., 2006, 2010). Chapter 5 takes a look at disease prevalences as a function of the host species' growth rate. Overall, it shows that the underlying type of disease transmission determines if the prevalence is a direct or indirect function of the 'weediness'. Prevalence always increases with larger host growth rates for density-dependent disease transmission. For frequency-dependent transmission, prevalence only depends on the epidemiological parameters of the pathogen. But the growth rate needs to be above a certain threshold, or the infectious disease will drive the population to extinction. This threshold depends on the host species' growth rate, revealing the indirect dependence on

'weediness'.

How does this knowledge help us answer question (iii)? First, a larger number of infected hosts generate a higher possibility to generate necessary, pre-jump mutations for a successful emergence in a recipient species. This is shown in chapter 3 where higher intermediate reproductive numbers yield larger outbreak sizes, which offer more possibilities of adaptive mutations and therefore a larger risk of emergence. In addition, the spatial homogeneity coefficient depends on the prevalence. That is, the likelihood to transmit a disease between two communities, of the same species or not, increases with the prevalence. Therefore, I can causally link the 'weediness' of species to the danger of originating zoonotic EIDs. Assuming that 'weedy' species are more successful in withstanding environmental changes, the answer to question (iii) is that a change or loss in biodiversity overall increases the risk of zoonotic EIDs. It is therefore in line with Cohen (2000)'s notion that 'change' appears to facilitate the risk of EIDs.

7.5 What was the Impact of an Exemplar Novel Infectious Disease in a Human Population?

My results seem to draw an alarming picture for public health. Biodiversity appears to be declining (Epstein, 2001; Keesing et al., 2006), and the rise of deadly pathogens like SARS might become more common. But this is only the worst case scenario. In fact, a large fraction of all humans' pathogens are of zoonotic origin (Woolhouse, 2002; Woolhouse and Gowtage-Sequeria, 2005). Yet, even if all zoonotic EIDs were as deadly as the 'Spanish flu' from 1918, humans would still exist. Question (iv) addresses this by asking what actually happened when a novel, infectious disease arose in the population. In principle, this question has been addressed by every epidemiological study focusing on infectious disease spread: a novel pathogen arises in the population which has no or only little pre-existing immunity and causes an epidemic. Some literature focuses on the effect of host heterogeneity (Lloyd-Smith et al., 2005; Colizza and Vespignani, 2008), the influence of host immunity (Iwasa et al., 2004; Bansal et al., 2007), or the theoretical development of models (Newman, 2002; Keeling and Rohani, 2002). Instead of adding another theoretical

model, I use a different approach by looking at empirical data of the latest zoonotic EID epidemic - the 'Swine flu' influenza A(H1N1) pandemic of 2009.

The 'Swine flu' pandemic was novel in a way that researchers were trying to assess the danger in real-time to advise public health authorities and governments on the severity and impact of the first influenza pandemic for over 40 years (Fraser et al., 2009; Donaldson et al., 2009). Most interestingly, the predicted impact and the observed number of cases did not seem to match up. To understand this discrepancy, I compare published case numbers in England with a simple model to explain the observed behaviour in chapter 6. My model incorporates not only the number of infected, but also the number of ill people. I use data from a serological survey to estimate the pre-pandemic immunity within the English population, which reveals that almost one in six people were immune to the novel strain through cross-immunity previously induced by antigenically similar influenza strains. My model reveals that for every case that was reported there were approximately 18 infected hosts without symptoms. Furthermore, the growth rate of the second wave already reveals that the number of asymptomatic, not-reported cases must have been much higher than expected. The epidemiology of the 2009 influenza epidemic in England begins to make sense if my results are incorporated into the picture.

To re-formulate these results in plain words: being infected is something different than being ill. It is an important distinction which needs to be made. Especially theoretical modellers often oversee the clinical aspects of infectious diseases. Therefore, the results of chapter 6 illustrate two important insights. First, the impact of zoonotic EIDs cannot just be measured by epidemiological parameters or case numbers. These are valuable information, but it needs more to fully assess the inherent danger of a zoonotic EID. Studies of the pathogen's genome, its cellular processes, and its clinical symptoms are necessary to understand the full impact a zoonotic EID would have. The second result of chapter 6 is tied to the first one. Previously, I argued that there are three fundamental dimensions to assess a model: complexity, realism, and relevance (see section 2.1). Chapter 6 shows that the models used to predict the impact of the pandemic at its start lacked some relevance. By showing that the initial estimates about the 'Swine flu' epidemic were overall correct, I point out that the theoretical models delivered precise predictions, but might not have

been an answer to what was asked for. Indeed, the motivation for chapter 6 is the apparent difference in public perception of what researchers apparently predicted, and what the actual outcome was. I argue that researchers were overall correct in their estimates, but need to be precise in communicating the results they gain. If the public asks how bad will it be, delivering the probable number of infected might not be the best answer. More relevant models would have been ones that differentiate between infected, for example include the severity of symptoms and case fatality rates.

7.6 Summary

However, it is an important fact about theoretical models that they only try to describe what can be observed. They are based on previous experiences, and therefore no model should be assumed to be irrevocable and the inviolable truth. One needs to be careful not to over-interpret the results. Keeping models focused on the question to answer (complexity) and incorporating relevant information about the infectious disease (realism) will lead to the best insights models can provide (relevance). I aimed to design all the models used in this thesis according to these principles. As every theory, they might prove wrong. But as new, deadly zoonotic EIDs like Ebola and SARS make an appearance into the human population, the role of mathematical models in disease control and decision making is becoming more important. For example, linking the probability of an emergence to the number of already observed cases is most helpful in public health decision making, and insights about the relevance of a complex biodiversity on EID risks might have implications for climate change policy making. Finally, modelling biology allows for great insights but also offers many pitfalls. Or, as William Cameron once wrote (Cameron, 1963):

> "Not everything that can be counted counts, and not everything that counts can be counted."

REFERENCES

Abdel-Ghafar AN, Chotpitayasunondh T, Gao Z et al. (2008). Update on avian influenza A (H5N1) virus infection in humans. *N Engl J Med* 358(3): 261–73.

Alberts B, Bray D, Hopkin K et al. (2009). *Essential cell biology*. Garland Science, London.

Alexander HK and Day T (2010). Risk factors for the evolutionary emergence of pathogens. *J R Soc Interface* 7: 1455–1474.

Anderson RM and May RM (1992). *Infectious diseases of humans: dynamics and control*. Oxford University Press, Oxford.

Anderson RM, May RM and McLean AR (1988). Possible demographic consequences of AIDS in developing countries. *Nature* 332(6161): 228–234.

André JB and Day T (2005). The effect of disease life history on the evolutionary emergence of novel pathogens. *P Roy Soc B Biol Sci* 272(1575): 1949.

Anishchenko M, Bowen Ra, Paessler S et al. (2006). Venezuelan encephalitis emergence mediated by a phylogenetically predicted viral mutation. *P Natl Acad Sci USA* 103(13): 4994–9.

Antia R, Regoes RR, Koella JC and Bergstrom CT (2003). The role of evolution in the emergence of infectious diseases. *Nature* 426(6967): 658–61.

Antolin MF (2008). Unpacking β : within-host dynamics and the evolutionary ecology of pathogen transmission. *Annu Rev Ecol Evol S* 39(1): 415–437.

Arinaminpathy N and McLean AR (2009). Evolution and emergence of novel human infections. *Proc Biol Sci* 276(1675): 3937–43.

Arino J, Brauer F, van Den Driessche P, Watmough J and Wu J (2006). Simple models for containment of a pandemic. *J R Soc Interface* 3(8): 453–7.

Athreya K and Ney P (2004). *Branching processes*. Dover Publications, Mineola.

Baigent SJ and McCauley JW (2003). Influenza type A in humans, mammals and birds: determinants of virus virulence, host-range and interspecies transmission. *Bioessays* 25(7): 657–71.

Balcan D, Colizza V, Gonçalves B et al. (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. *P Natl Acad Sci USA* 106(51): 21484–9.

Bansal S, Grenfell BT and Meyers LA (2007). When individual behaviour matters: homogeneous and network models in epidemiology. *J R Soc Interface* 4(16): 879–91.

Bansal S, Pourbohloul B and Meyers LA (2006). A comparative analysis of influenza vaccination programs. *PLoS Med* 3(10): e387.

Beck G and Habicht GS (1996). Immunity and the invertebrates. *Sci Am* 60(66).

Becker N (1981). A general chain binomial model for infectious diseases. *Biometrics* 37(2): 251–8.

Beigel JH, Farrar J, Han AM et al. (2005). Avian influenza A (H5N1) infection in humans. *N Engl J Med* 353(13): 1374–85.

Benmayor R, Hodgson DJ, Perron GG and Buckling A (2009). Host mixing and disease emergence. *Curr Biol* 19(9): 764–7.

Biek R, Walsh PD, Leroy EM and Real LA (2006). Recent common ancestry of Ebola Zaire virus found in a bat reservoir. *PLoS Pathog* 2(10): e90.

Blight KJ (2000). Efficient initiation of HCV RNA replication in cell culture. *Science* 290(5498): 1972–1974.

Boots M and Sasaki A (1999). 'Small worlds' and the evolution of virulence: infection occurs locally and at a distance. *Proc Roy Soc Lond Series B* 266(1432): 1933–1938.

Boots M and Sasaki A (2003). Parasite evolution and extinctions. *Ecol Lett* 6(3): 176–182.

Bouvier N and Palese P (2008). The biology of influenza viruses. *Vaccine* 26: D49–D53.

Briese T, Paweska JT, McMullan LK et al. (2009). Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. *PLoS Pathog* 5(5): e1000455.

Brockmann D, Hufnagel L and Geisel T (2005). Dynamics of modern epidemics. Mclean A and May RM, editors, *SARS: a case study in emerging infections*, Oxford University Press, I. pages 81–91.

Brockmann D, Hufnagel L and Geisel T (2006). The scaling laws of human travel. *Nature* 439(7075): 462–5.

Brown V and White KAJ (2010). The HPV vaccination strategy: could male vaccination have a significant impact? *Comput Math Methods Med* 11(3): 223–37.

Cameron WB (1963). *Informal sociology: a casual introduction to sociological thinking*. Random House.

Cauchemez S, Donnelly CA, Reed C et al. (2009). Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States. *N Engl J Med* 361(27): 2619–27.

Census and Statistics Department (2001). *2001 population census*. Technical report.

Chen H (2009). Serologic survey of pandemic (H1N1) 2009 virus, Guangxi province, China. *Emerg Infect Dis* (6): 10–13.

Cliff AD, Smallman-Raynor MR, Haggett P, Stroup DF and Thacker SB (2009). *Emergence and re-emergence infectious diseases - a geographical analysis*. Oxford University Press, Oxford.

Cohen ML (2000). Changing patterns of infectious disease. *Nature* 406(6797): 762–7.

Coleman PG and Welburn SC (2004). Are fitness costs associated with resistance to human serum in Trypanosoma brucei rhodesiense? *Trends Parasitol* 20(7): 311–5.

Colizza V, Barrat A, Barthélemy M and Vespignani A (2006). The role of the airline transportation network in the prediction and predictability of global epidemics. *P Natl Acad Sci USA* 103(7): 2015–20.

Colizza V and Vespignani A (2008). Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: theory and simulations. *J Theor Biol* 251(3): 450–67.

Dawood FS, Jain S, Finelli L et al. (2009). Emergence of a novel swine-origin influenza A(H1N1) virus in humans. *N Engl J Med* 360(25): 2605–15.

Day T, André JB and Park A (2006). The evolutionary emergence of pandemic influenza. *Proc Biol Sci* 273(1604): 2945–53.

de Silva UC, Warachit J, Waicharoen S and Chittaganpitch M (2009). A preliminary analysis of the epidemiology of influenza A(H1N1)v virus infection in Thailand from early outbreak data, June-July 2009. *Euro Surveill* 14(31): 6–8.

DEFRA - Department for Environment, Food and Rural Affairs (2011). *Zoonoses report 2010*. DEFRA, London.

Delangue J, Salez N, Ninove L et al. (2011). Serological study of the 2009 pandemic due to influenza A(H1N1) in the metropolitan French population. *Clin Microbiol Infect* : 4–10.

Dennehy JJ (2009). Bacteriophages as model organisms for virus emergence research. *Trends Microbiol* 17(10): 450–7.

Dennehy JJ, Friedenberg Na, McBride RC, Holt RD and Turner PE (2010). Experimental evidence that source genetic variation drives pathogen emergence. *Proc Biol Sci* 277: 3113–3121.

Diekmann O, Heesterbeek J and Metz J (1990). On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J Math Biol* 28(4): 417–21.

Diekmann O, Heesterbeek JaP and Roberts MG (2010). The construction of next-generation matrices for compartmental epidemic models. *J R Soc Interface* 7(47): 873–85.

Directgov (2009). *School term dates*. Technical report.

Dobson AP (2004). Population dynamics of pathogens with multiple host species. *Am Nat* 164 Suppl(November 2004): S64–78.

Dodd PJ and Ferguson NM (2007). Approximate disease dynamics in household-structured populations. *J R Soc Interface* 4(17): 1103–6.

Donaldson LJ, Rutter PD, Ellis BM et al. (2009). Mortality from pandemic A/H1N1 2009 influenza in England: public health surveillance study. *BMJ* 339: b5213–b5213.

Duffy S, Shackelton La and Holmes EC (2008). Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet* 9(4): 267–76.

Epstein PR (2001). Climate change and emerging infectious diseases. *Microbes Infect* 3(9): 747–54.

Eubank S, Guclu H, Kumar VSA et al. (2004). Modelling disease outbreaks in realistic urban social networks. *Nature* 429(6988): 180–4.

Ezenwa VO, Godsey MS, King RJ and Guptill SC (2006). Avian diversity and West Nile virus: testing associations between biodiversity and infectious disease risk. *Proc Biol Sci* 273(1582): 109–17.

Ezzell C (2000). Care for a dying continent. *Sci Am* 282(5): 96–105.

Faust K (1999). Spatial arrangement of social and economic networks among villages in Nang Rong District, Thailand. *Soc Networks* 21(4): 311–337.

Ferguson NM, Cummings DAT, Cauchemez S et al. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* 437(7056): 209–14.

Frank Sa and Jeffrey JS (2001). The probability of severe disease in zoonotic and commensal infections. *Proc Biol Sci* 268(1462): 53–60.

Fraser C, Donnelly CA, Cauchemez S et al. (2009). Pandemic potential of a strain of influenza A(H1N1): early findings. *Science* 324(5934): 1557–61.

Freed EO (1998). Hiv-1 gag proteins: diverse functions in the virus life cycle. *Virology* 251(1): 1 – 15.

Fujii Y, Goto H, Watanabe T, Yoshida T and Kawaoka Y (2003). Selective incorporation of influenza virus RNA segments into virions. *P Natl Acad Sci USA* 100(4): 2002–7.

Futuyma DJ and Moreno G (1988). The evolution of ecological specialization. *Annu Rev Ecol Syst* 19(1): 207–233.

Gago S, Elena SF, Flores R and Sanjuán R (2009). Extremely high mutation rate of a hammerhead viroid. *Science* 323(5919): 1308.

Gillespie DT (1977). Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81(25): 2340–2361.

Girard MP, Tam JS, Assossou OM and Kieny MP (2010). The 2009 A(H1N1) influenza virus pandemic: a review. *Vaccine* 28(31): 4895–4902.

Gokhale CS, Iwasa Y, Nowak MA and Traulsen A (2009). The pace of evolution across fitness valleys. *J Theor Biol* 259(3): 613–20.

Guan Y, Zheng BJ, He YQ et al. (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* 302(5643): 276–8.

Hagenaars TJ, Donnelly CA and Ferguson NM (2004). Spatial heterogeneity and the persistence of infectious diseases. *J Theor Biol* 229(3): 349–59.

Hamer WH (1906). On epidemic disease in England. *Lancet* 167(4307): 733–739.

Hartfield M, White KAJ and Kurtenbach K (2011). The role of deer in facilitating the spatial spread of the pathogen Borrelia burgdorferi. *Theor Ecol* 4(1): 27–36.

Hatta M, Gao P, Halfmann P and Kawaoka Y (2001). Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. *Science* 293(5536): 1840–2.

Hay SI, Guerra CA, Tatem AJ, Atkinson PM and Snow RW (2005). Tropical infectious diseases: urbanization, malaria transmission and disease burden in africa. *Nat Rev Micro* 3(1): 81–90.

Hays JN (2005). *Epidemics and pandemics: their impacts on human history*. ABC-CLIO, Santa Barbara, 1st edition.

Health Protection Agency (2009a). *Method used to estimate new pandemic (H1N1) 2009 influenza cases in England in the week 3 August to 9 August 2009*. Technical report, Health Protection Agency.

Health Protection Agency (2009b). *Synopsis of the method used to estimate the number of pandemic influenza (H1N1) 2009 cases in England in the week 21 to 27July 2009*. Technical report, Health Protection Agency.

Health Protection Agency (2009c). *Weekly national influenza reports*. Technical report, Health Protection Agency.

Health Protection Agency (2009d). *Weekly pandemic flu media update*. Technical report.

Heffernan JM, Smith RJ and Wahl LM (2005). Perspectives on the basic reproductive ratio. *J R Soc Interface* 2(4): 281–93.

Holmes EC (2006). The evolution of viral emergence. *P Natl Acad Sci USA* 103(13): 4803–4.

Holmes EC (2009a). *The evolution and emergence of RNA viruses.* Oxford University Press, Oxford, 1st edition.

Holmes EC (2009b). The evolutionary genetics of emerging viruses. *Annu Rev Ecol Evol S* 40(1): 353–372.

Holmes EC (2010). The comparative genomics of viral emergence. *P Natl Acad Sci USA* 107 Suppl: 1742–6.

Holmes K, Zelus B, Schickli J and Weiss S (2001). Receptor specificity and receptor-induced conformational changes in mouse hepatitis virus spike glycoprotein. *Adv Exp Med Biol* 494: 173–181.

Holt RD, Dobson AP, Begon M, Bowers RG and Schauber EM (2003). Parasite establishment in host communities. *Ecol Lett* 6(9): 837–842.

Holt RD and Lawton JH (1993). Apparent competition and enemy-free space in insect host-parasitoid communities. *Am Nat* 142(4): 623–645.

Holt RD and Pickering J (1985). Infectious disease and species coexistence: a model of Lotka-Volterra form. *Am Nat* 126(2): 196–211.

Horimoto T and Kawaoka Y (2005). Influenza: lessons from past pandemics, warnings from current incidents. *Nat Rev Microbiol* 3(8): 591–600.

Hoskisson Pa and Trevors JT (2010). Shifting trends in pathogen dynamics on a changing planet. *A Van Leeuw J* 98(4): 423–7.

Hufnagel L, Brockmann D and Geisel T (2004). Forecast and control of epidemics in a globalized world. *P Natl Acad Sci USA* 101(42): 15124–9.

Instituto Brasileiro de Geografia e Estatistica (2000). *Censos demograficos 2000*. Technical report.

Isaacs A and Burke DC (1958). Mode of action of interferon. *Nature* 182(4642): 1073–1074.

Isaacs A and Lindenmann J (1957). Virus interference 1: the interferon. *P Natl Acad Sci USA* 147(927): 258–267.

Iwasa Y, Michor F and Nowak MA (2004). Some basic properties of immune selection. *J Theor Biol* 229(2): 179–88.

Jacob C (2010). Branching processes: their role in epidemiology. *Int J Environ Res Public Health* 7(3): 1186–204.

Jacob C and Viet aF (2003). Epidemiological modeling in a branching population. Particular case of a general SIS model with two age classes. *Math Biosci* 182(1): 93–111.

Janeway CA, Travers P, Walport M and Shlomchik MJ (2005). *Immunobiology*. Garland Science Publishing, New York.

Jansen Vaa, Stollenwerk N, Jensen HJ et al. (2003). Measles outbreaks in a population with declining vaccine uptake. *Science* 301(5634): 804.

Johnson NPaS and Mueller J (2002). Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. *Bull Hist Med* 76(1): 105–115.

Jones KE, Patel NG, Levy Ma et al. (2008). Global trends in emerging infectious diseases. *Nature* 451(7181): 990–3.

Kao RR, Gravenor MB, Baylis M et al. (2002). The potential size and duration of an epidemic of bovine spongiform encephalopathy in British sheep. *Science* 295(5553): 332–335.

Karlin S and Taylor HM (1975). *A first course in stochastic processes*. Academic Press, Inc., San Diego, second edi edition.

Keeling MJ (1999). The effects of local spatial structure on epidemiological invasions. *Proc Biol Sci* 266(1421): 859–67.

Keeling MJ and Eames KTD (2005). Networks and epidemic models. *J R Soc Interface* 2(4): 295–307.

Keeling MJ and Rohani P (2002). Estimating spatial coupling in epidemiological systems: a mechanistic approach. *Ecol Lett* 5(1): 20–29.

Keeling MJ and Rohani P (2007). *Modelling infectious diseases in humans and animals*. Princeton University Press, Princeton, 1st edition.

Keeling MJ and Ross JV (2008). On methods for studying stochastic disease dynamics. *J R Soc Interface* 5(19): 171–81.

Keesing F, Belden LK, Daszak P et al. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* 468(7324): 647–652.

Keesing F, Holt RD and Ostfeld RS (2006). Effects of species diversity on disease risk. *Ecol Lett* 9(4): 485–98.

Kermack WO and McKendrick AG (1927). A contribution to the mathematical theory of epidemics. *P R Soc Lond A Conta* 115(772): 700–721.

Kiss IZ, Green DM and Kao RR (2005). Disease contact tracing in random and clustered networks. *Proc R Soc B* 272(1570): 1407–1414.

Kiss IZ, Green DM and Kao RR (2006). Infectious disease control using contact tracing in random and scale-free networks. *J R Soc Interface* 3(6): 55–62.

Kubiak RJ, Arinaminpathy N and McLean AR (2010). Insights into the evolution and emergence of a novel infectious disease. *PLoS Comput Biol* 6(9): e1000947.

Kubiak RJ and McLean AR (2012). Why was the 2009 infuenza pandemic in England so small? *PLoS ONE* 7(2): e30223.

Kuiken T, Holmes EC, McCauley J et al. (2006). Host species barriers to influenza virus infections. *Science* 312(5772): 394–7.

Leinbach TR (1983). Rural transport and population mobility in Indonesia. *J Dev Areas* 17(3): 349–63.

Leroy EM, Kumulungui B, Pourrut X et al. (2005). Fruit bats as reservoirs of Ebola virus. *Nature* 438(7068): 575–6.

Lessler J, Reich NG, Cummings DaT et al. (2009). Outbreak of 2009 pandemic influenza A(H1N1) at a New York City school. *N Engl J Med* 361(27): 2628–36.

Li W, Shi Z, Yu M et al. (2005). Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310(5748): 676–9.

Li W, Wong Sk, Li F et al. (2006). Animal origins of the severe acute respiratory syndrome coronavirus: insight from ACE2-S-protein interactions. *J Virol* 80(9): 4211–9.

Lipsitch M, Cohen T, Cooper B et al. (2003). Transmission dynamics and control of severe acute respiratory syndrome. *Science* 300(5627): 1966–70.

Litman GW (1996). Sharks and the origins of vertebrate immunity. *Sci Am* 60(66).

Lloyd AL and May RM (1996). Spatial heterogeneity in epidemic models. *J Theor Biol* 179(1): 1–11.

Lloyd-Smith JO, Schreiber SJ, Kopp PE and Getz WM (2005). Superspreading and the effect of individual variation on disease emergence. *Nature* 438(7066): 355–9.

Longini IM, Nizam A, Xu S et al. (2005). Containing pandemic influenza at the source. *Science* 309(5737): 1083–7.

Louz D, Bergmans HE, Loos BP and Hoeben RC (2010). Emergence of viral diseases: mathematical modeling as a tool for infection control, policy and decision making. *Crit Rev Microbiol* 36(3): 195–211.

Matrosovich MN, Matrosovich TY, Gray T, Roberts Na and Klenk HD (2004). Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *P Natl Acad Sci USA* 101(13): 4620–4.

May RM (2004). Uses and abuses of mathematics in biology. *Science* 303(5659): 790–3.

May RM and Anderson RM (1983). Epidemiology and genetics in the coevolution of parasites and hosts. *Proc Roy Soc Lond Series B* 219(1216): 218–313.

May RM and Anderson RM (1984). Spatial heterogeneity and the design of immunization programs. *Math Biosci* 72(1): 83–111.

May RM, Anderson RM and McLean AR (1988). Possible demographic consequences of HIV/AIDS epidemics. *Math Biosci* 90: 475–505.

May RM, Gupta S and McLean AR (2001). Infectious disease dynamics: what characterizes a successful invader? *Philos Trans R Soc Lond B Biol Sci* 356(1410): 901–10.

McBryde E, Bergeri I, Gemert CV et al. (2009). Early transmission characteristics of influenza A(H1N1)v in Australia: Victorian state, 16 May - 3 June 2009. *Euro Surveill* 14(42): 1–6.

McCallum H, Barlow N and Hone J (2001). How should pathogen transmission be modelled? *Trends Ecol Evol* 16(6): 295 – 300.

McMichael AJ (2004). Environmental and social influences on emerging infectious diseases: past, present and future. *Philos Trans R Soc Lond B Biol Sci* 359(1447): 1049–1058.

McMichael T (2001). *Human frontiers, environments and disease*. Cambridge University Press, Cambridge.

Meyer BJ, de la Torre JC and Southern PJ (2002). Arenaviruses: genomic RNAs, transcription, and replication. *Curr Top Microbiol Immunol* 262: 139–157.

Meyers LA, Pourbohloul B, Newman MEJ, Skowronski DM and Brunham RC (2005). Network theory and SARS: predicting outbreak diversity. *J Theor Biol* 232(1): 71–81.

Miller E, Hoschler K, Hardelid P et al. (2010). Incidence of 2009 pandemic influenza A H1N1 infection in England: a cross-sectional serological study. *Lancet* 375(9720): 1100–1108.

Mims CA, Nash A and Stephen J (2000). *Mims' pathogenesis of infectious disease*. Academic Press.

Molinari NAM, Ortega-Sanchez IR, Messonnier ML et al. (2007). The annual impact of seasonal influenza in the US: measuring disease burden and costs. *Vaccine* 25(27): 5086 – 5096.

Monath TP (1994). Dengue: the risk to developed and developing countries. *P Natl Acad Sci USA* 91(7): 2395–2400.

Morse SS (1995). Factors in the emergence of infectious diseases. *Emerg Infect Dis* 1(1): 7–15.

Munayco CV, Gómez J, Arrasco J et al. (2009). Epidemiological and transmissibility analysis of influenza A(H1N1)v in a southern hemisphere setting: Peru. *Euro Surveill* 14(32): 1–5.

Nagy PD and Simon aE (1997). New insights into the mechanisms of RNA recombination. *Virology* 235(1): 1–9.

Newman MEJ (2002). Spread of epidemic disease on networks. *Phys Rev E* 66(1): 1–11.

Newman MEJ (2003). The structure and function of complex networks. *SIAM Review* 45(2): 167.

Nie Y, Wang P, Shi X et al. (2004). Highly infectious SARS-CoV pseudotyped virus reveals the cell tropism and its correlation with receptor expression. *Biochem Biophys Res Commun* 321(4): 994–1000.

Nishiura H, Chowell G, Safan M and Castillo-Chavez C (2010). Pros and cons of estimating the reproduction number from early epidemic growth rate of influenza A (H1N1) 2009. *Theor Biol Med Model* 7: 1.

Office for National Statistics (2001). *Census 2001*. Technical report.

Office for National Statistics (2009). *2009 population estimates in England*. Technical report.

Olofsson S, Kumlin U, Dimock K and Arnberg N (2005). Avian influenza and sialic acid receptors: more than meets the eye? *Lancet Infect Dis* 5(3): 184–8.

Palmer SR, Soulsby EJL and Simpson DIH (1998). *Zoonoses*. Oxford Medical Publications, Oxford, 1st edition.

Parrish CR, Holmes EC, Morens DM et al. (2008). Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiol Mol Biol Rev* 72(3): 457–70.

Pepin KM, Lass S, Pulliam JRC, Read AF and Lloyd-Smith JO (2010). Identifying genetic markers of adaptation for surveillance of viral host jumps. *Nat Rev Microbiol* 8(11): 802–13.

Pianka ER (1970). On r- and K-selection. *Am Nat* 104(940): 592–597.

Pilgrim ES, Crawley MJ and Dolphin K (2004). Patterns of rarity in the native British flora. *Biol Conserv* 120(2): 161 – 170.

Pourbohloul B, Ahued A, Davoudi B et al. (2009). Initial human transmission dynamics of the pandemic (H1N1) 2009 virus in North America. *Influenza Other Respi Viruses* 3(5): 215–22.

Power AG and Mitchell CE (2004). Pathogen spillover in disease epidemics. *Am Nat* 164 Suppl(November 2004): S79–89.

Prachayangprecha S, Makkoch J, Payungporn S et al. (2010). Serological analysis of human pandemic influenza (H1N1) in Thailand. *J Health Popul Nutr* 28(6): 537–44.

Prochaska JO, Redding Ca, Harlow LL, Rossi JS and Velicer WF (1994). The transtheoretical model of change and HIV prevention: a review. *Health Educ Behav* 21(4): 471–486.

Read A and Skorping A (1995). The evolution of tissue migration by parasitic nematode larvae. *Parasitology* 111 (Pt 3: 359–71.

Read JM, Eames KTD and Edmunds WJ (2008). Dynamic social networks and the implications for the spread of infectious disease. *J R Soc Interface* 5(26): 1001–7.

Read JM and Keeling MJ (2003). Disease evolution on networks: the role of contact structure. *Proc Biol Sci* 270(1516): 699–708.

Reid AH, Fanning TG, Hultin JV and Taubenberger JK (1999). Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene. *P Natl Acad Sci USA* 96(4): 1651–6.

Reluga T, Meza R, Walton DB and Galvani aP (2007). Reservoir interactions and disease emergence. *Theor Popul Biol* 72(3): 400–8.

Renault M (2008). Andre's actual method and its application to the generalized ballot problem. *Am Math Mon* 115(4): 358–363.

Riley S and Ferguson NM (2006). Smallpox transmission and control: spatial dynamics in Great Britain. *P Natl Acad Sci USA* 103(33): 12637–42.

Roberts MG, Baker M, Jennings LC, Sertsou G and Wilson N (2007). A model for the spread and control of pandemic influenza in an isolated geographical region. *J R Soc Interface* 4(13): 325–30.

Rodriguez JJ, Parisien JP and Horvath CM (2002). Nipah virus V protein evades alpha and gamma interferons by preventing STAT1 and STAT2 activation and nuclear accumulation. *J Virol* 76(22): 11476–11483.

Rogers DJ, Wilson AJ, Hay SI and Graham AJ (2006). The global distribution of yellow fever and dengue. *Adv Parasitol* 62: 181–220.

Roossinck MJ (1997). Mechanisms of plant virus evolution. *Annu Rev Phytopathol* 35(1): 191–209.

Scherer A and McLean AR (2002). Mathematical models of vaccination. *Br Med Bull* 62: 187–199.

Scholtissek C (1995). Molecular evolution of influenza viruses. *Virus Genes* 11(2-3): 209–15.

Scholtissek C, Rohde W, Von Hoyningen V and Rott R (1978). On the origin of the human influenza virus subtypes H2N2 and H3N2. *Virology* 87(1): 13–20.

Schrag SJ and Wiener P (1995). Emerging infectious disease: what are the relative roles of ecology and evolution? *Trends Ecol Evol* 10(8): 319–24.

Schultz U, Kaspers B and Staeheli P (2004). The interferon system of non-mammalian vertebrates. *Dev Comp Immunol* 28(5): 499–508.

Seo SH, Hoffmann E and Webster RG (2002). Lethal H5N1 influenza viruses escape host anti-viral cytokine responses. *Nat Med* 8(9): 950–4.

Skowronski DM, Astell C, Brunham RC et al. (2005). Severe acute respiratory syndrome (SARS): a year in review. *Annu Rev Med* 56: 357–81.

Smith GJD, Vijaykrishna D, Bahl J et al. (2009). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459(7250): 1122–1125.

Smith H and Sweet C (1988). Lessons for human influenza from pathogenicity studies with ferrets. *Rev Infect Dis* 10(1): 56–75.

Smolinski MS, Hamburg MA and Lederberg J (2003). *Microbial threats to health - emergence, detection and response*. The National Academies Press, Washington, D.C.

Statistics Bureau Japan (2005). *2005 population census*. Technical report.

Stearns SC (1992). *The evolution of life histories*. Oxford University Press, Oxford.

Taubenberger JK and Morens DM (2006). Molecular virology: was the 1918 flu avian in origin? *Emerg Infect Dis* 12(1): 15–22.

Taylor LH, Latham SM and Woolhouse MEJ (2001). Risk factors for human disease emergence. *Philos Trans R Soc Lond B Biol Sci* 356(1411): 983–9.

Thackray LB and Holmes KV (2004). Amino acid substitutions and an insertion in the spike glycoprotein extend the host range of the murine coronavirus MHV-A59. *Virology* 324(2): 510–24.

Thanh HX, Anh DN and Tacoli C (2005). Livelihood diversification and rural-urban linkages in Vietnam's Red River Delta. *IIED Working Paper Series on Rural-Urban Interactions and Livelihood Strategies* .

Tildesley MJ and Keeling MJ (2009). Is R_0 a good predictor of final epidemic size: foot-and-mouth disease in the UK. *J Theor Biol* 258(4): 623–9.

Trapman P and Bootsma MCJ (2009). A useful relationship between epidemiology and queueing theory: the distribution of the number of infectives at the moment of the first detection. *Math Biosci* 219(1): 15–22.

Trapman P, Meester R and Heesterbeek H (2004). A branching model for the spread of infectious animal diseases in varying environments. *J Math Biol* 49(6): 553–76.

Traulsen A, Iwasa Y and Nowak MA (2007). The fastest evolutionary trajectory. *J Theor Biol* 249(3): 617–23.

Tuite AR, Greer AL, Whelan M et al. (2010). Estimated epidemiologic parameters and morbidity associated with pandemic H1N1 influenza. *Can Med Assoc J* 182(2): 131–6.

US Census Bureau (2000). *Census 2000*. Technical report.

Van Tienderen PH (1991). Evolution of generalists and specialist in spatially heterogeneous environments. *Evolution* 45(6): 1317–1331.

Vazquez A (2007). Epidemic outbreaks on structured populations. *J Theor Biol* 245(1): 125–9.

Wallinga J, van Boven M and Lipsitch M (2010). Optimizing infectious disease interventions during an emerging epidemic. *P Natl Acad Sci USA* 107(2): 923–8.

Watts DJ, Muhamad R, Medina DC and Dodds PS (2005). Multiscale, resurgent epidemics in a hierarchical metapopulation model. *P Natl Acad Sci USA* 102(32): 11157–62.

Watts DJ and Strogatz SH (1998). Collective dynamics of 'small-world' networks. *Nature* 393(6684): 440–2.

Wearing HJ, Rohani P and Keeling MJ (2005). Appropriate models for the management of infectious diseases. *PLoS Med* 2(7): e174.

Webby R, Hoffmann E and Webster RG (2004a). Molecular constraints to interspecies transmission of viral pathogens. *Nat Med* 10(12 Suppl): S77–81.

Webby R, Perez DR, Coleman JS et al. (2004b). Responsiveness to a pandemic alert: use of reverse genetics for rapid development of influenza vaccines. *Lancet* 363(9415): 1099–103.

Wein LM, Craft DL and Kaplan EH (2003). Emergency response to an anthrax attack. *P Natl Acad Sci USA* 100(7): 4346–51.

Weiss RA and McMichael AJ (2004). Social and environmental risk factors in the emergence of infectious diseases. *Nat Med* 10(12): S70–6.

White KAJ and Gilligan CA (1998). Spatial heterogeneity in three-species, plant-parasite-hyperparasite, systems. *Phil Trans R Soc Lond B Biol Sci* 353(1368): 543–557.

Wiklund C and Friberg M (2009). The evolutionary ecology of generalization: among-year variation in host plant use and offspring survival in a butterfly. *Ecology* 90(12): 3406–17.

Wilf HS (1994). *Generatingfunctionology*. Academic Press, Inc., New York, 2nd edition.

Wolfe ND, Dunavan CP and Diamond J (2007). Origins of major human infectious diseases. *Nature* 447(7142): 279–83.

Woolhouse MEJ (2002). Population biology of emerging and re-emerging pathogens. *Trends Microbiol* 10: S3–S7.

Woolhouse MEJ and Dye C (2001). Preface. *Phil Trans R Soc Lond B* 356: 981–982.

Woolhouse MEJ and Gowtage-Sequeria S (2005). Host range and emerging and reemerging pathogens. *Emerg Infect Dis* 11(12): 1842–7.

Woolhouse MEJ, Haydon DT and Antia R (2005). Emerging pathogens: the epidemiology and evolution of species jumps. *Trends Ecol Evol* 20(5): 238–44.

Woolhouse MEJ, Taylor LH and Haydon DT (2001). Population biology of multihost pathogens. *Science* 292(5519): 1109–12.

World Health Organization (2006). *SARS: how a global epidemic was stopped*. WHO Press, Geneva, 1st edition.

World Health Organization (2011). *Ebola*. Technical report.

Worobey M and Holmes EC (1999). Evolutionary aspects of recombination in RNA viruses. *J Gen Virol* 80 (Pt 10: 2535–43.

Xu H (2001). Commuting town workers: the case of Qinshan, China. *Habitat Int* 25(1): 35–47.

Yang Y, Sugimoto JD, Halloran ME et al. (2009). The transmissibility and control of pandemic influenza A (H1N1) virus. *Science* 326(5953): 729–33.

Yates A, Antia R and Regoes RR (2006). How do pathogen evolution and host heterogeneity interact in disease emergence? *Proc Biol Sci* 273(1605): 3075–83.

APPENDIX

Attached are a code sample of translating an ODE system into a stochastic simulation using the Gillespie algorithm in C and copies of my published research papers, Kubiak et al. (2010) and Kubiak and McLean (2012). Each author's contribution is stated at the beginning of this thesis.

```c
/* File: detectiondyn.c */

#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include "detectiondyn_global.h"
#include "detectiondyn_local.h"
#include "random.h"

/* Global variables */
// Variables with straw to python
int patient_zero = 0;
int population_size = 0;
int max_days = 300;


/* Functions used */

/* Function to read pi array from Python in */
void set_pi(double* in_array, int length) {
    beta_length = length;

    pi = (double*) malloc(beta_length * sizeof(double));
    if (pi == NULL) {
        printf("ERROR! pi is out of memory! \n");
    }

    for(i = 0; i < beta_length; i++) {
        pi[i] = in_array[i]; // Store the data in pi
    }
}

/* Function to read beta array from Python in */
void set_beta(double* in_array, int length) {
    if (beta_length != length) {
        printf("ERROR! beta has not the same length as pi! \n");
    }

    beta = (double*) malloc(beta_length * sizeof(double));
    if (beta == NULL) {
        printf("ERROR! beta is out of memory! \n");
    }

    for(i = 0; i < beta_length; i++) {
        beta[i] = in_array[i]; // Store the data in beta
    }
}

/* Function to read gamma array from Python in */
void set_gamma(double* in_array, int length) {
    if (beta_length != length) {
        printf("ERROR! gamma has not the same length as pi! \n");
    }

    gamma = (double*) malloc(beta_length * sizeof(double));
    if (gamma == NULL) {
        printf("ERROR! gamma is out of memory! \n");
    }
```

```
    for(i = 0; i < beta_length; i++) {
        gamma[i] = in_array[i]; // Store the data in gamma
    }
}

/* Function to get new_infected_overall array out to Python */
void get_new_infected(double* out_array, int d1, int d2) {
    // Check for errors in the dimensions
    if (d1 != beta_length) {
        printf("Error!!! DIM1 of get_new_infected is not beta_length!");
    }
    if (d2 != population_size) {
        printf("Error!!! DIM2 of get_new_infected is not population_size!"
            );
    }

    // Store the data in the out_array
    for(i = 0; i < d1; i++) {
        for(j = 0; j < d2; j++) {
            index_number = (int) i * d2 + j;
            out_array[index_number] = new_infected[i][j];
        }
    }
}

/* Function to get new_infected_overall array out to Python */
void get_outbreak_size(double* out_array, int d1) {
    // Check for errors in the dimensions
    if (d1 != beta_length) {
        printf("Error!!! DIM1 of get_outbreak_size is not beta_length!");
    }

    // Store the data in the out_array
    for(i = 0; i < d1; i++) {
        out_array[i] = outbreak_size[i];
    }
}

/* Function to round numbers */
int runden(double number) {
    return (number >= 0) ? (int)(number + 0.5) : (int)(number - 0.5);
}

/* Build the SIR array */
void construct_SIR(void) {
    SIR = (int*) malloc((beta_length+3) * sizeof(int)); // Define size of
        SIR
    if (SIR == NULL) {
        printf("ERROR! SIR is out of memory! \n");
    }

    // Reminder: it is SIR[state] S->Id->Iu->R
    for (i = 1; i < beta_length+3; i++) {
        SIR[i] = 0; // Write it in
    }
    SIR[0] = population_size - 1; // Start with the number of hosts in the
        population
    SIR[patient_zero+1]++; // Infect one person from one of the strains
        (patient_zero can random 0-1 number from python)
```

```
    }

/* Construct the transfer rates */
void construct_h(void) {
    // Create the arrays
    number_reactions = (int) (beta_length * 3);
    h = (double*) malloc(number_reactions * sizeof(double));
    if(h == NULL) {
        printf("ERROR! h is out of memory! \n");
    }
    plus_state = (int*) malloc(number_reactions * sizeof(int));
    if(plus_state == NULL) {
        printf("ERROR! plus_state is out of memory! \n");
    }
    minus_state = (int*) malloc(number_reactions * sizeof(int));
    if(minus_state == NULL) {
        printf("ERROR! minus_state is out of memory! \n");
    }

    // Construct it
    // Reminder: it is SIR[state] S->Is->Ia->R
    counter = 0;

    // Iterate through infection types (s -> a)
    for(i = 0; i < beta_length; i++) {
        // Infections
        for(j = 0; j < beta_length; j++) {
            h[counter] = pi[i] * beta[j];
            minus_state[counter] = 0;
            plus_state[counter] = i+1;
            counter++;
        }
        // Recoveries
        h[counter] = gamma[i];
        minus_state[counter] = i+1;
        plus_state[counter] = i+3;
        counter++;
    }

    // Test
    if(counter != number_reactions) {
        printf("ERROR! The number of reactions does not match in
            construct_h()! \n");
    }
}

/* Construct the transfer weights */
void construct_a(void) {
    // Create the arrays
    a = (double*) malloc(number_reactions * sizeof(double));
    if(a == NULL) {
        printf("ERROR! a is out of memory! \n");
    }

    // Construct it
    // Reminder: it is SIR[state] S->Is->Ia->R
    counter = 0;

        // Iterate through infection types (s -> a)
```

```
    for(i = 0; i < beta_length; i++) {
        // Infections
        for(j = 0; j < beta_length; j++) {
            a[counter] = (double) SIR[0] * SIR[j+1] / population_size;
            counter++;
        }
        // Recoveries
        a[counter] = (double) SIR[i+1];
        counter++;
    }

    // Test
    if(counter != number_reactions) {
        printf("ERROR! The number of reactions does not match in
            construct_a()! \n");
    }
}

/* Construct new infected and its help arrays */
void construct_new_infected(void) {
    // First dimension
    new_infected = (double**) malloc(beta_length * sizeof(double*));

    // Second is number of days
    for (i = 0; i < beta_length; i++) {
        new_infected[i] = (double*) malloc((population_size) * sizeof
            (double));
    }

    // Check if everything worked good
    if (new_infected == NULL) {
        printf("ERROR! new_infected is out of memory! \n");
    }

    // Write a 0 in everything
    for(i = 0; i < beta_length; i++) {
        for(j = 0; j < population_size; j++) {
            new_infected[i][j] = 0.0;
        }
    }
}

/* Construct an array for the outbreak size*/
void construct_outbreak_size(void) {
    // First dimension
    outbreak_size = (int*) malloc(beta_length * sizeof(int*));

    // Check if everything worked good
    if (outbreak_size == NULL) {
        printf("ERROR! outbreak_size is out of memory! \n");
    }

    // Write a 0 in everything
    for(i = 0; i < beta_length; i++) {
        outbreak_size[i] = 0;
    }
}

/* Calculate a0 */
```

```
void calc_a0(void) {
    double_sum = 0.0;
    for (i = 0; i < number_reactions; i++) {
        double_sum += a[i] * h[i];
    }
    a0 = double_sum;
}

/* Calculate tau */
void calc_tau(void) {
    tau = (1.0 / a0) * log(1.0 / genrand_real1());
}

/* Calculate which reaction is realized */
void calc_reaction_id(void) {
    r = a0 * genrand_real1();
    double_sum = 0.0;

    // Get reaction
    i = 0;
    do {
        double_sum += a[i] * h[i];
        i++;
    } while(double_sum < r);
    reaction_id = i - 1;
}

/* Re-Calc SIR according to transfer and store the new infected */
void recalc_SIR(void){
    SIR[minus_state[reaction_id]]--;
    SIR[plus_state[reaction_id]]++;
}

/* Caclulate the number of infected */
int calc_infected(void) {
    // Reminder: it is SIR[state] S->Id->Iu->R
    // Check for new infection
    if(minus_state[reaction_id] == 0) {
        new_infected[plus_state[reaction_id]-1][outbreak_size[plus_state
            [reaction_id]-1]] += zeit;
        outbreak_size[plus_state[reaction_id]-1]++;
        int_sum++;
    }

    // Check for recoveries
    if(plus_state[reaction_id] >= beta_length+1) {
        int_sum--;
    }

    return int_sum;
}

/* Initialize the random number generator */
void init_mersennetwister(int seed_py) {
    init_genrand((unsigned long) seed_py);
}

/* Constructor */
void constructer(void) {
```

```
    // Initialize the main arrays
    construct_SIR();
    construct_h();
    construct_new_infected();
    construct_outbreak_size();
}

/* Destructor */
void destructer(void) {
    // Free memory
    free(SIR);
    free(h);
    free(plus_state);
    free(minus_state);
    for (i = 0; i < beta_length; i++) {
        free(new_infected[i]);
    }
    free(new_infected);
    free(outbreak_size);
}

/* Simulation itself */
void simulation(void) {
    // Construct stuff
    constructer();

    // Set needed values
    int_sum = 1;

    // Get all the needed data
    stop = 0;
    zeit = 0.0;

    // Run the update
    do {
        // Do all the updates
        construct_a();
        calc_a0();
        calc_tau();
        calc_reaction_id();
        recalc_SIR();

        // Add time and check for stop
        zeit += tau;
        if (zeit >= (double) max_days) {
            stop = 1;
            free(a); // Free memory used each loop
            break;
        }

        // Check if new infected occured and get the infected
        actual_inf_number = calc_infected();

        // Check for stop (no new infected)
        if (actual_inf_number == 0) {
            stop = 1;
            free(a); // Free memory used each loop
            break;
        }
```

```
            free(a); // Free memory used each loop
        } while (stop == 0);
    }
```

OPEN ACCESS Freely available online

PLoS COMPUTATIONAL BIOLOGY

Insights into the Evolution and Emergence of a Novel Infectious Disease

Ruben J. Kubiak*, Nimalan Arinaminpathy, Angela R. McLean

Institute for Emerging Infections, James Martin 21st Century School, Department of Zoology, University of Oxford, Oxford, UK

Abstract

Many zoonotic, novel infectious diseases in humans appear as sporadic infections with spatially and temporally restricted outbreaks, as seen with influenza A(H5N1). Adaptation is often a key factor for successfully establishing sustained human-to-human transmission. Here we use simple mathematical models to describe different adaptation scenarios with particular reference to spatial heterogeneity within the human population. We present analytical expressions for the probability of emergence per introduction, as well as the waiting time to a successful emergence event. Furthermore, we derive general analytical results for the statistical properties of emergence events, including the probability distribution of outbreak sizes. We compare our analytical results with a stochastic model, which has previously been studied computationally. Our results suggest that, for typical connection strengths between communities, spatial heterogeneity has only a weak effect on outbreak size distributions, and on the risk of emergence per introduction. For example, if $R_0 = 1.4$ or larger, any village connected to a large city by just ten commuters a day is, effectively, just a part of the city when considering the chances of emergence and the outbreak size distribution. We present empirical data on commuting patterns and show that the vast majority of communities for which such data are available are at least this well interconnected. For plausible parameter ranges, the effects of spatial heterogeneity are likely to be dominated by the evolutionary biology of host adaptation. We conclude by discussing implications for surveillance and control of emerging infections.

Citation: Kubiak RJ, Arinaminpathy N, McLean AR (2010) Insights into the Evolution and Emergence of a Novel Infectious Disease. PLoS Comput Biol 6(9): e1000947. doi:10.1371/journal.pcbi.1000947

Editor: Mark M. Tanaka, University of New South Wales, Australia

Received April 15, 2010; **Accepted** August 31, 2010; **Published** September 30, 2010

Funding: Ruben J. Kubiak gratefully acknowledges funding for this work from the James Martin 21st Century School (http://www.21school.ox.ac.uk) and the EPSRC (http://www.epsrc.ac.uk); Nimalan Arinaminpathy gratefully acknowledges funding from the James Martin 21st Century School; and Angela R. McLean is a Senior Research fellow of All Souls College (http://www.all-souls.ox.ac.uk) whose support is gratefully acknowledged. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ruben.kubiak@zoo.ox.ac.uk

Introduction

Zoonotic emergence of novel human infections poses a significant risk to global public health. For example, the 'Spanish flu' pandemic of 1918 probably originated in birds and caused millions of deaths worldwide [1]. While much less virulent, the subsequent influenza pandemics of 1957, 1968 and 2009 [2,3] are potent reminders of the capacity of the influenza virus to cross the species barrier into humans. Many other pathogens share this capacity: the SARS outbreak of 2003 [4–6] has been linked to bats and palm civets [7,8]. In 2008, a novel arenavirus which killed four out of five patients in South Africa was linked to rodents [9].

Previous work [10,11] has studied models of within-host evolution and between-host transmission, in which an initially poorly transmitting pathogen acquires adaptations to human hosts, following repeated zoonotic introductions until it achieves pandemic potential. These make the natural, simplifying assumption that the host population is homogeneous, so that changes in infection parameters entirely reflect adaptations in the biology of host infection. In reality, however, factors such as human contact patterns [12] and other host heterogeneity [13,14] may also shape the risk and speed of emergence events. We concentrate here on the heterogeneity in the spatial structure of the human host population, an area which has hitherto received little attention in the context of adapting pathogens. We model spatially separated communities with varying types and strengths of interconnections, for example between a village and a city. Our aim is to study under what regimes such 'ecological' structure could have a strong effect on emergence, in comparison with 'evolutionary' factors governing the biology of infection.

In the following section we give an overview of the modelling approach. We then present new analytical results for the simple models studied previously, which ignored spatial host population structure. We use these expressions to provide useful answers to important questions: if we knew how a zoonotic pathogen would adapt to human physiology, could we anticipate its emergence? How reliable would such predictions be? Furthermore, can we predict which zoonoses will cause outbreaks which do not turn into epidemics? Next, we ask: how large does a single, finite host population have to be, for population size to have a negligible effect? We then incorporate spatial heterogeneity by separating the human host population into communities. We present a model in which a small village is connected, by human travel, with a large city as an example of the general case of two interconnected communities. We use this model to ask: how strong do community interconnections have to be for us to safely ignore the separation of a population into spatially structured communities, such as cities and villages? We review available commuting data to ask how these thresholds compare with typical human mobility patterns? We close with a discussion of implications for public health.

Author Summary

Emerging infections are a continuing global public health issue, the most recent example being last year's 'Swine flu' influenza pandemic. However, for many zoonotic pathogens, some adaptation is required to cross the species barrier from an animal reservoir into humans and cause sustained transmission. Previous work has explored the relationship between the evolutionary biology of an adapting pathogen, and the epidemiology of cases that may arise before such a pathogen becomes pandemic-capable. Here, we extend this work to incorporate what is often an important host ecological feature, the spatial distribution of the host population. Many zoonoses occur away from large population centres. For example, HIV is thought to have entered the human population through bushmeat hunters in the sparsely populated jungles of Central Africa. We ask: when a pathogen is evolving to adapt for human transmission, under what circumstances does the spatial structure underlying the human population become important? We approach this question using mathematical models to explore regimes of connectedness between communities. Our results suggest that most communities are sufficiently interconnected to show no effect on the emergence process. We finish by discussing the implications of these findings for public health.

Materials and Methods

Modelling Evolutionary Adaptation

We build on a model of evolution and emergence originally presented by [10] in which a zoonotic pathogen infects humans, and initially has very poor onward transmissability. Thus for people who are infected by animals the average reproductive number, R_0, is well below one ($R_0 \ll 1$). We call this the first reproductive number α_1 for the wildtype strain. Occasionally, during such zoonotic infections, the pathogen acquires genetic changes that increase its ability to pass to other humans. During any chain of transmission the pathogen might adapt sufficiently that it achieves such ease of human-to-human transmission that $R_0 > 1$ and an epidemic becomes possible. Such a process can be characterised by a vector of reproductive numbers (α_i with $i = 1, \ldots, n$), and a vector of mutation probabilities (μ_i with $i = 1, \ldots, n$) where $n-1$ denotes the number of adaptive steps necessary to reach the fully adapted strain n.

In what follows we restrict our attention to the case of $n = 3$, $\mu_{1,2} = \mu$ and $\mu_3 = 0$, allowing us to model two routes of adaptation with opposite and distinct characteristics while minimising the overall complexity in number of required strains. For both routes of adaptation, the first, wildtype strain has very low transmissibility, the third has pandemic potential, and the second strain has intermediate transmissibility. This intermediate transmissibility is not enough to sustain the novel pathogen within the human host population, but secondary infections by humans are possible. Thus we have $\alpha_1 < \alpha_2 < 1$. Finally, the human adapted strain has a reproductive number $\alpha_3 > 1$. Further, we assume an identical mean infectious period for all strains.

Following [11], we first distinguish two routes to adaptation: the 'punctuated' scenario has an evolutionary course $\alpha_i = [0, 0.1, 2]$, while the 'gradual' scenario has $\alpha_i = [0, 0.9, 2]$, the only difference being α_2, the fitness of the intermediate stage.

This leads to the following SIR-model, normalized with respect to the mean infectious period,

$$\begin{aligned} \frac{dS}{dt} &= -\sum_{i=1}^{3} \alpha_i I_i \frac{S}{N} \\ \frac{dI_i}{dt} &= \alpha_i I_i \frac{S}{N} + \mu_{i-1} I_{i-1} - (\mu_i + 1) I_i \\ \frac{dR}{dt} &= \sum_{i=1}^{3} I_i \end{aligned} \tag{1}$$

where S is the number susceptible, I_i is the number infected with strain i, and R the number of recovered or removed. We do not include births and deaths as we expect a zoonotic emergence, or extinction, to happen on a much shorter timescale than the human lifespan. We translate this model to a stochastic simulation of a multitype branching process, using the Gillespie algorithm [15]. The infection is seeded in a single random, susceptible host with the wildtype strain.

In general, every introduction has only two possible outcomes: emergence or extinction. Extinction happens if the novel pathogen dies out because it fails to adapt for human transmission or just by stochastic extinction. Hence, the introduction only leads to a limited number of infectious hosts, which we refer to as the 'outbreak size'.

Conversely, a novel pathogen of zoonotic origin emerges if it is sufficiently adapted for human transmission and begins to spread in a self-sustaining way. Formally, in an unlimited host population the cumulative number of infectious hosts is unbounded as time goes to infinity. Computationally we use a threshold of $I_3 = 100$ infectious hosts with the fully adapted strain to distinguish between emergence and extinction. This threshold ensures a probability of extinction less than $(1/R_3)^{I_3} = 7.9 * 10^{-31}$ [16]. Therefore, the number of falsely identified emergences, which would truly be extinctions, is negligibly small. Moreover, these arbitrarily small probabilities ensure that our simulation results are insensitive to the precise choice of threshold used. In situations where the host population is very small we relax our emergence threshold to smaller numbers of infectious hosts as some population sizes are below 100.

Probability of Emergence

In the special case $N \to \infty$ and $S/N \to 1$ it is possible to calculate the probability of emergence per introduction into the human host population, given the evolutionary course of the pathogen, and the mutation rate with which the pathogen adapts. Our derivations start with assuming one homogeneous human host population of infinite size. This assumption can be easily relaxed as we show later. To calculate the probability of emergence, we define next event probabilities of infection p_i, mutation m_i, and recovery r_i for each individual infected host, therefore the probabilities for what type of event will come next for each infectious host are

$$\begin{aligned} p_i &= \frac{\alpha_i}{\alpha_i + \mu_i + 1} \\ m_i &= \frac{\mu_i}{\alpha_i + \mu_i + 1} \\ r_i &= \frac{1}{\alpha_i + \mu_i + 1} \end{aligned} \tag{2}$$

Note that in general, we can extend this adaptation process to arbitrarily many adaptive steps. The mutations are uni-directional towards the adapted strain. Using a branching-process approach similar to [10], we derive the probability of emergence per introduction as follows (see Text S1, A.1.1, for more details)

$$p_{em} = 1 - \begin{cases} \sum_{x=1}^{n} r_x \prod_{y=2}^{x} \frac{m_{y-1}}{1 - p_y q_y} & \text{if } p_1 = \alpha_1 = 0 \\ \frac{1}{2p_1} - \sqrt{\frac{1}{4p_1^2} - \sum_{x=1}^{n} \frac{r_x}{p_1} \prod_{y=2}^{x} \frac{m_{y-1}}{1 - p_y q_y}} & \text{otherwise} \end{cases} \tag{3}$$

The probability of emergence can be expressed by the next event probabilities and the probability of extinction q_i given an index infection of strain i. This expression can be solved analytically for all possible routes of adaptation.

Waiting Time to Emergence

Regardless of the underlying population structure and the pathogen's biology, we can make an estimate of the number of introductions needed before an emergence arises M, given the probability of emergence per introduction p_{em} with $0 \leq p_{em} \leq 1$ (see Text S1, A.1.3, for more details)

$$\langle M \rangle = \frac{1}{p_{em}} - 1 \tag{4}$$

Note that this is the average number of introductions without an emergence. The average number of introductions needed for an emergence event is $\langle M \rangle + 1$.

In addition, the variance can be obtained in a similar way (see Text S1, A.1.3)

$$\mathrm{var}(M) = \frac{1}{p_{em}^2} - \frac{1}{p_{em}} \tag{5}$$

This variance leads to a standard deviation of the same order as the average number of introductions, $< M >$. This makes the number of introductions before an emergence inherently unpredictable if the probability of emergence per introduction is small ($p_{em} \ll 1$).

Outbreak Size Distribution

Again, in the special case $S/N \rightarrow 1$, the branching-process approach can be extended to derive the probabilities of outbreak sizes before emergence (see Text S1, A.1.2, for more details). In general, the probability of an outbreak of size x_i is defined as $g_i(I_{i,0}, x_i)$ with i denoting the strain. The number of infected hosts to start with is denoted by $I_{i,0}$. Furthermore, the overall outbreak size probability can be derived using conditional probabilities

$$p_{out}(X) = g_1(I_{1,0}, x_1) \sum_{I_{2,0},\ldots,I_{n,0}=0}^{I_{1,0}+x_1,\ldots,I_{n-1,0}+x_{n-1}} \prod_{i=2}^{n} \zeta_i(I_{i,0}) g_i(I_{i,0}, x_i) \tag{6}$$

where $X = I_{1,0} + \sum_{i=1}^{n} x_i$ is the total outbreak size, and $\zeta_i(I_{i,0})$ is the probability of getting $I_{i,0}$ patient zeros to start with in strain i. The summation in the derivation of the overall outbreak size probability is over all possible subsets of infectious hosts to start with.

Incorporating Spatial Heterogeneity

We use a metapopulation model to explore the effects of spatial host heterogeneity, effectively dividing the human population into interconnected communities. As an example of the general case of spatially structured communities, we focus on a simple village - city model to approximate the spatial host heterogeneity in rural areas connected, by human mobility, to bigger cities. There are many different types of human mobility between communities such as villages and cities, including short-term commuting and long-term labour migration. Particularly in developing countries, however, information on dominant patterns is sparse. Nonetheless, anecdotal evidence from Vietnam [17], for example, suggests that short term commuting plays an important role: here, a subset of the village residents collects agricultural produce for trading in local markets in the city, and travels to the city on a daily to weekly basis. Accordingly we present a model in which the residence time of villagers in the city is typically less than the infectious period. However, in the supplementary information we also present a model incorporating migration on longer timescales (see Text S1, A.2). These two different models illustrate that our results appear qualitatively robust to different types of human movement.

As before, we have a wildtype pathogen capable of acquiring adaptations for human transmission. Assume a finite number of hosts in the village, and an effectively infinite number in the city. To allow for daily commuting, we label individuals in the city according to whether they are commuters from the village or not (neglecting commuters originating from the city and present in the village). The superscripts (v),(c) represent village and city inhabitants respectively, while (vm) denotes villagers in the city. Village members commute to the city at a per capita rate χ_1, and return at per capita rate χ_2. At any one time, a proportion w of villagers, the commuters, are in the city with w being set by χ_1 and χ_2 as described below. Further, we neglect susceptible village commuters acquiring infection in the city (this arises formally from the infinite number of hosts in the city). The number of village residents is fixed at $N = N^{(v)} + N^{(vm)} = 1000$, and we define the average number of commuters as

$$< c > = N^{(vm)} = N * w \tag{7}$$

Figure 1 shows a schematic representation of this commuting model. Normalising time with respect to the mean infectious period, the governing equations are

Village, non-commuting :

$$\begin{aligned} \frac{dS^{(v)}}{dt} &= -\sum_{i=1}^{3} \alpha_i I_i^{(v)} \frac{S^{(v)}}{N^{(v)}} - \chi_1 S^{(v)} + \chi_2 S^{(vm)} \\ \frac{dI_i^{(v)}}{dt} &= \alpha_i I_i^{(v)} \frac{S^{(v)}}{N^{(v)}} + \mu_{i-1} I_{i-1}^{(v)} \\ &\quad -(\chi_1 + \mu_i + 1) I_i^{(v)} + \chi_2 I_i^{(vm)} \end{aligned} \tag{8}$$

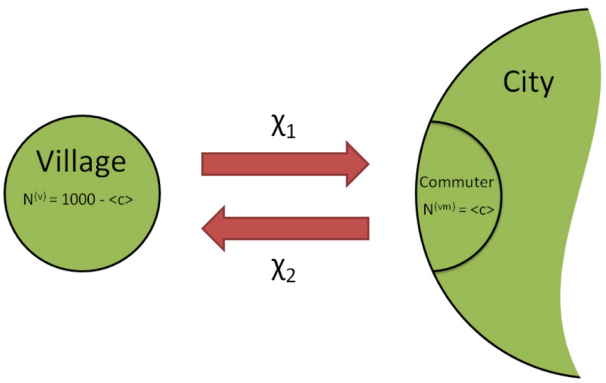

Figure 1. Schematic representation of the short-term commuting model. $N^{(v)}$ is the number of residents present in the village and $N^{(vm)} = < c >$ the number of commuters in the city. The city has an infinite number of residents. χ_1 is the per capita commuting rate from the village to the city and χ_2 the return rate. Together, both determine the number of commuters $< c > = N * w = N(\chi_2/\chi_1 + 1)^{-1}$.
doi:10.1371/journal.pcbi.1000947.g001

Village, commuting :

$$\begin{aligned}\frac{dS^{(vm)}}{dt} &= \chi_1 S^{(v)} - \chi_2 S^{(vm)} \\ \frac{dI_i^{(vm)}}{dt} &= \chi_1 I_i^{(v)} - \chi_2 I_i^{(vm)} \\ &\quad + \mu_{i-1} I_{i-1}^{(vm)} - (\mu_i + 1) I_i^{(vm)}\end{aligned} \tag{9}$$

City :

$$\frac{dI_i^{(c)}}{dt} = \alpha_i \left(I_i^{(c)} + I_i^{(vm)} \right) + \mu_{i-1} I_{i-1}^{(c)} - (\mu_i + 1) I_i^{(c)} \tag{10}$$

Commuting equilibrium :

$$\begin{aligned}\frac{N^{(vm)}}{N^{(v)}} &= \frac{\chi_1}{\chi_2} = \frac{w}{1-w} \\ \text{with } N &= N^{(vm)} + N^{(v)} = \text{const.}\end{aligned} \tag{11}$$

Note that equation (11) arises from the fact that $w = N^{(vm)}/(N^{(v)} + N^{(vm)})$. To represent daily commuting, with an infectious period of 5 days, we set $\chi_2 = 5$, and choose w to give the required average number of commuters $< c >$. To seed a wildtype infection in the village we set $I_1^{(v)} = 1$. In the village-city model, an emergence event is defined as having 100 infectious hosts with the fully adapted strain in the city.

Results

Validation of Analytical Results

We use the three strain model described before to study the impact of the mutation rate, μ, and the average reproductive number of the intermediate strain, α_2, on the probability of emergence per introduction in a single, infinite population. We assume $0 < \mu \leq 10^{-1}$ as an illustrative spectrum of possible mutation rates. Figure 2 shows the probability of emergence for different mutation rates and average reproductive numbers of the intermediate strain. Not surprisingly, the probability of emergence grows non-linearly with α_2 and μ. The probability of having no mutation in the second strain is $(1+\mu)^{-x}$ where x is the number of infected hosts with strain 2. While the intermediate reproductive number affects the exponent, the mutation rate has a direct influence on the base. We validate our analytical results by comparing them with the average probability of emergence of 10^3 simulated emergence processes, using one homogeneous population as described in (1). Figure 2 reveals an excellent agreement between the analytical results and simulations.

Effect of Limited Host Population Size

For small host communities, the depletion of susceptible hosts can play a significant role in limiting an ongoing outbreak. What is the effect of a finite population size on these analytical results which assume an infinite host population? Figure 3A compares the simulated outbreak size distribution of different sized populations with our analytical predictions. Note that, for populations greater than 500, there is close agreement between numerical and analytical results. When considering populations of size 1,000 or more, we do not expect population size dependence to have a substantial effect.

Effect of Spatial Heterogeneity

The village's number of residents in our commuting model is sufficient to avoid finite size effects on the outbreak size. Furthermore, it is independent of any spatial heterogeneity. The number of residents only has a limiting effect on the outbreak size distribution. Figure 3B shows the outbreak size distribution for our

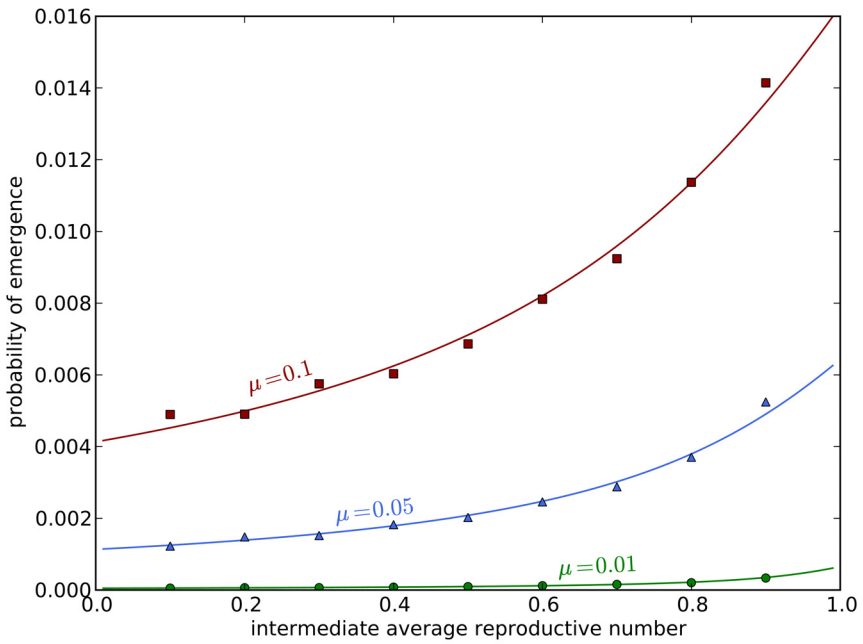

Figure 2. Comparison of the probability of emergence per introduction. Shown is the probability of emergence as a function of the mutation rate μ and the intermediate reproductive number α_2. The average reproductive number for each strain is $\alpha_i = [0, \alpha_2, 2]$ and the mutation rate $\mu_i = [\mu, \mu, 0]$, $i \in \{1, 2, 3\}$. The solid lines are analytical results. The data points represent the average probability of emergence over 1000 simulated emergences with a host population size of $N = 1000$. The agreement between analytical calculations and simulations is excellent. Further, the probability of emergence grows non-linearly with α_2 and μ.
doi:10.1371/journal.pcbi.1000947.g002

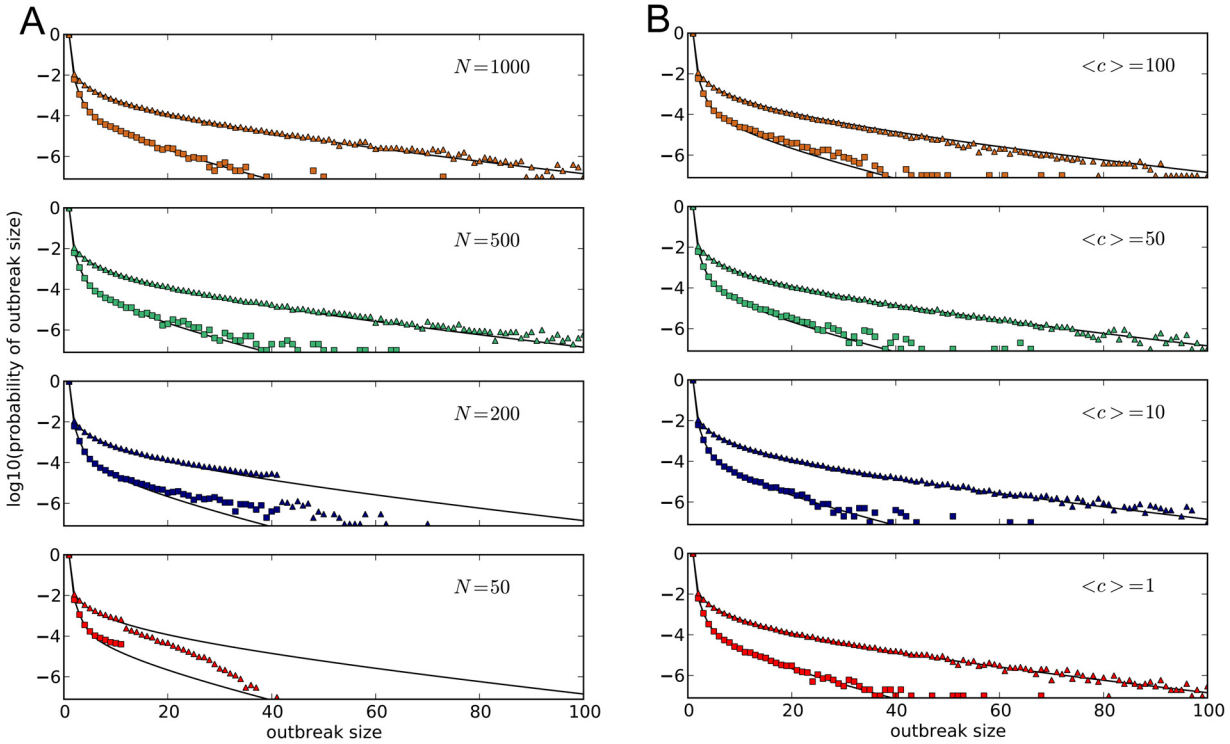

Figure 3. Outbreak size distribution. The solid black line represents the analytical result for an infinite population. Data points are the average probability of 10^7 simulations. Squares represent the punctuated route (I), and triangles the gradual route (II), both with the mutation probability $\mu = 0.1$. **A** Outbreak size distribution as a function of host population size. For a population size, N, of 50, 200 and 500, an emergence event is defined as having at least 10% of the population size infected with the fully adapted strain, while the number was fixed at 100 infected with the fully adapted strain for $N = 1000$. While host populations with $N < 500$ clearly show finite size effects, even small host populations with $N \approx 500$ can be effectively treated as infinite as the outbreak size is small compared to the population size. The figure reveals the effect of the pathogen's evolution on the outbreak size distribution as the distributions group according to the route of adaptation. **B** Probability of emergence in the city with short-term commuting. The probability is a function of the overall number of infectious hosts in the village-city model with $N = 1000$. The red color represents $<c> = 1$, blue $<c> = 10$, green $<c> = 50$ and gold $<c> = 100$ commuters. As for homogeneous populations, the outbreak size distributions group according to the evolutionary route of adaptation. Spatial heterogeneity does not have an influence as all simulations do not show a significant variation from the analytical results.
doi:10.1371/journal.pcbi.1000947.g003

short-term commuting model. The average number of commuters ranges from 1 to 100. As we expect, no significant effect on the outbreak size distribution can be seen, even for $<c> = 1$. It validates our assumption of the independence of infectious hosts, necessary for a branching-process formulation, as the simulations closely match the analytical predictions. It is noteworthy here that only the biology of the novel pathogen determines the emergence process, as outbreak sizes group according to the intermediate average reproductive number α_2. The minimal deviation for $<c> = 100$ commuters is based on the fact that the effective village size is only $N^{(v)} = 900$ due to the absent commuters.

In Figure 4, we extract the probability of emergence per introduction given a certain number already infected. These data are easily calculated using the outbreak size distribution and the probability of emergence per introduction. Assume an introduction has caused x infectious hosts already. The probability of extinction is the cumulative probability of getting an outbreak size equal to or larger than x, renormalized by all possible outcomes (extinctions and emergences) once x hosts are infectious. This yields the probability of emergence per number of infected. In Figure 4, the effect of spatial heterogeneity can be seen directly. For $<c> \geq 10$, the village-city simulations agree very well with the analytical results assuming a single, infinite population. But for $<c> = 1$, the probability of emergence converges to approximately 0.42.

While Figure 4 shows the probability of emergence as a function of the number infected, the actual outcome is highly unpredictable even if the probability of an event is known as the average waiting time to an emergence shows (see equations 4 and 5). It can be generalized for the probability of emergence given x infectious hosts. For example, the probability of emergence given five infectious hosts in the gradual route (II) is $p_{em}(x = 5) = 0.567$. It follows on average every 1.736 times this happens an emergence will happen. The standard deviation is ± 1.161, which leads to the conclusion that even if the probability is known, it is inherently unpredictable when this will actually lead to an emergence.

Spatial Homogeneity Coefficient

This confirms that a pathogen needs a sufficient connection between communities to emerge, despite its ability to cause outbreaks, regardless of the spatial structure. Hence we expect the existence of a threshold where spatial heterogeneity effectively does not matter any more. Previous research has shown that the effect of heterogeneity in spatially structured population models depends on the interconnectivity with a threshold effectively allowing the pathogen to spread between communities [18–20]. Our approach allows new insights, as we do not need to specify the actual number of infectious hosts migrating to a new community. We measure connectivity between communities in terms of the average number of commuters $<c>$ for which rich empirical

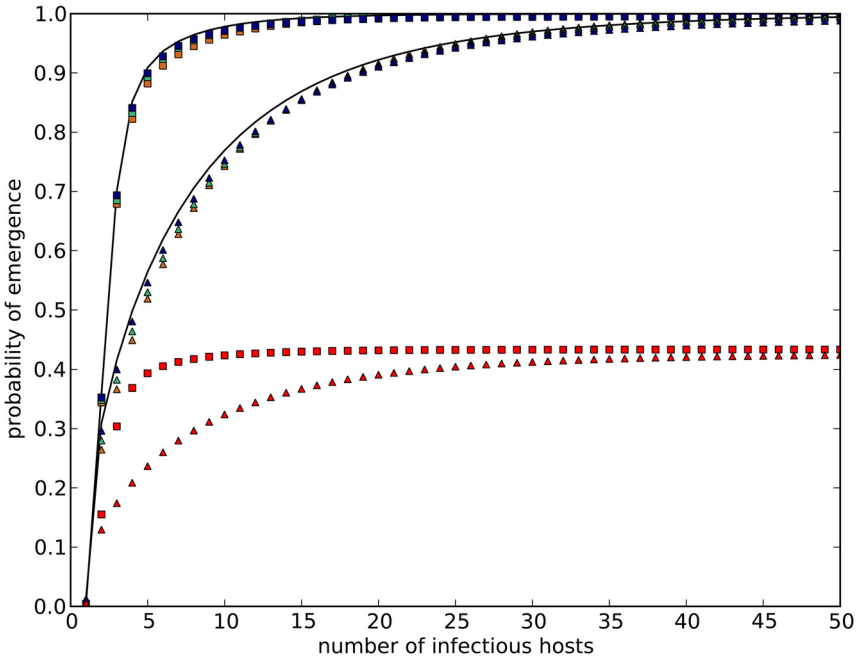

Figure 4. Probability of emergence in the city with short-term commuting. The probability is a function of the overall number of infectious hosts in the village-city model. The solid black lines represent the analytical results for an infinite population without spatial heterogeneous structure. Data points are the average probability of 10^7 simulations. Squares represent the punctuated route (I), and triangles the gradual route (II), both with the mutation probability $\mu = 0.1$. As with the outbreak size distribution, probabilities group according to the route of adaptation instead of connection strength in number of commuters. However, the probability of emergence does not converge to 1 for $<c> = 1$ commuters, revealing the effect of spatial heterogeneity when the number of commuters is small relative to the average reproductive number for the well-adapted strain.
doi:10.1371/journal.pcbi.1000947.g004

datasets can be found. Figure 5 presents illustrative examples of empirical data of commuting patterns in different parts of the world. Most data has been collected by Offices of Statistics of five countries on three continents [21–25]. A further, two independent studies have been used to estimate commuting patterns of towns in Indonesia [26] and China [27].

We next attempt to quantify the regimes in $<c>$ for which spatial heterogeneity may be neglected. We approach this question using a simple analytical derivation for the effect of spatial heterogeneity, which considers only the adapted strain. Assume a connected community such as the village in our village-city model, with a fully adapted pathogen introduced into the village. Given an emergence and epidemic in the village, the probability that this causes an emergence and epidemic in the city is

$$H_{spatial} = 1 - \left(\frac{1}{R_0}\right)^{<c> f} \tag{12}$$

Therefore, $H_{spatial}$ is a spatial homogeneity coefficient, measuring the impact of spatial heterogeneity on an emergence process. It ranges from 0, leading to two isolated communities, to 1, effectively removing any spatial heterogeneity and forming one homogeneous population. f is the fraction of the village residents becoming infectious. It can be derived using [28]

$$f = 1 - e^{-R_0 f} \tag{13}$$

The spatial homogeneity coefficient depends only on the connection strength expressed in commuters $<c>$ and the average reproductive number $R_0 = \alpha_3$ of the fully adapted strain. Though it only considers the fully adapted strain, we expect this coefficient to be a good approximation for a multi-strain model as the vast majority of infectious hosts will transmit the fully adapted strain in the case of an emergence.

Figure 6 gives an overview of the influence of spatial heterogeneity as a function of $<c>$. Effectively, spatial heterogeneity is negligible once a critical number of ten commuters connect the two communities. This is a very low threshold, and empirical data shows that the probability of having a community with less than the critical number of ten commuters is approximately 1% for all our data combined.

As illustrated by the close fit between analytical and numerical results in figure 6, there is only a small error in the analytical expression arising from neglecting infections with mal-adapted strains. This error is greatest for the gradually adapting pathogen, because an intermediate strain with $\alpha_2 = 0.9$ tends to cause larger outbreaks than one with $\alpha_2 = 0.1$. Nevertheless, the deviation remains small.

In light of this agreement, how does the critical average number of commuters vary with the adapted reproductive number? While for $R_0 = 2$ ten average commuters are sufficient to dissolve spatial heterogeneity, this changes dramatically for smaller average reproductive numbers (see Figure 7). If a well-adapted strain is only just pandemic capable (i.e. R_0 just above 1) villages with only ten commuters are only 50% likely to seed an epidemic in their local city, and spatial structure becomes important again. For example, the critical number of commuters is close to 100 for $R_0 = 1.2$. For $R_0 = 2$ the spatial homogeneity coefficient is approximately 0.42 for one commuter. This agrees with what we find in Figure 4 using simulated results.

Discussion

In this article we first present analytical results to calculate epidemiological parameters of a novel disease, adapting to

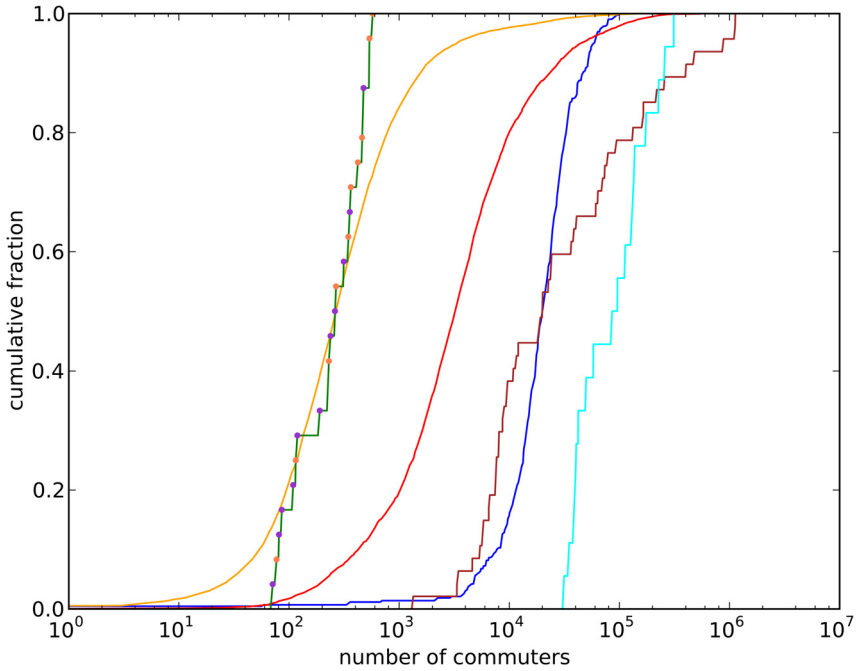

Figure 5. Data of commuting patterns in different parts of the world. Shown is the cumulative fraction of all communities with equal or less than the specified number of commuters. The data was mostly collected by the National Statistical Offices of the respective countries. The gold line represents commuting data from Brazil [21], the red line data from the USA [22], the blue line data from the UK [23], the brown line data from Japan [25], the cyan line data from Hong Kong [24], and the green line data from two independent sources. The green line has orange [27] and pink data points [26], corresponding to its data sources. Our data represents the commuting flows between administrative units. The definition of administrative units varies highly between countries. For example, the US data is on a granularity of 3,141 counties, while the data from Japan is based on its 47 prefectures. However, heterogeneity can also be found within countries datasets. The Brazilian data is on a level of 5,507 municipalities with resident sizes ranging from 1,166 to 10, 435, 548.
doi:10.1371/journal.pcbi.1000947.g005

humans. We explore the influence of spatial host contact structure, and validate our result with stochastic simulations of simple village-city models as an example of interconnected communities within a spatially structured population. Our study reveals that for plausible parameter ranges, spatial heterogeneity only has very limited impact on the probability of emergence, as well as the outbreak size distribution. Neither a change in strength of spatial heterogeneity (e.g. number of commuters), nor in its quality (e.g. short term versus long term commuting) shows a significant influence. Our results suggest that only the most remote rural communities would be subject to epidemiological isolation. In particular, the available empirical data suggests that communities tend to be highly interconnected with relatively high connection strengths. Of course, it is the most remote communities of the world for whom we have the least relevant data. More empirical research on spatial heterogeneity is needed to form a better understanding of its effect, and this need is greatest in developing parts of the world.

In addition, population size becomes an important factor only when that population is relatively small — fewer than $N \approx 500$ individuals. Only a small number of infectious hosts are actually involved in the emergence process, which relates to the small reservoir of susceptibles needed for a successful emergence. Moreover, biological processes such as the speed of evolution and the adaptive route show a strong influence on the overall emergence process. We show that epidemiological parameters such as the outbreak size group according to the evolutionary route. Previous research has shown the effect of the pathogen's route of evolutionary adaptation and mutation rate on the probability of emergence per introduction [10,11]. Our theoretical derivation of the probability of emergence extends this and offers the benefit of being analytically solvable for any possible route of adaptation and any mutation probabilities. We note here that previous work has highlighted the role of other, significant types of heterogeneity in emergence of a novel infection. For example, [29] describe the effect of the pathogens life history, such as the length of infection, on the emergence of a novel pathogen. Further, heterogeneity in human-to-human transmission within a population may have an influence on the course and probability of emergence and outbreaks [13,14], usually lowering the probability of emergence. While we have concentrated here on simple types of spatial heterogeneity, a significant question for future research is the role of mixed heterogeneities, for example spatially structured populations with additional heterogeneity in the human-to-human transmission.

We also find that the waiting time for an outcome of a novel pathogen's introduction is highly unpredictable, even if the probability for such an event is known. Conversely, this means that an estimate of the underlying epidemiological parameters from observed data will be highly uncertain. Unfortunately, a large number of observations will be necessary to achieve confidence in the parameters, and even a large number of introductions gone extinct do not rule out the possibility of emergence for a pathogen. We came to a similar result [11] using a measurement on the upper bound of the probability of emergence. Moreover, the probability of emergence given a certain number of infectious hosts can be surprisingly low. Even a comparatively large number of infectious hosts can end in extinction, especially for low mutation rates and intermediate-stage average reproductive numbers just below one.

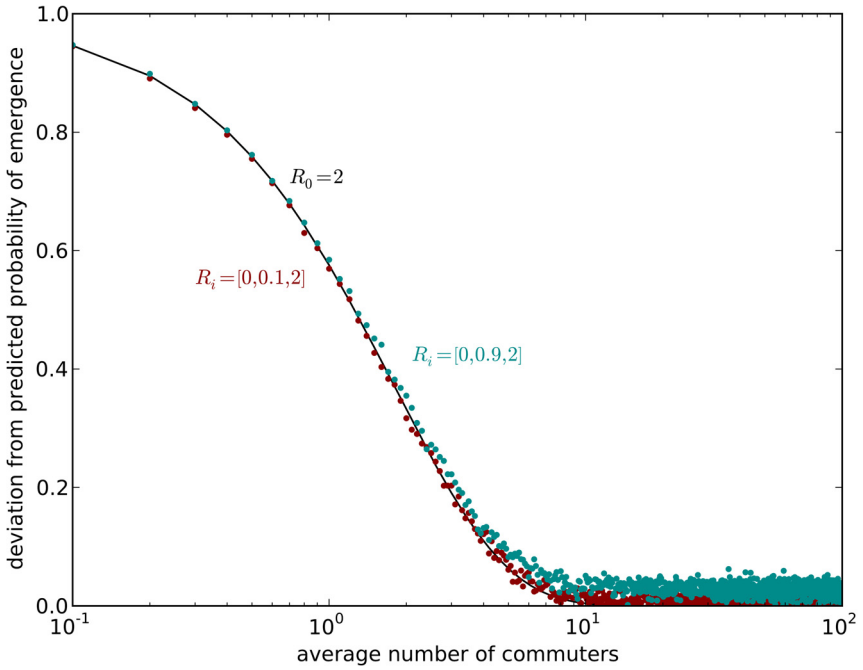

Figure 6. Deviation between simulated and analytical predicted probability of emergence. The deviation is a function of the average number of commuters $<c>$. The deviation is defined as $|p_{em}^{(ana)}-p_{em}^{(sim)}|/p_{em}^{(ana)}$. The analytical probability of emergence $p_{em}^{(ana)}$ is for an infinite ulation without spatial structure. The simulated probability of emergence $p_{em}^{(sim)}$ is for short-term commuting with the blue data points representing the gradual route (II) of adaptation, and the orange data points representing the punctuated route (I). The solid black line is $1-H_{spatial}(1/\alpha_3)^{<c>f}$, as defined in equation 12 in the main text. It is the analytical expected deviation for spatial heterogeneity as a function of the spatial homogeneity coefficient. The simulations agree very well with the analytical expected deviation. The gradual route (II) is slightly more off from the theoretical prediction as a result of the small but significant number of infected with the intermediate strain. Effectively, it lowers the number of commuters infected with the fully adapted strain and therefore the probability of transmission from the village to the city. Nevertheless, the analytical prediction as well as the simulations show no significant impact of spatial heterogeneity from a critical commuter threshold of $<c>=10$.
doi:10.1371/journal.pcbi.1000947.g006

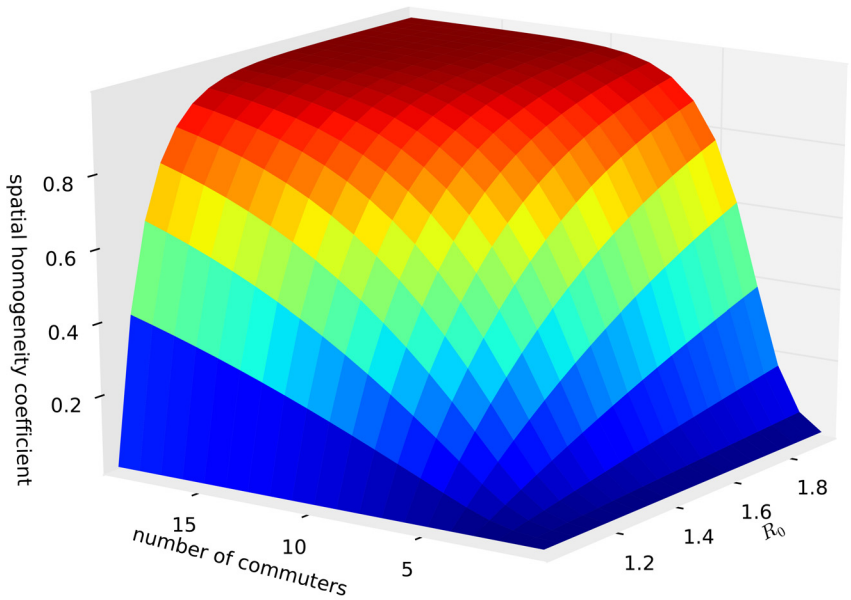

Figure 7. Impact of spatial heterogeneity on disease transmission between communities (I). The impact is measured with the spatial homogeneity coefficient with $0\leq H_{spatial}\leq 1$. Given $H_{spatial}=1$ every emergence in the village automatically leads to an emergence in the city, and $H_{spatial}=0$ represents no chance of successfully transmitting the pathogen into the city. The figure reveals that spatial structure becomes especially important for small average reproductive numbers. In addition, the average number of commuters needed to show an effect of spatial heterogeneity is surprisingly small.
doi:10.1371/journal.pcbi.1000947.g007

Our work has relevance for important public health issues: if a novel disease is detected in a rural setting, and it appears to be spreading, how feasible is it to contain infection by restricting movements to and from the village? Our results suggest that first, an infeasibly tight level of quarantine would be required for any chance of containment, corresponding to enforcing a low level of $<c>$ in Figure 6. To all intents and purposes isolation would have to be absolute to be effective. In most circumstances such extreme intervention would not be acceptable. Second, given typical mobility patterns, it is likely that once there is a detectable number of cases in the village, there may already be a significant number of cases outside of it. Therefore quarantining interventions are likely to come too late.

Our work raises important questions for future research: where should surveillance be focused to detect an emergence as early as possible, especially if resources are limited? Given emergence of a novel infection in a rural setting, how much time can we buy through limiting travel to and from major urban centres? These and other questions will undoubtedly benefit from more systematic studies of emergence in the context of population distributions. Nonetheless, theoretical models such as those presented here can offer useful, fundamental insights to guide such studies.

Supporting Information

Text S1 Supplementary material with figures for 'Insights into the Evolution and Emergence of a Novel Infectious Disease'.
Found at: doi:10.1371/journal.pcbi.1000947.s001 (0.16 MB PDF)

Author Contributions

Conceived and designed the experiments: RJK NA ARM. Performed the experiments: RJK. Analyzed the data: RJK. Contributed reagents/materials/analysis tools: RJK. Wrote the paper: RJK NA ARM.

References

1. Reid AH, Fanning TG, Hultin JV, Taubenberger JK (1999) Origin and evolution of the 1918 Spanish influenza virus hemagglutinin gene. P Natl Acad Sci USA 96: 1651–1656.
2. Horimoto T, Kawaoka Y (2005) Influenza: lessons from past pandemics, warnings from current incidents. Nat Rev Micro 3: 1740–1526.
3. Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team (2009) Emergence of a novel swine-origin influenza A (H1N1) virus in humans. New Engl J Med 360: 2605–2615.
4. Brockmann D, Hufnagel L, Geisel T (2005) SARS: a case study in emerging infections. Oxford: Oxford University Press, 3rd edition.
5. Meyers LA, Pourbohloul B, Newman MEJ, Skowronski DM, Brunham RC (2005) Network theory and SARS: predicting outbreak diversity. J Theor Bio 232: 71–81.
6. Skowronski DA, Astell C, Brunham RC, Low DE, Petric M, et al. (2005) Severe acute respiratory syndrome (SARS): A year in review. Annu Rev Med 56: 357–381.
7. Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, et al. (2003) Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China. Science 302: 276–278.
8. Li WD, Shi ZL, Yu M, Ren WZ, Smith C, et al. (2005) Bats are natural reservoirs of SARS-like coronaviruses. Science 310: 676–679.
9. Briese T, Paweska JT, McMullan LK, Hutchison SK, Street C, et al. (2009) Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from Southern Africa. PLoS Pathog 5: 8.
10. Antia R, Regoes RR, Koella JC, Bergstrom CT (2003) The role of evolution in the emergence of infectious diseases. Nature 426: 658–661.
11. Arinaminpathy N, McLean AR (2009) Evolution and emergence of novel human infections. Proc R Soc B 276: 3937–3943.
12. Woolhouse ME, Haydon DT, Antia R (2005) Emerging pathogens: the epidemiology and evolution of species jumps. Trends Ecol Evol 20: 238–244.
13. Yates A, Antia R, Regoes RR (2006) How do pathogen evolution and host heterogeneity interact in disease emergence? Proc R Soc B 273: 3075–3083.
14. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM (2005) Superspreading and the effect of individual variation on disease emergence. Nature 438: 355–359.
15. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81: 2340–2361.
16. May RM, Gupta S, McLean AR (2001) Infectious disease dynamics: what characterises a successful invader? Philos T R Soc Lon B 356: 901–910.
17. Thanh HX, Anh DN, Tacoli C (2005) Livelihood diversification and rural-urban linkages in Vietnam's Red River Delta. IIED Working Paper Series on Rural-Urban Interactions and Livelihood Strategies.
18. Watts DJ, Muhamad R, Medina DC, Dodds PS (2005) Multiscale, resurgent epidemics in a hierarchical metapopulation model. P Natl Acad Sci USA 102: 11157–11162.
19. Vazquez A (2007) Epidemic outbreaks on structured populations. J Theor Bio 245: 125–129.
20. Colizza V, Vespignani A (2008) Epidemic modeling in metapopulation systems with heterogeneous coupling patter: theory and simulations. J Theor Bio 251: 450–467.
21. Instituto Brasileiro de Geografia e Estatística, Brazil (2000) Censos demográficos 2000.
22. US Census Bureau, United States of America (2000) Census 2000.
23. Office for National Statistics, United Kingdom (2001) Census 2001.
24. Census and Statistics Department, Hong Kong (2001) 2001 population census.
25. Statistics Bureau, Japan (2005) 2005 population census.
26. Leinbach TR (1983) Rural transport and population mobility in Indonesia. J Dev Areas 14: 349–364.
27. Xu H (2001) Commuting town workers: the case of Quinshan, China. Habitat Int 25: 35–47.
28. Anderson RM, May RM (1991) Infectious diseases of humans: dynamics and control. Oxford: Oxford University Press, 1st edition.
29. André JB, Day T (2005) The effect of disease life history on the evolutionary emergence of novel pathogens. Proc R Soc B 272: 1949–1956.

OPEN ACCESS Freely available online

PLoS one

Why Was the 2009 Influenza Pandemic in England So Small?

Ruben J. Kubiak*, Angela R. McLean

Institute for Emerging Infections, Oxford Martin School, Department of Zoology, University of Oxford, Oxford, United Kingdom

Abstract

The "Swine flu" pandemic of 2009 caused world-wide infections and deaths. Early efforts to understand its rate of spread were used to predict the probable future number of cases, but by the end of 2009 it was clear that these predictions had substantially overestimated the pandemic's eventual impact. In England, the Health Protection Agency made announcements of the number of cases of disease, which turned out to be surprisingly low for an influenza pandemic. The agency also carried out a serological survey half-way through the English epidemic. In this study, we use a mathematical model to reconcile early estimates of the rate of spread of infection, weekly data on the number of cases in the 2009 epidemic in England and the serological status of the English population at the end of the first pandemic wave. Our results reveal that if there are around 19 infections (i.e., seroconverters) for every reported case then the three data-sets are entirely consistent with each other. We go on to discuss when in the epidemic such a high ratio of seroconverters to cases of disease might have been detected, either through patterns in the case reports or through even earlier serological surveys.

Citation: Kubiak RJ, McLean AR (2012) Why Was the 2009 Influenza Pandemic in England So Small? PLoS ONE 7(2): e30223. doi:10.1371/journal.pone.0030223

Editor: Alessandro Vespignani, Northeastern University, United States of America

Received September 28, 2011; **Accepted** December 14, 2011; **Published** February 10, 2012

Funding: RJK gratefully acknowledges funding for this work from the Oxford Martin School and the Engineering and Physical Sciences Research Council. ARM is a Senior Research Fellow of All Souls College whose support is gratefully acknowledged. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ruben.kubiak@zoo.ox.ac.uk

Introduction

In the spring of 2009 an influenza pandemic arose in North America. The first 'Swine flu' influenza A(H1N1) cases were recorded in March 2009 in Mexico [1]. Infection soon became a global phenomenon: the first laboratory confirmed cases in the UK were on 27 April [2], and by the end of May more than 50 countries had confirmed cases. The World Health Organisation formally declared the outbreak a pandemic on 11 June 2009 [3].

As the pandemic spread around the world it was natural to ask "how bad will it be?", and attempts to forecast epidemic impacts were rapidly assembled, broadcast and then published. Several of these studies focused on estimating the new influenza's basic reproductive number R_0, defined as the number of secondary cases caused by each case as an epidemic spreads into a population with no pre-existing immunity [4]. A related parameter, the effective reproductive number, R_{eff}, pertains when an epidemic spreads in a population in which some proportion are already immune [5].

Estimates of the basic reproductive number ranged from 1.2 to 2.1 (see Table 1). A mid-range estimate of 1.5 implied that the epidemic would grow until 1/3 of the population were infected or immune, then turnover and come to an end when around 60% of the population were immune. If all infections lead to illness prior to the acquisition of immunity this implied that a very large fraction of the population would become ill during the pandemic's early waves. Some alarming predictions were indeed made, particularly for the UK.

In England, the Health Protection Agency (HPA) monitored case estimates over the whole period of the pandemic, and published its findings online in its 'Weekly Pandemic Flu Media Updates' and its 'Weekly National Influenza Reports' [2,6]. The last pandemic flu media update was published on 14 January 2010, the official end of the pandemic in England [2,6]. By this time HPA's estimates totalled 910,000 cases over the course of the pandemic in England [2]. This amounts to less than 2% of the English population [7], a much lower figure than expected even with a basic reproductive number from the lower end of the published estimates.

This large discrepancy between the predicted and observed epidemics could have several causes. The basic reproductive number might have been very much smaller than estimated, a large part of the population might have been immune before the pandemic arrived, or many susceptible individuals might have been infected and become immune without being ill enough to register in the case estimates.

In what follows we use simple mathematical models to reproduce the HPA case estimates and try to learn how each of these three factors contributed to the surprisingly small epidemic. We then ask at what point in time during 2009 it should have become apparent that the English epidemic would be so small.

Methods

Phases of the Epidemic

We use the HPA case estimates which were made available to the public via the internet [2] as a data source for estimated numbers of symptomatic cases with influenza A(H1N1) in England. On 27 April 2009, the HPA announced the first laboratory confirmed infected case in the UK. It updated the

Table 1. Reproductive numbers estimates for influenza A(H1N1) from several independent studies.

Reproductive number estimates
1.22–1.58 (CI: 1.05–2.04) [1]
1.3–1.7 [17]
1.21–1.35 [18]
1.75 (CI: 1.64–1.88) [19]
1.37 (CI: 1.21–1.41)† [20]
1.6 (CI: 1.5–1.8)† [21]
1.78–2.07 (CI: 1.67–2.22) [22]
1.44 (CI: 1.36–1.51)† [23]
1.31 (CI: 1.25–1.38) [24]

A range of possible reproductive numbers is presented when multiple estimations were made. Estimates marked with † are the effective reproductive number R_{eff}, while unmarked ones are the basic reproductive number R_0. The confidence interval (CI) is given wherever possible.
doi:10.1371/journal.pone.0030223.t001

number of new laboratory confirmed cases in England on a daily basis until 2 July 2009 when it switched to weekly updates of estimated numbers of new cases in the previous week. The last pandemic media update was published on 14 January 2010, when the influenza epidemic in England was officially declared to be over [6].

HPA's case estimates were an estimate of the number of people with symptomatic infections. The number of cases was estimated as the product of four factors: the number of GP visits with influenza-like-illness, the proportion of diagnosed cases that were confirmed as infected with A(H1N1), the proportion of symptomatic cases likely to contact the health services and the likely impact of the National Pandemic Flu Service when it became available [8,9].

The data fall into three distinct phases:

Phase I: The first phase is the initial growth phase. It starts in week 19 with the first infected cases in the UK, and ends in week 30. From week 19 to week 27, the HPA released case reports based on laboratory confirmed cases. It changed its method to estimate cases using surveillance measurements in week 28 [2], explaining the jump in new cases between week 27 and 28. However, we expect that the laboratory confirmed case numbers released before week 28 are under-estimates of actual case numbers, because of limited laboratory facilities and the time delay to test for influenza A(H1N1).

Phase II: From week 30 to week 36 the number of cases declined. This is the period of school summer holidays in England. State schools were closed from 22 July to 3 September 2009 [10], exactly corresponding to phase II and marking the borders of the three phases. It is very likely that school closures during the summer break caused the decline in transmission events as children and young adults have been identified as the main age groups that were infected (see Figure 1 and [11,12]). Most universities were also shut during this phase, reducing mixing amongst young adults.

Phase III: The last phase is the second epidemic wave after the school holiday. Schools reopened in week 36, and the number of new infections increased until week 44 after which numbers of new cases decreased again. The third phase ends in week 1 of 2010 with no more published HPA updates on estimated A(H1N1) cases.

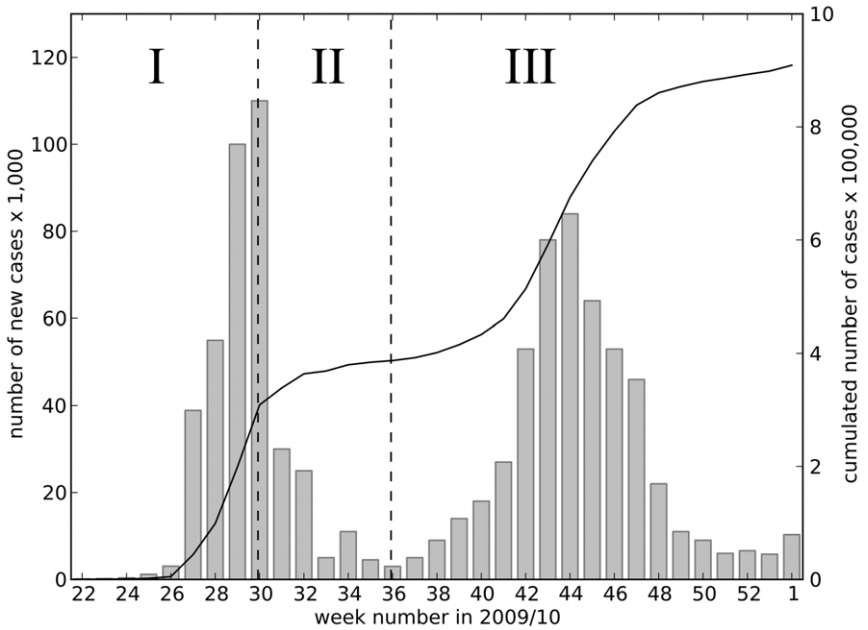

Figure 1. HPA case data and cumulative number of cases. The bars show the number of estimated new cases per week. The cumulative number of cases per week is presented as black line on a separate axis. The plot is separated into three phases, I–III, divided by dashed lines in week 30 and 36 between which state schools in England were closed [10]. Phase I is the initial growth phase before 22 July 2009, phase II is the school holiday phase until 3 September 2009, and at this date phase III with another growth and decline begins.
doi:10.1371/journal.pone.0030223.g001

Theoretical Model

We developed a simple version of the SIR mathematical model that allows for two routes to immunity; one with disease and one without. As we are interested in short-term epidemic behaviour it is a model for a closed population without any births. There are four types of individual: susceptibles, S, catch the infection at a rate that is proportional to the fraction of the population that is infectious. Upon infection they become infectious, either with symptoms, I_s, or without, I_a. The proportion that develop symptoms on infection is p. After a period of infectiousness they recover to their respective immune class, R_s or R_a. The rate of transmission of infection from those infected to the susceptibles is reduced by some fixed factor r_h during phase II, the school holidays. At the start of the epidemic in England most people are susceptible so in class S, but some may already be immune so are already in class $R = R_s + R_a$. This model has five parameters: the proportion initially immune, $R(t=0)/(R(t=0)+S(t=0))$, the transmission rate, β, the reduction in transmission during the holidays, r_h, the duration of infectiousness, $1/\gamma$, and the proportion of the population who become symptomatic if infected, p:

$$\begin{aligned}
\frac{dS}{dt} &= -r_h(t)\,\beta\,(I_s+I_a)\,\frac{S}{N}\\
\frac{dI_s}{dt} &= p\,r_h(t)\,\beta\,(I_s+I_a)\,\frac{S}{N}-\gamma\,I_s\\
\frac{dI_a}{dt} &= (1-p)\,r_h(t)\,\beta\,(I_s+I_a)\,\frac{S}{N}-\gamma\,I_a\\
\frac{dR_s}{dt} &= \gamma\,I_s\\
\frac{dR_a}{dt} &= \gamma\,I_a\\
N &= S+I_s+I_a+R_s+R_a
\end{aligned} \tag{1}$$

where the reduction in transmission is a time dependent function:

$$r_h(t)=\begin{cases} r_h & \text{if within school holiday}\\ 1 & \text{otherwise}\end{cases} \tag{2}$$

Furthermore, the basic reproductive number can be defined as $R_0=\beta/\gamma$, and the effective reproductive number is just $R_{eff}=R_0\,S/N$. In the special case $p=r_h=1$ our model reduces to the standard, textbook epidemiological model called the "SIR model" [4].

Serological Data

We used serological data based on samples drawn in England between July and September 2009. These were published in 2010 [11]. Table S1 shows the reported serological data and the demographic data that we used to calculate numbers immune before and after the first wave of the pandemic. The original report of the serological data split England into two areas, one with a high risk of seroconversion and one with a low risk. For our purposes we combine these using population data of the English population from the Office of National Statistics [7]. This yields estimates of the proportion of the population immune before the pandemic arrived and the proportion immune at the end of the first wave of the pandemic. The simple model that we use to understand these data is age-independent, and we therefore need to calculate proportions of the total population immune without age-stratification. To do that from age-stratified data we used demographic data on the current age and regional distribution [7].

Qualitative Analysis

Our Method proceeds in four steps:

First we calculate how many people were immune to A(H1N1) before the pandemic and in September 2009. The difference between these two numbers yields the total number of people who were infected during the English epidemic up to the time of sampling. This is a conservative estimate as people might have been infected between their sampling, which started in mid-July, and September 2009. Our method therefore underestimates the actual number of infected cases.

Second we compare the number of seroconverters with the cumulative number of estimated symptomatic cases by the beginning of September to get an estimate of the proportion of all infections that manifest themselves in disease (parameter p in the mathematical model). This parameter is calculated as a central value with upper and lower confidence bounds.

We then use these two pieces of information (proportion immune before the pandemic arrived and proportion of all infections manifesting as reported cases) and fit our model to the case report data. In fitting the model to the data we infer values for the three free model parameters: the basic reproductive number (R_0), the infectious period ($1/\gamma$) and the reduction in transmission during the summer holidays (r_h).

In a fourth step we use our model to ask if it could have been obvious from case-reports alone that the English epidemic would be so small, and if so, when. If we assume $S/N \approx const.$, the number of symptomatic infected can be expressed through an exponential growth:

$$I_s(t) \approx I_0\, e^{\left(p\,R_{eff}-\gamma\right)t} \tag{3}$$

The HPA data on symptomatic cases is based on weekly updates. This weekly sampling rate might be too broad to estimate the correct number of symptomatic cases at any time point, especially if the generation time is smaller than a week. But it allows us to evaluate the accumulated sum of symptomatic cases (new infections plus recovered ones), which grows with:

$$I_s(t)+R_s(t) \approx \frac{\beta\,I_0}{R_{eff}-\gamma}\,e^{\left(p\,R_{eff}-\gamma\right)t}+\text{const.} \tag{4}$$

where the last term expresses the normalisation constant. As before, the growth is described with an exponential function. We can determine the effect of the depletion of susceptible hosts, or in other words the difference in S/N between the two waves, by taking the natural logarithm of the accumulated sum of HPA's estimates. The slope of the logarithmic growth periods is ($p\,R_{eff}$ - γ), so differences in the slope of the two epidemic waves reflect the drop in R_{eff} caused by a depletion of susceptible hosts from one wave to the next.

Results

The English epidemic unfolded in three phases (Figure 1). The first phase of rapid growth was followed by a second phase when the number of cases fell as schools closed for the summer holiday. The third phase saw the start of the second wave as schools reopened in the autumn. After a short period of exponential

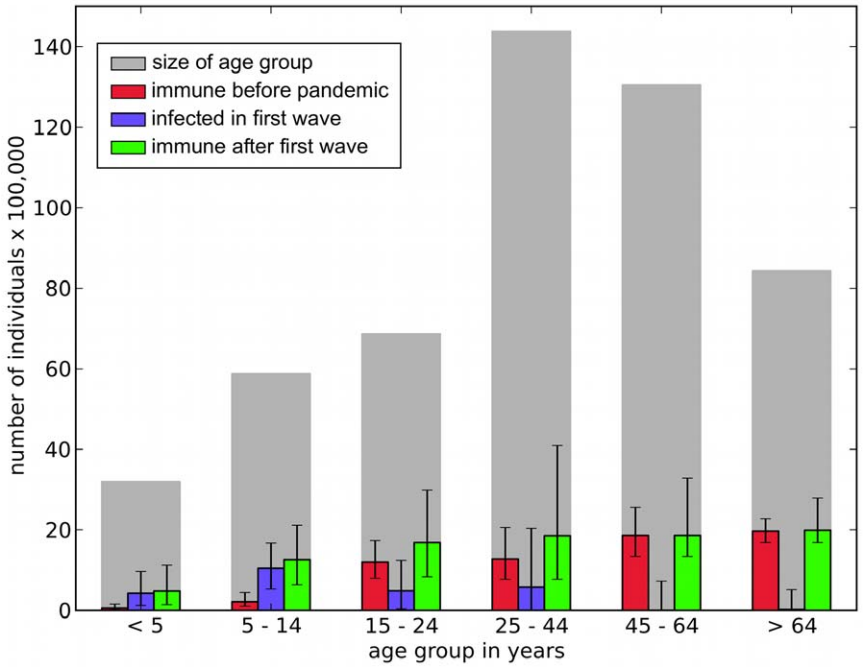

Figure 2. Comparison of serological data by age group. The number of individuals in each age group is shown with large, grey bars. Red bars show the number of individuals immune before the pandemic. Blue bars represent numbers who seroconverted between April and September 2009. Green bars show the number immune by early September 2009. All numbers are calculated using serological data and demographic data from England (see Table S1).
doi:10.1371/journal.pone.0030223.g002

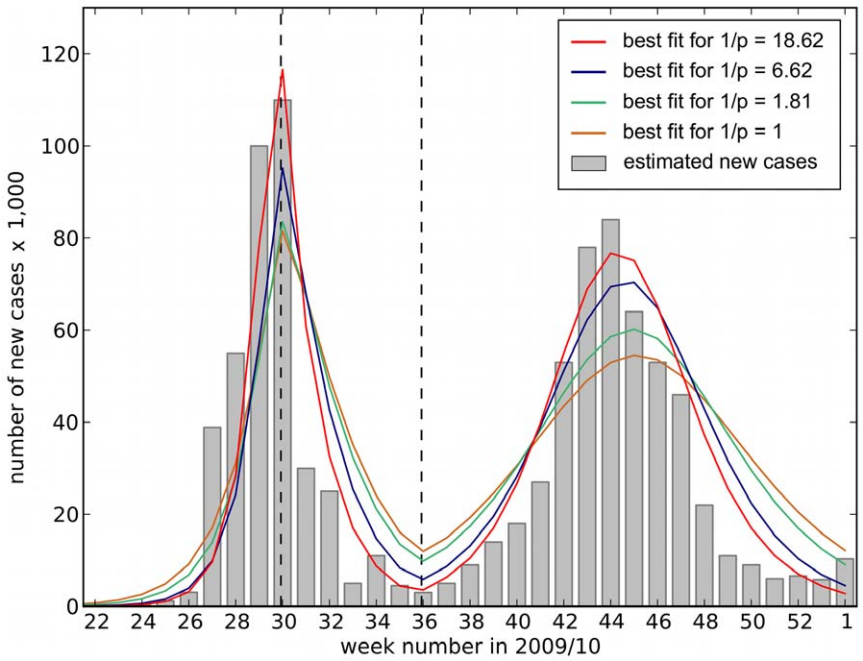

Figure 3. Comparison of HPA case estimates with theoretical predictions of new cases. The HPA case estimates per week are represented in a bar plot. The model in equation (1) was fitted to these data with state variable I_s representing reported cases. The initial conditions were set at the observed value of 12.7% of the population immune before week 22. Four different values of parameter p were assumed corresponding to four different fits: $1/p=1$ - brown line, $1/p=1.81$ - green line, $1/p=6.62$ - blue line and $1/p=18.62$ - red line. The figure illustrates that $1/p=18.62$ gives the best agreement between model and data with the smallest least square error, and allows our model to reproduce the case estimates. The fitting procedure yields estimates for the model's three parameters, shown in Table 2. Only when $1/p=18.62$ do these parameter estimates agree with previously published values, compare Table 2 with Tables 1, 3 and 4.
doi:10.1371/journal.pone.0030223.g003

Table 2. Estimated parameters using a least square fit.

Estimated parameters			
$1/p$	R_0	$1/\gamma$	r_h
1.00	1.16	0.18	0.98
1.81	1.18	0.28	0.97
6.62	1.27	0.82	0.89
18.62	1.59	2.42	0.65

All parameters were estimated assuming an immune fraction of 12.7% at the beginning of the pandemic. A least square fit was used to minimise the error between our model and HPA case estimates.
doi:10.1371/journal.pone.0030223.t002

growth, the second wave peaked in mid-November and then tailed off. This was surprisingly early, given the predicted attack rates.

These data raise four questions:

1. How transmissible was this strain of influenza amongst the population of England (i.e. how big was R_0)?
2. By how much did transmission decline during the summer holidays (i.e. what was r_h)?
3. For each case detected how many seroconverters were there (i.e. what was p)?
4. How soon could the failure to develop into a large epidemic have been recognised?

These questions can be addressed by combining the case report data with serological data (Figure 2). Those serological data show (red bars) that a substantial number of people were immune before the epidemic arrived and that this was particularly true for those over the age of 45. Correspondingly there were many seroconversions by early September (blue bars) in the young and very few in those over 45. Finally by early September a large part of the English population was already immune (green bars), regardless of age.

Adding up the number of serconverters across ages gives an estimate of 2,552,000 infections in England by the beginning of September 2009 (range 696,000–7,186,000 based on 95% confidence intervals for the published serological studies). Yet, by that time, only 385,368 symptomatic cases had occurred according to HPA case estimates. We can take the ratio of symptomatic cases to all infections as an estimate of the proportion of cases that are symptomatic, parameter p in our model. The serological data imply that $p = 0.15$ (range 0.05–0.55). This is more easily interpreted as the inverse of p; the number for seroconverters for each recorded case. This has a mean of 6.62 with range 1.81–18.62. In what follows we use four different values of $1/p$: $1/p = 1$, every seroconverter is symptomatic; $1/p = 1.81$, the lower bound estimate from the serological data; $1/p = 6.62$, the mean from the serological data; and $1/p = 18.62$, the upper bound from the serological data.

The serological data also yield an estimate for the proportion of the population who were immune to the pandemic influenza strain before it arrived. The estimate is 12.7% and is used to define initial conditions for the model in equation (1).

Having made estimates for the initial conditions and the ratio of seroconverters to cases we proceed to fit the model to the case data. This amounts to finding a set of parameter values that are consistent both with the case reports and with the serological data. Figure 3 presents the best fits of our mathematical model to the case reports for four different values of $1/p$ ranging from the situation where all seroconverters present as cases ($1/p = 1$) to the upper limit of estimated values in which for each reported case there are more than 18 seroconverters ($1/p = 18.62$) [11]. All fits assume that 12.7% of the population are immune before the pandemic arrives, reflecting pre-existing immunity in the population. Our fitting procedure is to fix p to one of our four values, fix the percentage immune before the pandemic to 12.7% and then estimate the three parameters R_0, $1/\gamma$ and r_h by making a least square fit of the solution of the model (equations (1) and (2)) to the HPA's case report data. Table 2 reports best fit parameter estimates for each value of $1/p$. Overall, the best fit is for $1/p = 18.62$. The parametric fit yields $R_0 = 1.59$, very much in agreement with previous estimates (see Table 1). In addition, an infectious period of 2.42 days lies well within previous independent estimates for influenza A(H1N1) (see Table 3 and Table 4). The best-estimate reduction factor, $r_h = 0.65$ reveals that transmission is 35% lower over the period of school holidays. While there is no empirical data to validate this result, it seems plausible as the number of infected cases was highest in the age group of 5 to 14 years old (see Figure 2). This reconciliation of the case report data and the serological data concludes that the 2009 English epidemic of symptomatic cases was small for the following reasons:

1. a large fraction (12.7%) of the population were immune before the pandemic arrived in England,
2. for every reported case there were many (approx. 18) seroconversions that were not reported as cases so the susceptible pool was very quickly diminished,
3. initial estimates for R_0 of around 1.5 were roughly right.

Table 3. Infectious period estimates for influenza A(H1N1) from two independent studies with confidence interval (CI).

Infectious period
2.5 (CI: 1.1–4.0) [19]
3.38 (CI: 2.06–4.69) [24]

The infectious period is presented in days.
doi:10.1371/journal.pone.0030223.t003

Table 4. Generation time estimates for influenza A(H1N1) from several independent studies.

Generation time
6–8 [25]
1.91 (CI: 1.30–2.71) [1]
2.6 (CI: 2.2–3.5) [26]
2.6–3.2 [17]
2.7 (CI: 2.0–3.5) [27]
2.9 [21]
1.9–2.6 (CI: 1.3–3.0) [22]
6 (CI: 5–7) [23]

The generation time is presented in days. A range of possible generation times is shown when multiple estimations were made. The confidence interval (CI) is given wherever possible. A median latent period of 1.4 days (CI: 1.0–1.8) [27] and 2.62 days (CI: 2.28–3.12) [24] has been reported. The infectious period can be estimated using the generation time as latent period and half the infectious period are on average the generation time [4].
doi:10.1371/journal.pone.0030223.t004

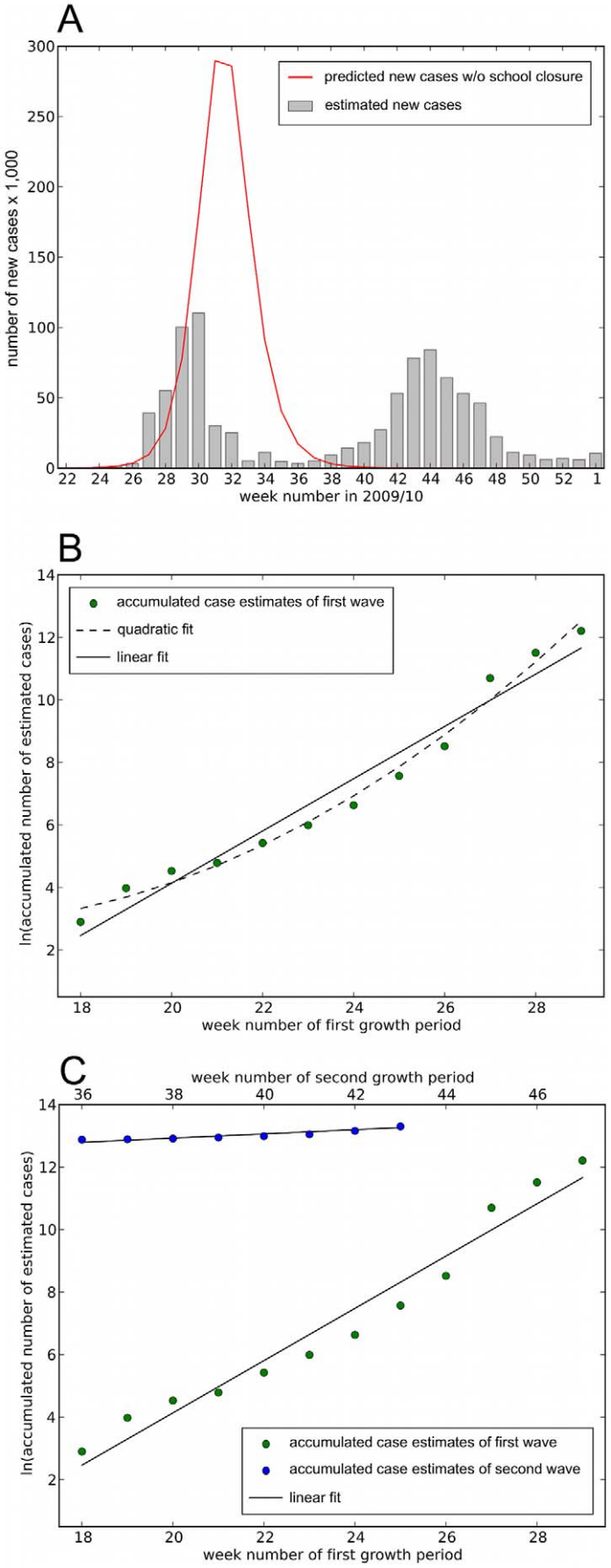
A
predicted new cases w/o school closure
estimated new cases
number of new cases x 1,000
week number in 2009/10
B
accumulated case estimates of first wave
quadratic fit
linear fit
ln(accumulated number of estimated cases)
week number of first growth period
C
week number of second growth period
accumulated case estimates of first wave
accumulated case estimates of second wave
linear fit
ln(accumulated number of estimated cases)
week number of first growth period

Figure 4. When did growth data reveal that the English epidemic would be small? A. shows that it was unlikely that the first wave would indicate that susceptible hosts were becoming exhausted. If there had been no school break we expect that the epidemic would have gone on growing for several weeks after week 30 when they actually shut. The red line shows predicted dynamics of the English epidemic if the schools had not closed (i.e. if $r_h = 1$ and other parameters are the best fits for the model of equation (1)). B. This is confirmed in regression analyses of the natural logarithm of the cumulative number of cases. Data are filled circles, linear fit is the solid line and quadratic fit is the dashed line. Linear fit: ln(cumulative cases) $= -13.5 + 0.84 * \text{week}$. Quadratic fit: ln(cumulative cases) $= -17.16 + 0.22 * \text{week} + 0.05 * (\text{week})^2$. The quadratic term is significant ($p = 0.002$) but small and positive, signifying that, if anything, the epidemic was accelerating just before the schools broke up for summer. C. compares growth rates of the first wave (green dots) and second wave (blue dots). An analysis of covariance reveals that the two slopes are significantly different ($p < 0.001$). Linear fit for second wave: ln(cumulative cases) $= 10.58 + 0.06 * \text{week}$.
doi:10.1371/journal.pone.0030223.g004

This raises the following question: how early would it have been possible to detect that there were so many seroconverters for every reported case? Was the epidemic already starting to slow in late July, before schools broke up? Figure 4 A answers this question with a clear "no". If the schools had not broken up, using our best-fit parameters, we believe that the epidemic would have continued to grow rapidly for several more weeks. It is therefore not surprising that a logarithmic plot of cumulated cases during the weeks before schools closed (Figure 4 B) shows that the epidemic was clearly still in its exponential growth phase with no indication of a falling off in the growth rate. However, when schools reconvened in September, the second pandemic wave clearly grew at a much slower rate than before the holiday. Figure 4 C plots cumulated case numbers in a logarithmic plot that compares the rising phases of the first and second waves. The second wave is clearly growing very much more slowly than the first wave, which is consistent with the large fraction of the population that had already seroconverted by the end of the first wave. In short, the very slow growth of the second wave of the pandemic was a clear indicator that many people had already seroconverted during the first wave.

Discussion

The HPA published their case estimates on influenza A(H1N1) infections over the period of the pandemic in England. Overall, the HPA estimated that approximately 910,000 individuals were ill with the pandemic influenza virus during 2009, amounting to less than 2% of the English population [2]. This surprisingly small number of symptomatic cases only makes sense in the light of the HPA's serological survey which found that around 17.7% of the English population were already immune by the end of the first wave of the pandemic in late September 2009. This observation, combined with knowledge of the level of immunity prior to the pandemics' arrival and the dates of the English summer holidays combine together to tell a coherent story about the 2009 influenza pandemic in England. The pandemic had a comparatively small impact because for every case that was reported there were around 18 seroconversions with no symptoms at all.

In retrospect, it was clear from the very slow growth of the second epidemic wave that something was different as the known cases up until September were too few to explains such slow growth. Depletion of susceptibles through asymptomatic cases would always have been a strong contender as a mechanism for this slow observed growth.

The HPA's serology study was organised, collected, analysed and published in a very short space of time. No other country has, to our knowledge, published such extensive data [13–15]. In principle, the same pattern that was picked up in these data in September would have been seen in samples drawn even earlier. This modelling study is a ringing endorsement of the explanatory power of serological surveys in a pandemic setting. However, it is not obvious that a serological survey could have been organised and executed any earlier in a pandemic than this one was.

The distinction between infection and disease is an early lesson in the teaching of infectious disease biology [16]. Predictions of the likely impact of the 2009 influenza pandemic were made using models that did not distinguish between cases of infection and cases of disease. A very simple model that allows asymptomatic seroconversion is presented here and yields a coherent reconciliation of the estimated numbers of disease versus the estimated number of infections through the course of the 2009 pandemic in England. Once it is clear how many asymptomatic cases there were, the epidemiology of the 2009 influenza pandemic in England makes sense.

Supporting Information

Table S1 Shown are the serological data of [11] combined with the demographic data of England given by [7]. Values in brackets show the 95% confidence intervals as estimated by [11]. All numbers shown are in 1,000s. Numbers might not sum due to rounding.
(PDF)

Author Contributions

Conceived and designed the experiments: RJK ARM. Performed the experiments: RJK. Analyzed the data: RJK ARM. Contributed reagents/materials/analysis tools: RJK. Wrote the paper: RJK ARM.

References

1. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, et al. (2009) Pandemic potential of a strain of inuenza A (H1N1): early findings. Science 324: 1557–61.
2. Health Protection Agency (2009) Weekly National Inuenza Reports. URL http://www.hpa. org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1244442494458.
3. Donaldson LJ, Rutter PD, Ellis BM, Greaves FEC, Mytton OT, et al. (2009) Mortality from pandemic A/H1N1 2009 inuenza in England: public health surveillance study. BMJ 339: b5213–b5213.
4. Anderson RM, May RM (1992) Infectious Diseases of Humans: Dynamics and Control (Oxford Science Publications). Oxford: Oxford University Press. 768 p. URL http://www.amazon.co.uk/Infectious-Diseases-Humans-Dynamics-Publications/dp/019854040X.
5. Scherer A, McLean AR (2002) Mathematical models of vaccination. Br Med Bull 62: 187–199.
6. Health Protection Agency (2009) Weekly Pandemic Flu Media Update. URL http://www.hpa.org.uk/NewsCentre/NationalPressReleases/2009PressReleases/.
7. Office for National Statistics (2009) 2009 Population Estimates in England. URL http://www. statistics.gov.uk/popest.
8. Health Protection Agency (2009) Synopsis of the method used to estimate the number of pandemic inuenza (H1N1) 2009 cases in England in the week 21 to 27July 2009. URL http://www.hpa. org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1244442494458.
9. Health Protection Agency (2009) Method used to estimate new pandemic (H1N1) 2009 inuenza cases in England in the week 3 August to 9 August 2009.

URL http://www.hpa.org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1244442494458.
10. Directgov (2009) School Term Dates. URL http://www.direct.gov.uk/en/Parents/Schoolslearninganddevelopment/SchoolLife/DG_4016103.
11. Miller E, Hoschler K, Hardelid P, Stanford E, Andrews N, et al. (2010) Incidence of 2009 pandemic inuenza A H1N1 infection in England: a cross-sectional serological study. Lancet 375: 1100–1108.
12. Girard MP, Tam JS, Assossou OM, Kieny MP (2010) The 2009 A (H1N1) inuenza virus pandemic: A review. Vaccine 28: 4895–4902.
13. Chen H (2009) Serologic Survey of Pandemic (H1N1) 2009 Virus, Guangxi Province, China. Emerg Infect Dis 41: 59–80.
14. Prachayangprecha S, Makkoch J, Payungporn S, Chieochansin T, Vuthitanachot C, et al. (2010) Serological analysis of human pandemic inuenza (H1N1) in Thailand. J Health Popul Nutr 28: 537–44.
15. Delangue J, Salez N, Ninove L, Kieffer a, Zandotti C, et al. (2011) Serological study of the 2009 pandemic due to inuenza A H1N1 in the metropolitan French population. Clin Microbiol Infect. pp 4–10.
16. Mims CA, Nash A, Stephen J (2000) Mims' Pathogenesis of Infectious Disease Academic Press. pp 474. URL http://www.amazon.co.uk/Mims-Pathogenesis-Infectious-Disease-Cedric/dp/0124982654/ref=sr_1_3?s=books&ie=UTF8&qid=1309779096&sr=1-3#.
17. Yang Y, Sugimoto JD, Halloran ME, Basta NE, Chao DL, et al. (2009) The transmissibility and control of pandemic inuenza A (H1N1) virus. Science 326: 729–33.
18. Nishiura H, Chowell G, Safan M, Castillo-Chavez C (2010) Pros and cons of estimating the reproduction number from early epidemic growth rate of inuenza A (H1N1) 2009. Theor Biol Med Model 7: 1.
19. Balcan D, Hu H, Goncalves B, Bajardi P, Poletto C, et al. (2009) Seasonal transmission potential and activity peaks of the new inuenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. BMC Med 7: 45.
20. Munayco CV, Gómez J, Arrasco J, Kochel TJ, Fiestas V, et al. (2009) Epidemiological and trans- missibility analysis of inuenza A(H1N1)v in a southern hemisphere setting: Peru. Euro Surveill 14: 1–5.
21. McBryde E, Bergeri I, Gemert CV, Rotty J, Headley EJ, et al. (2009) Early transmission characteristics of inuenza A(H1N1)v in Australia: Victorian state, 16 May–3 June 2009. Euro Surveill 14: 1–6.
22. de Silva UC, Warachit J, Waicharoen S, Chittaganpitch M (2009) A preliminary analysis of the epidemiology of inuenza A(H1N1)v virus infection in Thailand from early outbreak data, June–July 2009. Euro Surveill 14: 6–8.
23. Pourbohloul B, Ahued A, Davoudi B, Meza R, Meyers LA, et al. (2009) Initial human transmission dynamics of the pandemic (H1N1) 2009 virus in North America. Inuenza Other Respi Viruses 3: 215–22.
24. Tuite AR, Greer AL, Whelan M, Winter AL, Lee B, et al. (2010) Estimated epidemiologic parameters and morbidity associated with pandemic H1N1 inuenza. Can Med Assoc J 182: 131–6.
25. Dawood FS, Jain S, Finelli L, Shaw MW, Lindstrom S, et al. (2009) Emergence of a novel swineorigin inuenza A (H1N1) virus in humans. N Engl J Med 360: 2605–15.
26. Cauchemez S, Donnelly CA, Reed C, Ghani AC, Fraser C, et al. (2009) Household transmission of 2009 pandemic inuenza A (H1N1) virus in the United States. N Engl J Med 361: 2619–27.
27. Lessler J, Reich NG, Cummings DaT, Nair HP, Jordan HT, et al. (2009) Outbreak of 2009 pandemic inuenza A (H1N1) at a New York City school. N Engl J Med 361: 2628–36.