AGENT-BASED MODELING OF THE SPREAD OF THE 1918-1919 SPANISH FLU IN THREE CANADIAN FUR TRADING COMMUNITIES

A Thesis presented to the Faculty of the Graduate School

University of Missouri- Columbia

In Partial Fulfillment

Of the Requirements for the Degree

Master of Arts

---

by

CAROLINE AHILLEN

Dr. Lisa Sattenspiel, Thesis Supervisor

MAY 2006

The Undersigned, appointed by the Dean of the Graduate School,
have examined the thesis entitled:

AGENT-BASED MODELING OF THE SPREAD

OF THE 1918-1919 SPANISH FLU

IN THREE CANADIAN FUR TRADING COMMUNITIES

Presented by Caroline Ahillen

A Candidate for the degree of Master of Arts

And Hereby Certify that in their opinion it is worthy of acceptance.

Professor Lisa Sattenspiel

Professor Craig Palmer

Professor Karen Cone

ACKNOWLEDGEMENTS

I would like to thank all of those who have helped to make this project possible. I would first like to thank Dr. Lisa Sattenspiel for helping me to formulate this project, guiding me along the way, and for enduring my repeated editing problems and long-distance communications. I would also like to thank Connie Carpenter who introduced me to agent-based modeling and whose Norway House model formed the basis for this project. Without her work, this project would not have been possible. I would like to thank Nate Green, my computer programmer, for helping me to write the program code and for being so easy to work with. I would have been lost without his assistance. I would also like to thank my committee members, Dr. Craig Palmer and Dr. Karen Cone for helping me to evaluate and edit this rather lengthy paper. Finally, I would like to thank my family and friends for being so supportive and enduring my moods as I worked on this project. This research was supported by NSF Grant No. SBR-0094449.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (continued)

Figure **Page**

LIST OF TABLES

LIST OF TABLES (continued)

Chapter 1: Introduction

The 1918-1919 Spanish influenza pandemic caused millions of deaths as it raced around the globe at the close of World War I. The virulence and high mortality rates associated with this pandemic have been unmatched by any recorded flu epidemic before or since (Crosby, 2003). While the impact of the 1918-1919 flu was felt in almost every corner of the world, it proved especially devastating for the small Aboriginal fur trading communities of the central Canadian subarctic, where survival was already precarious. Among these, the community of Norway House, Manitoba appears to have been one of the hardest hit, and is estimated to have lost 18% of its population in just six weeks time (Herring, 1994). Previous analyses of historical and archival data from Norway House and the neighboring communities of Oxford House and God's Lake have provided a tantalizing glimpse into the 1918-1919 flu epidemic in this region of Canada, but have left some important questions unanswered. In this project, agent-based computer simulation is used to model the spread of the 1918-1919 flu epidemic at Norway House, Oxford House, and God's Lake in an attempt to address these questions.

Modeling is a powerful tool for the study of epidemic disease and can provide insights that would otherwise be impossible to obtain. The first model that was used to study the 1918-1919 flu at Norway House, Oxford House, and God's Lake was a deterministic mathematical model developed by Lisa Sattenspiel, D. Ann Herring, and colleagues (Sattenspiel and Dietz, 1995; Sattenspiel, 1990; Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002; Sattenspiel and Herring, 2003; Sattenspiel et al., 2000). While this Norway House Ordinary Differential Equation (NHODE) model significantly

expanded knowledge about the 1918-1919 flu epidemic in this region, concerns regarding the ability of the model to accurately represent small or diverse populations ultimately led to the development of a stochastic agent-based computer simulation of the 1918-1919 flu at Norway House: The Norway House model.

Because of its agent-based format, the original Norway House model (Carpenter, 2004) has provided a more realistic representation of the Norway House community, its diverse residents, local mobility patterns, and human social interaction than was possible with the NHODE model. Further, because it was stochastic rather than deterministic, the Norway House model was also able to account for the influence of stochasticity upon the epidemic patterns. Although the Norway House model was an advancement over Sattenspiel's NHODE model for these reasons, it had one major limitation. Because it simulated the 1918-1919 flu at only Norway House, the Norway House model provides no information regarding the spread of the epidemic between communities or regarding epidemic impact on a larger geographic scale, and thus, the insights it has provided are somewhat limited in scope. This focus on only Norway House also meant that the results from the Norway House model could also not be directly compared to Sattenspiel and Herring's previous work, which had considered the spread of the 1918-1919 flu epidemic among Norway House and the two adjacent communities of Oxford House and God's Lake (Herring and Sattenspiel, 2003; Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002; Sattenspiel and Herring, 2003; Sattenspiel et al., 2000)

The primary goal of this project is to address this limitation by expanding the one-post Norway House model into a three-post model. This requires that two additional communities, to represent Oxford House and God's Lake, be added to the original model

program and that fundamental additions be made to the original program code in order to allow for inter-community mobility. The resulting model, the Norway House, Oxford House, God's Lake (NHOHGL) model will simulate the spread of the 1918-1919 flu among, as well as within these three communities and will thus allow for the 1918-1919 flu epidemic to be studied on a larger, regional geographic scale. Following extensive testing, the new NHOHGL model can then be used to address several questions which could not have been addressed with a single community simulation.

For this project, the new NHOHGL model will be used to investigate three central research questions, two of which concern mysteries that continue to surround the 1918-1919 flu epidemic in this region of Canada. First, the Cree and Métis groups which inhabited this region of Canada practiced a seasonal pattern of social aggregation and dispersal that, in combination with seasonal differences in travel rates, may have had a strong influence on local epidemic patterns. The NHOHGL model will be used to determine what, if any, effect these seasonal patterns may have had upon the 1918-1919 flu epidemic in this region. Second, while it is clear that Norway House was severely impacted by the 1918-1919 flu epidemic, Oxford House and God's Lake appear to have been relatively unaffected (Herring, 1994; Herring and Sattenspiel, 2003). The NHOHGL model will be used to provide insight into why the epidemic may have failed to spread to these neighboring communities.

The third and final question to be addressed in this project pertains more to the practice of epidemic modeling than to the archival data. The agent-based approach employed by the NHOHGL model in order to simulate the 1918-1919 flu epidemic at Norway House, Oxford House, and God's Lake provides an alternative technique to the

mathematical modeling approach that was used in the NHODE model. While a comparison of the two modeling techniques could provide useful insights into abilities of each to accurately simulate an epidemic among a small population, such a comparison was not possible with the original Norway House model because this modeled the epidemic at one community (Norway House) while the NHODE model profiled three (Norway House, Oxford House, and God's Lake). With the three-post format of the NHOHGL model, a more direct comparison of the two modeling techniques may now be made.

The insights gained through the testing of the NHOHGL model and the investigation of these research questions have the potential to provide important insights into the 1918-1919 flu epidemic in the Norway House region of Manitoba, and, in addition, they are also more widely applicable. The data produced by the NHOHGL model simulations can be used to help determine which social and epidemiological factors may help or hinder the spread of an influenza epidemic or increase or decrease its severity. Such knowledge could potentially aid in the study of other historical virgin-soil epidemics for which less documentation is available; for example, the massive mortality that followed the initial introduction of European crowd diseases to the Native Americans. Such knowledge could also prove invaluable in the event of a future influenza pandemic by providing information that could be used to help minimize or halt epidemic spread. Today, as the H5N1 "Avian" flu virus threatens to spark another major influenza pandemic, perhaps even larger than the 1918-1919 Spanish flu, such knowledge could be more important than ever before.

Chapter 2: Influenza and the 1918-1919 Influenza Pandemic

The 1918-1919 Spanish influenza was a major global pandemic that was responsible for an estimated 20 to 40 million deaths worldwide. In Canada alone, the disease was responsible for about 50,000 deaths (Sattenspiel and Herring, 2002) and in the U.S. influenza caused 550,000 deaths, about ten times the number of American casualties in World War I (Karlen, 1995). During the winter of 1918-1919, the flu raced around the globe, spreading even to remote locations, and often following the paths of soldiers returning home from war. Wherever it struck, it infected large portions of the population and left many dead in its wake. This chapter profiles the flu virus, the emergence of pandemic influenza viruses, and the 1918-1919 pandemic.

Influenza and the Influenza Virus

Influenza is a respiratory infection that is caused by the influenza virus. A virus is simply a small portion of genetic material, either RNA or DNA, wrapped in a lipid-protein coat (Crawford, 2000; Taubenberger et al., 2005a) and is completely dependent on the host for survival and reproduction. Influenza viruses are RNA viruses, and are classified as orthomyxoviruses (CDC, 2004; Cox and Subbarao, 1999). There are three types of influenza: influenza A, B, and C. Infections with the types B and C viruses are relatively mild and have never been associated with pandemics. Infections with type A viruses, however, are more severe and lead to yearly outbreaks of influenza as well as occasional pandemics (CDC, 2004; Herring, 2000; Stevens et al., 2004; Taubenberger et al., 2005a; WHO, 2003).

The influenza virus has two primary surface proteins which extend like spikes from its lipid-protein coat: hemagglutinin (HA), which attaches the virus to its host cell, and neuraminidase (NA), which cleaves off newly formed virus cells from the host cell (Stevens et al., 2004; Taubenberger et al., 2005a). These proteins play a vital role in infection and viral reproduction and are also the major surface antigens, the structures that stimulate the production of antibodies by the immune system (Couch and Kasel, 1983; Taubenberger et al., 2005a). Because each type of influenza virus has a distinctive HA and NA protein configuration, these proteins are also used to identify particular viral subtypes. For example, the 1918-1919 flu pandemic was caused by an H1N1 virus. While numerous subtypes of the influenza virus have been identified, only a few are known to infect humans. The majority of influenza viruses may instead be found in aquatic birds, which are the natural reservoir, or natural host, for the virus and tend to carry influenza as a harmless infection of the gut (CDC, 2004; Cox and Subbarao, 1999; Crawford, 2000)

Influenza is transmitted from person to person via the small airborne droplets that are emitted during coughing, sneezing, or breathing. The latent period, or the period of time between infection and when a victim becomes contagious, is short and generally lasts between 2 and 5 days (Hart, 2004). An infected person can become contagious a full day before symptoms develop, and may remain contagious for up to seven days (WHO, 2003). This combination of an airborne mode of transmission and a short latent period allows influenza to spread quickly and easily, especially in crowded conditions. Cold weather also helps to facilitate the spread of influenza because the virus can survive

for longer periods of time outside of the body when temperatures are cool (WHO, 2003). This is the reason why outbreaks of influenza tend to occur during winter.

Once the virus enters a new host, the incubation period, or the period between infection and when an infected individual begins to show symptoms, is usually short. It generally lasts about 1-5 days (Cox and Subbarao, 1999; Hart, 2004) but was often 48 hours or less during the 1918-1919 pandemic (Crosby, 2003). Symptoms generally develop quickly and include fever, chills, headache, muscle aches, myalgia, prostration, malaise, dry cough, sore throat, rhinitis, and loss of appetite (Cox and Subbarao, 1999; Hart, 2004; Karlen, 1995; WHO, 2003). Because the symptoms of influenza are similar to those of other respiratory diseases, diagnosis by symptoms alone is difficult (Hart, 2004; WHO, 2003) and lab tests are usually necessary to confirm a diagnosis of influenza (Cox and Subbarao, 1999; Hart, 2004).

Most flu victims recover in about a week and require no medical treatment (WHO, 2003), but influenza can cause pneumonia and death in the elderly, the very young, and in those with preexisting medical conditions (Cox and Subbarao, 1999; Hart, 2004; WHO, 2003). Because of this, at-risk groups are encouraged to be vaccinated yearly in order to help prevent infection. Treatment with antiviral drugs may also help to prevent infection and may reduce the duration of symptoms if infection occurs (Cox and Subbarao, 1999; WHO, 2003). In addition, antiviral drugs may help to contain an emerging epidemic (Longini et al., 2004).

Following a bout with influenza, most people develop some form of immunity to the specific virus that attacked them. Immunity is induced by the creation of antibodies that are able to recognize influenza surface proteins, especially the HA (Couch and Kasel,

1983; Stevens et al., 2004; Taubenberger et al., 2005a). Although the extent and specificity of such immunity is debated, studies have indicated that immunity to a particular HA subtype can last between 4 and 7 years (Couch and Kasel, 1983). However, immunity against one HA subtype is not effective against others and, because the flu virus continually mutates, immunity to one year's strain often does not confer immunity to the strains circulating the following year.

The yearly reappearance of influenza has made it a familiar presence, and has led to an under-appreciation of the total health impact of the virus (CDC, 2004; Karlen, 1995). Although influenza is generally a mild disease, with mortality rates of about 0.01% (Karlen, 1995), it is also highly infectious and infects millions every winter. This can lead to significant mortality and even larger losses in productivity. According to the World Health Organization (2003), influenza attacks between 5 and 15% of the world's population each year, causing an estimated 3 to 5 million cases of severe illness and between 250,000 and 500,000 deaths. If such figures are generated by an average flu outbreak, the impact of a major pandemic would be unthinkable.

The Making of a Pandemic

Influenza pandemics occur when a new subtype of influenza A virus emerges that is able to cause disease, able to replicate in humans, and able to spread from person-to-person (Lazzari and Stöhr, 2004). Because humans have had no prior exposure to these viruses, they can cause massive virgin-soil epidemics, or epidemics to which no humans have any pre-existing immunity. When travel and contact rates are sufficient, these epidemics can quickly balloon into a worldwide pandemic, such as the 1918-1919 flu.

Two mechanisms can generate new human influenza viruses: antigenic drift and antigenic shift. Antigenic drift refers to the accumulation of minor mutations which frequently occur in the surface proteins of an existing strain of influenza. This process leads to continual changes in the structure of a virus's HA and NA surface proteins (Cox and Subbarao, 1999; Reid et al., 2003), which allows the virus to evade immune detection and to cause infection year after year. Although antigenic drift may mean that a new flu vaccine needs to be created each year, it does not lead to the production of pandemic viruses. That requires an antigenic shift.

An antigenic shift is a major genetic change in the structure of an influenza virus's HA and/or NA surface proteins. Antigenic shifts occur more rarely than antigenic drift, but are much more dangerous. Antigenic shifts can occur when an avian or swine flu virus gains the ability to infect humans and to be transmitted from person-to-person (Cox and Subbarao, 1999) or when different influenza viruses exchange genetic material in a process known as gene reassortment (CDC, 2004; Cox and Subbarao, 1999; Crawford, 2000; Karlen, 1995; Taubenberger et al., 2005a). Either method can produce a completely novel human influenza virus and trigger a pandemic.

Three influenza pandemics have occurred during the past century, each of which was associated with a major antigenic shift. The first, was the 1918-1919 Spanish influenza pandemic, which was triggered by the emergence of the H1N1 virus; the second, was the 1957 Asian Flu pandemic, which signaled the arrival of the H2N2 virus; and the third was the 1968 "Asian flu" pandemic, which followed the arrival of the H3N2 subtype. Each of these pandemics caused significant morbidity and mortality, although the 1918-1919 pandemic was by far the most dramatic.

Many scientists agree that it is only a matter of time until we are faced with another influenza pandemic and there have been some recent close calls. In 1976, the reappearance of the H1N1 subtype caused concern that it could trigger an epidemic similar to the 1918-1919 flu (Crawford, 2000; Karlen, 1995), but the ensuing outbreak was relatively minor and was easily controlled. In 1997, the alarm was again sounded following an outbreak of influenza in Hong Kong that was triggered by an avian H5N1 virus (Hart, 2004; Holmes et al., 2005). Although the outbreak was quickly contained, H5N1 viruses continued to circulate among waterfowl and in 2003, caused another outbreak in Hong Kong (Holmes et al., 2005; WHO, 2003). This avian flu virus continues to cause limited outbreaks in Asia and has recently been receiving a large amount of media attention. The virus is currently being monitored by the World Health Organization for signs that it has gained the ability to be transmitted from person-to-person (Appenzeller, 2005; Crawford, 2000; Lazzari and Stöhr, 2004; WHO, 2004) and public health officials have expressed concern that a modern influenza pandemic could be significantly more severe than even the 1918-1919 Spanish influenza due to the increased speed of international travel, an increasingly large and urbanized world population, and increasing numbers of elderly and immuno-compromised individuals (CDC, 2004; Crawford, 2000; WHO, 2004). Given the continued threat posed by pandemic influenza, it is useful to know as much as possible about the worst modern pandemic, the 1918-1919 Spanish flu.

The 1918-1919 Influenza Pandemic

In 1918, the emergence of the H1N1 influenza virus sparked a massive virgin soil epidemic which claimed millions of lives. The disease was dubbed the Spanish flu, not because it arose in Spain, but because it struck during World War I, and Spain was a neutral country with no media censorship to keep news of its health problems from reaching the rest of the world (Crosby, 2003; Taubenberger et al., 2005a). The pandemic spread around the globe in three waves. The first arose in the spring of 1918, possibly from a U.S. military camp in Kansas (Crawford, 2000; Crosby, 2003; Herring, 2000; Taubenberger et al., 2005a). This wave was relatively minor and caused only limited concern, but it was followed by a much more severe and wider-reaching pandemic wave in the fall and winter of 1918-1919. It is this wave that is most often associated with the 1918-1919 flu and that was responsible for most of the mortality. This main pandemic wave was then followed by a third, less severe and more limited wave in the winter of 1919.

The 1918-1919 flu exhibited many unusual features that distinguished it from normal yearly influenza epidemics. Especially notable was its exceptional virulence and high mortality rates. The flu tended to claim between 2.5% and 5% percent of its victims (Crawford, 2000; Herring, 2000; Taubenberger et al., 2005a) and even more in some locations, including the Canadian subarctic (Herring, 2000). It has been estimated that during the pandemic, between one third (Taubenberger et al., 2005a) and one half (Crawford, 2000; Johnson, 2003; Stevens et al., 2004) of the world's population was infected and that at least 20 million (McNeil, 1976; Taubenberger and Reid, 1997) and probably closer to 40 or 50 million people died (Crosby, 2003; Herring, 2000; Lazzari

and Stöhr, 2004; Reid et al., 2003; Skinner, 1909; Taubenberger et al., 2005a; Tumpey et al., 2005; WHO, 2004).

The 1918-1919 flu also exhibited distinctive symptoms. The virus attacked with exceptional speed, bringing healthy individuals to utter prostration and even death in less than 48 hours (Crosby, 2003). Symptoms were generally similar to those seen in normal influenza outbreaks and included headache, fever, eye, ear and back pain, and a "feeling of severe illness" (Corish, 1920). However, in some severe cases, the faces of the victims turned blue as their lungs filled with thick bloody fluid (Crosby, 2003; Taubenberger et al., 2005a), a condition known as heliotrope cyanosis (Johnson, 2003). A bluish tinge to the skin usually signaled imminent death and is a distinctive feature of the 1918 flu (Crawford, 2000; Crosby, 2003; Johnson, 2003; Taubenberger et al., 2005a). However, although many flu victims quickly succumbed to this form of viral pneumonia, most of the mortality during the 1918-1919 flu pandemic was attributable to secondary pneumonia caused by opportunistic bacterial infections in lung tissue that had been damaged by the flu virus (Corish, 1920; Crawford, 2000; Crosby, 2003; Taubenberger et al., 2005a). In the days before antibiotics, such infections were often fatal.

Another striking feature of the 1918 flu was the speed and breadth of its geographic spread. The virus managed to circle the globe and kill millions in just six months (Karlen, 1995). The pandemic reached all corners of the globe, leaving few areas unaffected. The flu struck hard in major cities and on the Western Front (Crosby, 2003), but it also managed to infiltrate more remote locations such as the Canadian wilderness (Herring, 2000), Alaska (Taubenberger and Reid, 1997; Taubenberger et al., 2005a), and remote islands in the Pacific (Crosby, 2003; Karlen, 1995; Taubenberger and Reid,

1997). The rapidity and extent of the 1918-1919 flu was likely due to the virulence of the strain, the effects of World War I, and the more efficient forms of transportation that were available in 1918 (Crawford, 2000). The flu was especially rampant in the military and many accounts suggest that the virus was often carried to new locations by soldiers returning home from war (Crawford, 2000; Crosby, 2003; Herring, 2000). The flu would have spread quickly along railways, roads and steamships to all corners of the globe with the returning soldiers, causing outbreaks wherever they happened to stop along the way.

Perhaps the most unusual characteristic of the 1918-1919 influenza pandemic was its choice of victims. While most influenza viruses tend to target the very young, the elderly, and those with compromised immune systems, the 1918-1919 H1N1 virus attacked people of all ages at nearly equal rates (Crawford, 2000; Crosby, 2003; Karlen, 1995; Taubenberger and Reid, 1997; Taubenberger et al., 2005a; Taubenberger et al., 2005b). For this reason, mortality among young adults and teenagers was proportionally higher than would otherwise have been expected. In 1918, influenza and pneumonia deaths among young adults were more than 20 times higher than in previous years while deaths among the elderly were comparatively low (Taubenberger and Reid, 1997). The reasons for this trend are uncertain. It is possible that the elderly possessed some immune resistance to the 1918 pandemic virus due to experiences with earlier flu epidemics (Taubenberger et al., 2005a). It is also possible that the high mortality among young adults was a consequence of World War I, which placed many young men into crowded and unsanitary conditions where influenza could be easily spread (Crosby, 2003). Another theory, developed by Sir MacFarlane Burnet, suggests that the robust immune system of young adults may have triggered a more severe immune system response upon

infection, leading to intense inflammation in the lungs and causing fluid to build up and cause suffocation (Crosby, 2003).

Other demographic groups also seemed to be favored by the 1918 virus. During the 1918 flu, mortality was especially high among pregnant women (Cox and Subbarao, 1999; Crosby, 2003; Karlen, 1995) and infection often led to spontaneous abortion (Crosby, 2003). Additionally, American Indians "suffered hideously in the pandemic" (Crosby, 2003), probably due to the deteriorating social conditions for many native groups at this time. African Americans, in contrast, appeared to be less susceptible than other groups, possibly because more of them had caught the flu in the earlier, and milder, spring wave (Crosby, 2003).

When the flu arrived, it tended to quickly overwhelm and overburden local medical officials and public health systems (Crosby, 2003; McNeil, 1976) and there were often not enough healthy people to care for the sick and to bury the dead. Efforts to control the spread of the pandemic were numerous, but mostly unsuccessful. Many localities instituted quarantine policies, but these efforts were often begun too late to stop the rapid spread of the epidemic. Many cities also closed public meeting places such as schools, theatres, and churches (Crosby, 2003; Herring, 2000; McNeil, 1976) and some instituted even more extreme measures, such as requiring citizens to wear face masks (Crosby, 2003) or launching massive disinfection campaigns (Karlen, 1995), but these measures also did little to mitigate the spread of the 1918 pandemic. Another common response was vaccination. Several doctors developed vaccines and used them to immunize large numbers of people (Crosby, 2003; Karlen, 1995). These vaccines were essentially useless against the flu virus, but often appeared to work because they tended

to be administered as the pandemic was already fading (Crosby, 2003) and also helped to convince the public that local authorities were at least making an effort to stop the deadly disease.

For those who did become infected with the 1918-1919 flu, various remedies were suggested to help prevent death but these too were mostly ineffective. A contemporary home medical book recommended taking quinine or aspirin, keeping the feet warm, keeping the bowels open, drinking plenty of hot lemonade or hot milk, drinking teas made from dried raspberries, camphor, or cayenne pepper, and also suggested prevention measures including letting in fresh air, rinsing the mouth and nostrils after visiting overcrowded places, and taking care not to “inhale the breath of others" (Corish, 1920). The only proven remedy for the 1918-1919 flu was good nursing care (Crosby, 2003; Herring, 2000), something that was at a premium during the pandemic due to its coincidence with World War I.

The Identity of Influenza

For decades following the 1918-1919 influenza pandemic, the true identity of the infectious agent remained unknown. Viruses were not discovered until the 1930s, and influenza was not identified as the causal agent for the 1918-1919 pandemic until 1933 (Cox and Subbarao, 1999; Herring, 2000). Contemporary medical professionals were baffled by this killer disease which seemed to have appeared out of nowhere. Initially, some blamed the pandemic on the Germans, or on poor conditions brought on by the war (Crosby, 2003; Karlen, 1995). However, a more common explanation was that the 1918-1919 flu was caused by a previously unknown disease, “a new plague” (Taubenberger et

al., 2005a). The most popular candidate was Pfeiffer's bacillus (*Haemophilus influenzae*), which was often discovered in lung tissue samples taken from victims of the pandemic (Crawford, 2000; Crosby, 2003). Other bacteria, including streptococci and pneumococci, were also frequently found in the lungs of victims, leading some doctors to finger them as the culprit (Corish, 1920; Crosby, 2003). However, we now know that these bacteria caused only opportunistic pneumonia in those already infected with the 1918-1919 flu virus, which was the real monster behind the 1918-1919 Spanish flu.

Even though the identity of the infectious agent behind the pandemic has now been firmly established, the reasons for the 1918-1919 flu virus's extreme virulence are just beginning to be elucidated. A research team led by Jeffrey Taubenberger, working with tissue samples from victims of the pandemic, has analyzed the genetic structure of the 1918 flu virus in an effort to locate genes or proteins that could have been responsible for the virulence and severity of the 1918-1919 flu (Taubenberger et al., 2005a). This team has recently published the entire genetic sequences of all 8 RNA segments of the flu virus (Taubenberger et al., 2005b) and Tumpey et al. (2005) have managed to resurrect the virus and transmit it to mice. This research suggests that the 1918 virus possessed specific genetic features that can explain its high virulence. Tumpey and his team have discovered that the virus had the unique ability to replicate in the absence of trypsin, a protease that is most common in lung cells, and thus the virus could grow in other cells. They also found that certain genetic features of the 1918 virus allowed it to replicate particularly efficiently in bronchial cells, a feature which may explain the speed of the virus's attack (Tumpey et al., 2005). In addition, they have determined that the 1918-1919 influenza pandemic was triggered by an avian virus that had adapted to humans

(Taubenberger et al., 2005b). These findings make the study of the 1918-1919 flu particularly relevant as we face the possibility of the H5N1 avian flu gaining a similar adaptation.

Chapter 3: The Study Communities and the 1918-1919 Flu Pandemic

In order to ensure that the NHOHGL model would be as accurate a representation of the 1918-1919 flu epidemic at Norway House, Oxford House, and God's Lake as possible, it was necessary to become familiar with the inhabitants of these communities, and the social and environmental landscape. It was also necessary to obtain particular information concerning the experiences of these communities with the 1918-1919 flu pandemic. These data, derived from historical, archival, and ethnographic sources, were used to construct a framework for the model that would approximate local social organization, economic practices, and settlement patterns around the time of the 1918-1919 influenza pandemic. They were also used to test the new model.

Life in Central Manitoba

Norway House, Oxford House, and God's Lake are all former fur trading posts that are located in Central Manitoba, in what was once the Hudson's Bay Company Keewatin Fur Trade District (Herring and Sattenspiel, 2003). The three communities occupy a portion of the Canadian Shield, a large mass of crystalline rock which underlies much of central Canada. The regional landscape is characterized by tundra, bogs, swamps, and coniferous boreal forest (Hallowell, 1992; Lytwyn, 2002) and is crisscrossed by a complex network of rivers and lakes. Even though the region was unsuitable for agriculture because the ground was either frozen or swampy for most of the year, it had long been inhabited by native hunter and gatherer groups, mostly Cree

and Ojibwa, who had taken advantage of its abundant game resources and subsisted on a mainly hunting and fishing economy (Smith, 1981).

Life in this region revolved around the changing seasons, which led to dramatic variations in climate, terrain, and subsistence resources at different times of the year (Hallowell, 1992; Herring, 1994). Winter began in November and lasted through March, with temperatures routinely reaching -20°F in February (Hallowell, 1992). The first snowfall often occurred in the fall, and the region averaged between 50 and 100 inches of snow per year (Hallowell, 1992). During the winter months, the Cree hunted large and small game and trapped fur-bearing animals. The lakes and rivers remained frozen and travel was limited to snowshoe, toboggan, or dogsled, often following the frozen river trails. The snow usually began to melt sometime in April, and the ice covering the rivers and lakes usually began to break up in early May (Hallowell, 1992). In the springtime, waterfowl returned from their winter migrations and the waterways once again opened up for fishing, providing new subsistence resources. Travel was especially difficult during the spring, as meltwater made the ground swampy and the rivers treacherous (Herring and Sattenspiel, 2003). Summer lasted from June through August and brought temperatures that could surpass +72°F (Hallowell, 1992). During the summertime, the waterways remained open and the canoe was the preferred means of travel (Flannery, 1995; Hallowell, 1992), allowing for larger groups to travel more rapidly and frequently than was possible at any other time of the year.

The Cree and Ojibwa had adapted to these changing seasonal conditions by practicing a seasonal pattern of residence that resulted in lower and dispersed population densities in the winter, when resources were scarce and widespread, and larger population

densities in the summer, when resources were plentiful and localized. This seasonal residence pattern revolved around the basic social and economic unit of Cree and Ojibwa society, the extended-family hunting band. These semi-autonomous groups generally consisted of a few patrilaterally or bilaterally related kinsmen and members of their households (Bishop, 1974; Hallowell, 1976; Lytwyn, 2002) and usually numbered about 15-30 individuals. In the late fall, the hunting bands dispersed across the hinterland, each traveling to their respective family hunting and trapping grounds (Bishop, 1974; Hallowell, 1992; Hanks, 1982; Herring, 1994; Ray, 1974; Ray, 1976), the rights to which were traditionally established by customary usage. During the summer months, several extended-family hunting groups would come together and cluster near larger lakes and rivers to take advantage of fishing opportunities (Flannery, 1995; Hallowell, 1992; Ray, 1976). Summer gatherings provided increased opportunities for social interaction with non-kin and were often the site of marriages, festivals, and ceremonies (Flannery, 1995).

The Influence of the Fur Trade

During the eighteenth and nineteenth centuries, Central Manitoba became an important focus of the Northern Fur trade due to its richness in fur-bearing animals (Hallowell, 1992; Sattenspiel and Herring, 1998). The Hudson's Bay Company (HBC) established numerous posts across the Canadian subarctic in order to facilitate the trade and transport of furs and supplies. Even before these posts were established, fur trapping had taken priority over hunting and fishing as a means of subsistence for most of the native inhabitants of the Canadian Shield, and the European items for which furs were exchanged had become an essential part of their ecological adaptation and everyday lives

(Hallowell, 1992; Hanks, 1982; Long, 1995). Furs were collected from the interior by trappers, who were primarily of aboriginal or mixed descent. They were then brought to the HBC posts, which were staffed by European employees, where they could be exchanged for European goods and foodstuffs (Hallowell, 1992; Hanks, 1982). The furs at the posts were then collected by brigades of HBC employees known as trippers, who would carry them along the rivers and waterways to Hudson's Bay, where they could be exported for sale in Europe (Sattenspiel and Herring, 1998; Sattenspiel et al., 2000).

The HBC operated on a credit and debt system. The posts would issue supplies to the trappers on credit prior to their winter trapping trips and in the late spring, the trappers would then bring back the furs they had collected to repay these debts (Hallowell, 1992; Hanks, 1982). As the fur trade came to dominate local economy and native demand for European goods increased, the Cree and Ojibwa became increasingly mired in this system and were less able to pay off their debts (Hallowell, 1992). This was especially the case after the HBC established a monopoly over the fur trade in 1821 (Hanks, 1982), and the problem became even more pronounced in the late nineteenth century, as the populations of fur-bearing animals declined and the demand for furs decreased. Individual trappers became increasingly tied to the particular HBC post where they held their debt and more sedentary populations had begun to grow up around the posts by the early 1900s.

In addition to these economic changes, the fur trade and the Hudson's Bay Company also brought many social changes. Frequent intermarriage between the European HBC employees and native, mostly Cree, women produced a new category of people, the Métis, or mixed bloods. Additionally, as the HBC became more and more

powerful, it began to establish more formalized hunting territories (Hallowell, 1976; Hanks, 1982; Ray, 1974) often with little regard for traditional land rights. The HBC also established formal leadership positions, such as the "trading chief," which were previously unknown among the traditionally egalitarian Cree and Ojibwa. Along with the fur trade also came missionaries, and by the late nineteenth century, most of the natives had converted to Christianity (Hallowell, 1992). The incorporation of the region under the Dominion of Canada in 1867 and the signing of Treaty 5 in 1875 brought further social change as it required the Cree and their neighbors to formally surrender much of their territory and political autonomy (Hallowell, 1992).

Even though the influence of the Euro-Canadian government, missionaries, and the fur trade led to the alteration or abandonment of many traditional cultural patterns, the fur trade served to reinforce others, like the traditional seasonal movement patterns (Long, 1995). Because these patterns proved to be useful in the fur trade, they were retained (Hallowell, 1992; Herring, 1994). In the winter, the extended-family hunting bands would disperse across the hinterland to gather furs for the HBC and in the summer, they would return to the HBC posts to gather, trade their furs, pay their debts, and obtain supplies for the following winter. This seasonal pattern continued to be practiced well into the twentieth century (Hallowell, 1992; Flannery, 1995) and may have impacted local disease patterns (see chapter 7).

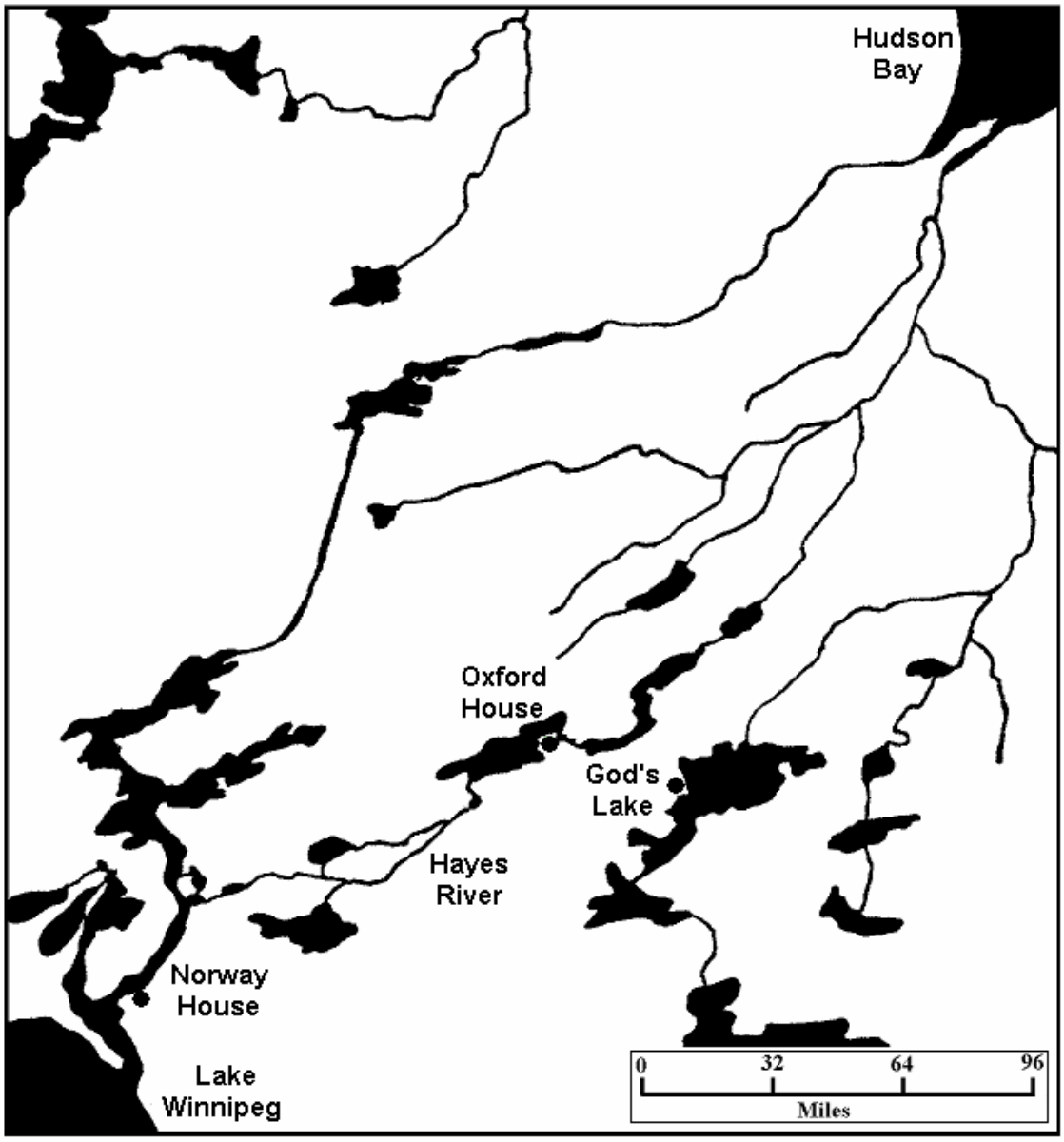
The Study Communities

Norway House, Oxford House, and God's Lake were three Cree/Métis fur-trading communities that had grown up around neighboring HBC posts. The geographic location

of the three communities is shown in Figure 3.1. Norway House was established as a supply base in 1801 and was located at a strategic position near the southern end of the Nelson River trade route, which was the main path between the interior and York Factory on the coast of Hudson Bay (Herring, 1994; Herring and Sattenspiel, 2003). By 1820, Norway House had become a central provisioning center at the crossroads of several major trade routes (Hallowell, 1992; Herring, 1994; Herring and Sattenspiel, 2003). Its importance began to wane during the late 19th century, as the fur trade went into decline and construction of railways and the increasing use of steamships caused a shift in trade routes. Yet, in 1918, Norway House remained a bustling community and an important focus of regional activity (Herring, 1994). It was the largest and most central of the three communities that are modeled in this study, and had a population of 736 in 1917 (Herring and Sattenspiel, 2003).

Oxford House was a much smaller and less central community than Norway House and had a population of 322 in 1917 (Herring and Sattenspiel, 2003). The Oxford House post was established in 1798 by the Hudson's Bay Company as a supply depot (Hanks, 1982). It is located along the Hayes River at the northeast end of Oxford Lake (Hanks, 1982). The Hayes River had been an important trade route between Norway House and York Factory from the end of the seventeenth century until the end of the nineteenth century, when the trade routes shifted (Hanks, 1982; Herring and Sattenspiel, 2003). Although the importance of the post at Oxford House had diminished by 1918, trade continued and the post was still in direct contact with Norway House (Herring and Sattenspiel, 2003).

Figure 3.1: Map of Central Manitoba showing the locations named in the text



God's Lake, the smallest and most peripheral of the three study communities, had a population of 299 in 1917 (Herring and Sattenspiel, 2003). The Post was located on the western coast of God's Lake, which is connected by small rivers to larger waterways in the North and to Island Lake in the South. Less central land trade routes connected God's Lake to both Norway House and Oxford House (Herring and Sattenspiel, 2003) and travel between the three posts was still occurring in 1918.

By the early twentieth century, even though the fur trade was in a severe state of decline and the importance of Norway House, Oxford House, and God's Lake was dwindling, economic life in these three communities continued to revolve around the fur trade and the HBC (Herring, 1994; Sattenspiel et al., 2000). By the time of the 1918-1919 flu pandemic, more permanent communities had begun to grow up around the posts (Hallowell, 1992; Hanks, 1982) and women and children would sometimes spend the winter in these communities rather than at the winter traplines with the men (Hallowell, 1992). Yet despite these changes, the summer aggregation and winter dispersal pattern continued to be practiced up until the mid 20th century and is thought to have perhaps influenced the experience that these communities had with the 1918-1919 flu.

The 1918-1919 Flu Comes to the Canadian Subarctic

Abundant historical and archival data on Norway House, Oxford House, and God's Lake and their experiences with the 1918-1919 flu pandemic have been compiled and examined by Herring (Herring, 1994; Herring, 2000; Herring and Sattenspiel, 2003). Information on mobility patterns during the epidemic has been extracted from the

Hudson's Bay Company Post Journals for Norway House, Oxford House, and God's Lake (Herring, 1994; Herring and Sattenspiel, 2003; Sattenspiel and Herring, 2002). The HBC post journals served as a daily log of the business activities at each post and recorded the daily comings and goings of those who came to the posts to trade or obtain supplies. Information on the mortality associated with the pandemic has been extracted from several sources including burial registers for the Anglican Church of Canada Jack River Mission at Norway House (Herring, 1994; Sattenspiel and Herring, 2002; Sattenspiel and Herring, 2003; Sattenspiel et al., 2000), government census data (Sattenspiel and Herring, 2002), and the Treaty Annuity Pay Lists for 1918 and 1919 (Sattenspiel et al., 2000). The Treaty Annuity Pay lists record the annual payments that were made by the Canadian government to the members of First Nations groups that had entered into a governmental treaty and list the names of family heads as well as the number and names of the family members for whom they were accepting payments.

From this information, a timeline for the 1918-1919 influenza pandemic in this region has been derived. The deadly second wave of the 1918-1919 influenza pandemic appears to have first reached Canada in early September of 1918. The flu quickly spread across the country with troops returning home from World War I and reached Winnipeg, the capital of Manitoba, on Sept. 30, 1918 with a train carrying infected troops westward (Herring, 1994; Herring and Sattenspiel, 2003). In Canada, the 1918-1919 flu pandemic exhibited the same distinctive features that were seen elsewhere: extreme virulence; high mortality, especially among young adults; a distinctive purple tinge to the faces of the dying; and rapid and wide geographic spread. It has been estimated that about one in six Canadians were infected during the pandemic and that between 30,000 and 50,000 died

(Herring and Sattenspiel, 2003). However, these figures are somewhat misleading, as the impact of the epidemic, especially on Aboriginal Canadian communities such as Norway House, Oxford House, and God's Lake, was highly variable (Herring, 1994; Sattenspiel and Herring, 2002).

The 1918-1919 flu reached Norway House in early December and appears to have arrived with a mail packet from Cross Lake, where the flu was raging, on Dec, 4 (Herring, 1994; Herring and Sattenspiel, 2003). The first to contract the illness was an unlucky HBC chore boy, who became sick 5 days later (Herring, 1994). This timing would have been consistent with the 2-5 day incubation period for influenza (Hart, 2004). The flu struck swiftly and spread rapidly throughout the community, often being carried by traders who had traveled to the post to pick up supplies or trade furs back to their respective winter traplines. The first death occurred on December 18th and by January 6, 1919, 50 people had died (Herring, 1994). Schools and churches were closed, mail delivery was suspended, and daily life deteriorated as more and more people fell ill (Herring, 1994). The epidemic at Norway House peaked in January and faded by mid-February, only six weeks after its arrival (Herring, 1994). The impact of the epidemic was staggering: about 20% of the population had died, and many of the deaths were among those aged 1-14 (Herring, 1994). Burial of the dead was not possible until after the spring thaw, and the bodies of the flu victims were often wrapped in sheets and placed on rooftops or stacked in sheds to keep them out of the reach of dogs (Herring, 1994; Herring and Sattenspiel, 2003).

Although mortality due to the 1918-1919 flu pandemic was high in many small Northern aboriginal communities that were involved in the fur trade (Herring, 1994;

Herring, 2000), the death toll at Norway House was significantly higher than the estimated 3% mortality among Aboriginal Canadians as a whole (Herring, 1994; Sattenspiel and Herring, 2003). Further, the impact of the 1918-1919 flu at Norway House was highly variable, and some families experienced no illness while others lost many members or were wiped out entirely (Herring et al., 2003).

The remarkably high mortality rate at Norway House can be attributed to several related factors which may have weakened the population and made them less able to survive a major epidemic like the 1918-1919 influenza. To begin with, survival in the Canadian subarctic was already precarious without the added threat of epidemic disease. The Cree still derived the bulk of their income from the fur trade and relied heavily on foods that were hunted or gathered in the bush, as well as some staples that were traded-for at the post (Herring, 1994; Herring, 2000; Herring et al., 2003; Herring and Sattenspiel, 2003; Sattenspiel et al., 2000). There was little in the way of stored supplies (Herring, 2000; Herring and Sattenspiel, 2003) and starvation was a constant threat (Hanks, 1982), especially in winter when the post, the only reliable source of provisions, could be hundreds of miles away. With so little in the way of stored supplies, the arrival of an epidemic such as the 1918-1919 influenza could deal a tremendous blow.

In addition to the threat of starvation, malnutrition and poor health were also common and would have rendered the Cree less resistant to infections such as the 1918-1919 flu. The decline of the fur trade had led to an overall decrease in the standard of living, and a decline in the populations of game animals after about 1820 had led to increased competition for natural food resources and an increased reliance on nutritionally inferior European foodstuffs (Hanks, 1982). Tuberculosis was also endemic

in the region and may have more directly contributed to the high mortality at Norway House, as tuberculosis is believed to have increased the likelihood of contracting deadly secondary pneumonia, the main killer during the 1918-1919 pandemic (Herring and Sattenspiel, 2003).

The high mortality at Norway House may also be related to the timing of the epidemic. The 1918-1919 flu struck in winter, after the Cree had already dispersed to their winter traplines (Herring and Sattenspiel, 2003). Although low population densities in the wintertime would have inhibited the spread of the flu between hunting bands, the relative isolation of these groups would have made outside assistance unlikely if the flu did come, and close quarters would have encouraged the spread of the flu to all members of the group. Further, the 1918-1919 flu caused high mortality among young adults, who were the primary hunters and providers and, in the wintertime, were often the only source of provisions for their hunting band. The epidemic thus left entire families alone in the hinterland with nobody to hunt for them or to maintain fires (Herring, 2000; Herring and Sattenspiel, 2003), which likely led to increased mortality. Many may also have contracted pneumonia after going out into the cold to attempt to collect firewood or food (Herring, 2000; Herring and Sattenspiel, 2003). This situation probably led to increased overall mortality rates and may also help to explain why some families at Norway House were wiped out entirely while others lost no members or did not even experience the flu (Herring and Sattenspiel, 2003).

While the 1918-1919 flu epidemic clearly took a large toll on the community at Norway House, the flu appears to have never even reached Oxford House or God's Lake (Herring and Sattenspiel, 2003). There are a few documented cases of men from Oxford

House and God's Lake who died at Norway House during the epidemic (Herring, 1994), but no recorded deaths at the communities themselves. It is not surprising that Norway House was the most severely impacted by the epidemic because it was the largest of the three communities and was located on more major trade routes. What is surprising is that the flu failed to spread from Norway House to Oxford House or God's Lake, even though contact among the communities continued throughout the epidemic. The reasons for the absence of the flu at Oxford House and God's Lake are unclear, but such a varied pattern of epidemic spread was not unique to this region of Canada and appeared to have been a global feature of the 1918-1919 flu pandemic. In this project, the NHOHGL model will be used to investigate why the flu failed to spread in this particular case (see chapter 7) and it is hoped that any insights it provides will also prove to be more widely applicable.

Chapter 4: Epidemic Modeling and Models of the 1918-1919 Flu Epidemic in the Norway House Region of Canada

Models can be powerful tools for the study of epidemic disease, especially when they are based upon sound historical and epidemiological data. This chapter will provide a brief introduction to models and modeling as well as a discussion of two models that have previously been used to investigate the 1918-1919 flu epidemic in the Norway House Region of Canada. The chapter will then conclude with an introduction to the model that was used in this thesis project.

Models and Modeling

At the most basic level, a model can be defined as any object or concept that is used to represent something else (Sattenspiel, 2003). Models are almost limitless in type and can range from simple drawings to complex computer simulations. All models, regardless of type or complexity, are simplifications of a real-life system (Cross and Moscardini, 1985; Gilbert and Troitzsch, 1999; Sattenspiel, 2003). Because they contain only some of the essential elements of a real-life system, as determined by the researcher, models are not exact reproductions of reality and can be interpreted by different people in different ways (Cross and Moscardini, 1985; Gilbert and Troitzsch, 1999; Kohler, 1999; Sattenspiel, 2003).

Models can serve many purposes (Cross and Moscardini, 1985; Gilbert and Troitzsch, 1999; James and McDonald, 1981). Primarily, they are used to facilitate understanding of a complex problem or system by breaking it down into simpler

components. They can also provide insight into how these different components influence the system as a whole. Furthermore, models can be useful for theory development and testing, as even the process of building a model forces researchers to rationalize and make explicit their theories and assumptions regarding the system and its major components (Cross and Moscardini, 1985; Gilbert and Troitzsch, 1999; Kohler, 1999). Models can be used to study optimization, or how a particular system can operate most efficiently operate within a set of known limitations, and can provide insight into how to control and manipulate systems or processes. They can also prove useful for training or entertainment.

Any system or phenomenon can be modeled using a variety of approaches, and the choice of which method to use depends upon the researcher's experience with modeling, time and data limitations, and the question that the model is intended to address (Cross and Moscardini, 1985; Sattenspiel, 2003). Many different types of models have been used to study epidemic disease and historical epidemics. Epidemic disease models have been present in the literature for decades (Sattenspiel and Herring, 1998) and interest in disease modeling has become increasingly popular since the 1970s (Sattenspiel, 1990). Although epidemic models are quite diverse in form, they fall into three basic categories: statistical models, mathematical models, and computer simulation (Sattenspiel, 2003).

Statistical models begin with a real data set and use a series of statistical formulas in order to determine mathematical correlations between different measured variables and to uncover the sources of patterns in the data. For epidemic models, this data set is often historical data from a real epidemic in a real population (Sattenspiel, 2003). A number of

statistical approaches have previously been used to study historical epidemics including time series analysis, regression analysis, correlation analysis, spatial distribution analysis, and linear models (Sattenspiel, 2003). While these statistical models are useful for the identification of patterns in an existing data set, they often have less explanatory potential than other types of epidemic models.

Mathematical models, unlike statistical models, begin with a set of assumptions and observations about a system rather than a set of real data. Mathematical models consist of a set of mathematical equations that specify the relationships among a set of variables and which together provide a mathematical description of the problem, question, or hypothesis to be addressed by the model (Burghes et al., 1982; Cross and Moscardini, 1985; James and McDonald, 1981). These equations are used to generate an artificial data set that is meant to reproduce the patterns that would be found in a real data set. This type of model is useful for studying the interrelated aspects of a complex problem, determining the relative importance of different parameters, and identifying the connections between them (Sattenspiel, 2003). This makes mathematical models well-suited for study of conditions that influence the spread of epidemic diseases (Sattenspiel and Herring, 2002).

Like mathematical models, computer simulation models also model the underlying relationships among a series of variables rather than the statistical correlations between them (Gilbert and Troitzsch, 1999). Computer simulation models work in much the same way as mathematical models, using these assumptions to generate a realistic artificial data set. The primary difference is that computer simulation models rely upon computer programs rather than mathematical equations to generate the data.

One major advantage of computer simulation over mathematical or statistical modeling is that, while the other techniques generally assume large and/or homogeneous populations, computer simulations are often agent-based. Agent-based modeling is also known as individual-based modeling, distributed artificial intelligence, multi-agent modeling, and artificial societies (Epstein and Axtell, 1996; Gilbert and Troitzsch, 1999; Kohler, 1999). Agent-based models generally consist of three basic components: a population of individual actors or "agents," an environment or social space, and a set of rules for action and interaction among the agents (Epstein and Axtell, 1996; Sattenspiel, 2003). All action in an agent-based model takes place through the agents, which are simple, self-contained programs that collect information from their surroundings and use it to determine how to act (Gilbert and Troitzsch, 1999; Kohler, 1999). The agents can be assigned individual properties that can change during the course of the simulation if certain conditions are met or certain events occur. For example, in an epidemic model, a susceptible (S) agent can become infectious (I) if it both encounters an infectious agent and becomes infected. The environment serves as the context for the action in the model and can be programmed to mimic any social or environmental context. In many agent-based models, the agents are able to interact with and modify the environment as well as to interact with one another. The rules for all interaction in the model are specified by the programmer. These rules determine how the action in a model is to proceed and provide specifications on how the agents are to respond to encounters with other agents and to certain conditions of their environment. The rules for agent action and interaction must be carefully constructed in order to ensure that they are specific, realistic, and that they provide for all circumstances that may occur in the model.

When the model is run, the interplay between these three elements produces a dynamic artificial world that is designed to model a specific process or phenomenon. Agents are designed to interact "intelligently" with each other and with their environment (Gilbert and Troitzsch, 1999) and it is this interaction, as specified by the rules for interaction, that cumulatively produces the emergent behavior of the model. For this reason, agent-based models may be considered to have a "bottom-up" structure (Epstein and Axtell, 1996). Insight into the system or problem that the model is meant to simulate comes from examination of how the model develops over time and how varying the initial conditions of the agents and the environment affects the outcome of the simulation (Gilbert and Troitzsch, 1999).

Although computer simulation has been used widely in the social sciences since the 1960s, agent-based modeling only became popular in the early 1990s (Epstein and Axtell, 1996; Gilbert and Troitzsch, 1999). Most of the software and techniques were originally developed in other disciplines, especially the field of artificial intelligence and the subfield of distributed artificial intelligence, which is concerned with the properties and design of networks of interacting agents (Gilbert and Troitzsch, 1999). These programs and techniques have been found to show tremendous potential for aiding in social scientific research (Dean et al., 1999; Epstein and Axtell, 1996; Gilbert and Troitzsch, 1999; Kohler, 1999; Sattenspiel, 2003) and over the past ten to fifteen years, agent-based models have grown increasingly popular in the social sciences. For example, agent-based models have been used in archaeology to model population growth and collapse (Axtell et al., 2002) and the impact of the availability of natural resources on settlement patterns and culture change (Dean et al., 1999; Kohler et al., 1999). Agent-

based models have also been used to model the evolution and influence of various social and spatial processes in human and primate societies (Kohler and Gumerman, 1999) and, more rarely, to study epidemic disease (Barrett et al., 2005; Carpenter, 2004).

Agent-based computer simulation is becoming popular due to the advantages that it possesses over other modeling techniques. The primary advantage of the agent-based approach is that, because each agent is individually constructed, each can be assigned its own individual properties and be treated differently during the simulation (Epstein and Axtell, 1996; Kohler, 1999). This feature allows agent-based models to better replicate small or diverse populations than population-based models, like mathematical and statistical models, which assume that all individuals of a population have identical characteristics (Sattenspiel, 2003). Agent-based modeling also provides a practical way to study systems that involve many coevolutionary interactions, which are difficult to analyze using traditional social science methods (Kohler, 1999) and can uniquely illustrate how simple local-level rules and elements can generate complex population-level phenomena (Epstein and Axtell, 1996; Gilbert and Troitzsch, 1999; Kohler, 1999). Additionally, tight control over the properties of the agents and the landscape can allow for a closer approximation of a complex modern or historical environment (Kohler, 1999). Another advantage of agent-based models is that they allow for controlled experiments on the effects of varying individual parameters in realistic human populations that would otherwise be impossible, unethical, or expensive (Epstein and Axtell, 1996; Gilbert and Troitzsch, 1999; Sattenspiel, 2003). It is these advantages that make agent-based models a promising tool for the study of epidemic disease and have led to its use in this project.

Models of the 1918-1919 Flu Epidemic in the Norway House Region of Canada

Two models have previously been created in order to investigate the 1918-1919 flu epidemic in the Norway House region of Canada: the Norway House Ordinary Differential Equation model and the Norway House agent-based model. Although these models employ different modeling techniques, both incorporate historical data from the HBC post journals and Anglican Church of Canada burial records, ethnographic data on the Cree and Métis people, and epidemiological data on the 1918-1919 flu pandemic into their overall design and both have provided important insights into the 1918-1919 flu epidemic in this region of Canada. These models served as the forerunners for the NHOHGL model project and will thus be briefly described in the following paragraphs.

- **The Norway House Ordinary Differential Equation (NHODE) Model**

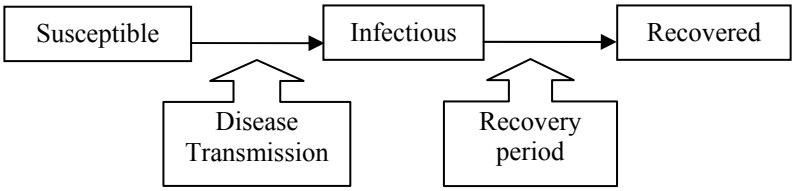
Previous work by Sattenspiel and colleagues has utilized mathematical modeling to study the 1918-1919 influenza epidemic at Norway House, Oxford House, and God's Lake (Sattenspiel and Dietz, 1995; Sattenspiel, 1990; Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002; Sattenspiel and Herring, 2003; Sattenspiel et al., 2000). As Sattenspiel and Herring (2002) note, the 1918-1919 influenza epidemic in the Norway House region of Manitoba is an ideal subject for this type of modeling because of the quality of the existing historical records and because the region is small and is representative of the situation throughout much of the Canadian subarctic. Data from the historical, ethnographic, and epidemiological record were used in construction of the model and in deriving parameter values (Sattenspiel, 1990; Sattenspiel and Herring, 2002) in order to make the model environment resemble reality in this region at the time

of the 1918-1919 pandemic. Furthermore, the historical data served as a means of testing the effectiveness of different epidemiological scenarios in explaining the patterns exhibited by the epidemic (Sattenspiel and Herring, 2002) because it provided a set of real values to which the artificial values produced by the model could be compared.

Because it uses a series of differential equations to model the 1918-1919 flu epidemic at Norway House, Oxford House, and God's Lake, it is known as the Norway House Ordinary Differential Equation (NHODE) model. Sattenspiel's NHODE model is a three-group mathematical model that combines a simple mobility model with a standard SIR epidemic model (Sattenspiel and Dietz, 1995; Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002; Sattenspiel and Herring, 2003; Sattenspiel et al., 2000).

In an SIR epidemic model (see Figure 4.1) the model population includes three types of people who correspond to three distinct stages in the infection process: susceptible (S) individuals, who are at risk for contracting the disease; infectious (I) individuals, who have contracted the disease and are capable of transmitting it to susceptibles; and recovered (R) individuals, who are no longer able to transmit the disease to others. At the beginning of the simulation, all of the agents are susceptible. When an infectious individual is added to the population, a proportion of the susceptibles may contract the disease and become infectious. These infectious individuals then infect other susceptibles and the cycle continues until the disease can no longer be transmitted and the epidemic ends. Infectious individuals have only a finite amount of time during which they can transfer the disease to others, a period that corresponds to the infectious/ recovery period for the disease. Once this period has elapsed, the infectious individuals become recovered and are generally assumed to be immune to reinfection. In some

Figure 4.1: Plan for a basic SIR epidemic model
Diagram shows the transition from one stage of infection to the next, with the parameters controlling each stage change shown in the arrow boxes.



models, however, recovered individuals are allowed to return to a susceptible state after a specified period of time (Sattenspiel, 1990). This process of infection and recovery is monitored as the model runs and produces statistics for the simulated epidemic that are similar to what would be produced by a real epidemic.

The mobility portion of the NHODE model simulates the movements of the human population, taking into account the probability of contact between individuals. Disease transmission is specified as a function of both the number of contacts per unit of time and the probability of disease transmission during a contact (Sattenspiel and Herring, 2003). The SIR portion of the model simulates the progression of influenza through its biological stages, taking into account the rates or probability of progression from one stage to the next (Sattenspiel, 2003; Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002; Sattenspiel et al., 2000).

Sattenspiel's NHODE model has provided insight into many of the factors which may have influenced the spread, timing, and intensity of the 1918-1919 influenza pandemic. These include: the influence of contact rates on epidemic severity (Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002); the impact of mobility patterns on the

epidemic's timing and extent of spread (Sattenspiel and Herring, 2002; Sattenspiel et al., 2000); the possible impact of quarantine (Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2003); and the possible impact of the seasonal pattern of residence that was practiced in this region (Herring and Sattenspiel, 2003).

Yet despite these insights, the NHODE model has been found to have two fairly significant limitations. First, the model is population-based, and thus assumes that all of the individuals in the model are identical (Sattenspiel, 2003). This means that the NHODE model is unable to account for individual heterogeneity. Second, the model is deterministic. In deterministic models, parameters which tend to vary in real populations are assumed to be consistent at some mean value that is estimated from the data (Sattenspiel, 2003). This means that each time the model is run under a specific set of parameter values, the result is the same. It also means that the model does not take into account stochastic fluctuations in model parameters or population size (Sattenspiel and Herring, 1998). This would not be a problem for a large and homogeneous population, but proves problematic for the study of small communities such as Norway House, Oxford House, and God's Lake where the random nature of real world events could have had a large impact on the outcome of an epidemic such as the 1918-1919 flu (Sattenspiel, 2003; Sattenspiel and Herring, 1998).

- **The Norway House Agent-Based Model**

Recognizing the limitations of Sattenspiel's NHODE model, it became clear that a shift to computer simulation, and more specifically to an agent-based modeling approach, was needed in order to more accurately model the 1918-1919 flu in the Norway House

region of Manitoba. This led to the creation of the Norway House agent-based model (Carpenter, 2004). The Norway House model is a stochastic, agent-based computer simulation that was designed to simulate the spread of the 1918-1919 influenza epidemic within the community of Norway House. Like the NHODE model, it combines a simple epidemic model with a mobility model and has incorporated historical, ethnographic, and epidemiological data into its design.

The Norway House agent-based model overcame the two primary limitations of Sattenspiel's NHODE model. First, because it is individual- rather than population-based, the Norway House model was better able to account for population diversity. Second, because it is stochastic rather than deterministic, the Norway House model was better able to account for the influence of random events. The Norway House model also has additional advantages in that it incorporates a more realistic spatial representation of the Norway House community and a more realistic simulation of the movements of its population than was possible in Sattenspiel's NHODE model. Furthermore, the Norway House agent-based model has added stages to the basic SIR epidemic sequence that allow it to more accurately represent the epidemiology of the flu virus. These include an exposed (E) stage, which is meant to represent the latent period of influenza, and, in a later version of the program, a dead (D) stage, which introduced mortality into the epidemic simulations.

The Norway House model was used to investigate two primary research questions: first, how did the seasonal population movements that were practiced by the Cree and Métis influence the spread of the 1918-1919 flu through the Norway House community and second, how would a summer epidemic differ from a winter epidemic?

The simulation data suggested that the traditional seasonal movement patterns did have a strong impact on the spread of the 1918-1919 flu and that a summer epidemic would have been more severe than a winter epidemic (Carpenter, 2004). Although these findings were informative and suggestive of factors that may have influenced the spread of the 1918-1919 flu within a community, the Norway House model was unable to simulate the spread of the epidemic between communities because it models only the Norway House community. This limited the scope of the model and also meant that the results of the simulations could not be directly compared to those of Sattenspiel's mathematical model, which simulates the 1918-1919 flu epidemic at Norway House, Oxford House, and God's Lake.

Objectives of This Project

This project represents an attempt to overcome the limitations of the Norway House agent-based model by significantly expanding its capabilities. The goal of this project is to extend the Norway House model by adding two additional communities to represent Oxford House and God's Lake and by creating a new model function to replicate inter-community mobility. The resulting model would have the same geographic focus as Sattenspiel's NHODE model and would be able to model the spread of the 1918-1919 flu epidemic both within as well as among the three communities, but would retain the advantages of the agent-based modeling technique for the study of small populations. Table 4.1 provides a comparison of the basic features of the three models discussed in this chapter. Details on the construction of the new Norway House, Oxford House, God's Lake (NHOHGL) model will be presented in the following chapter.

Table 4.1: A basic comparison of the three models discussed in the text

	NHODE	Norway House	NHOHGL
Creator	Sattenspiel	Carpenter	Ahillen
Epidemic modeled	1918-1919 flu	1918-1919 flu	1918-1919 flu
Communities Modeled	Norway House Oxford House God's Lake	Norway House	Norway House Oxford House God's Lake
Modeling Technique	Mathematical	Computer Simulation	Computer Simulation
Basic Format	Population-based	Individual-based	Individual-based
Treatment of Random Events	Deterministic	Stochastic	Stochastic
Epidemic Pattern	SIR	SEIR *	SEIRD

*Death was added in a later version of the model: the Norway House Mortality model

Chapter 5: Design and Construction of the NHOHGL Model

The primary goal in developing the NHOHGL agent-based model was to create a stochastic, individual-based model to study the spread of the 1918-1919 influenza epidemic both within and among the communities of Norway House, Oxford House, and God's Lake. This required that Carpenter's basic Norway House model (Carpenter, 2004) be extended to include the communities of Oxford House and God's Lake and that an additional level of mobility be created to allow the agents to move between the three communities. This, in turn, required fundamental modifications to the original program design and code as well as the addition of several new parameters. Changes were also made to enhance the speed of the simulations, to allow the model to be more easily manipulated, and to facilitate future expansion of the model. This chapter will discuss the design and construction of the new NHOHGL model, how it functions, and the basic objectives of its use in this project

Design of the NHOHGL Model

The development of any agent-based model should begin with a basic plan or outline which often takes the form of one or more visual models (Gilbert and Troitzsch, 1999). These models serve as a guideline for model construction and help the researcher to solidify his or her views and assumptions regarding the phenomenon to be modeled and the relevant parameters. For agent-based models, three elements have to be considered: the agents, the landscape, and the rules to govern agent action and interaction. In the NHOHGL model, the design of each of these elements rests heavily

upon their design in the earlier Norway House model, and wherever possible, data from the historical, ethnographic, and epidemiological record are incorporated in order to provide a more realistic representation of the communities, the Cree and Métis people, and the 1918-1919 flu epidemic.

The design for the agents in the NHOHGL model is nearly identical to their design in the original Norway House agent-based model. In both models, the size and composition of the agent population is controlled by three parameters. The first, population size, specifies the number of agents in the simulation and determines their initial location on the landscape. The second, age/gender composition, assigns each agent to one of three groups: adult males (aged 20-50 years); adult females (aged 20-50); and children/elderly (under 20 or older than 50). These groups are differentiated in the model due to the unique age preferences of influenza and the 1918-1919 influenza pandemic (Crawford, 2000; Crosby, 2003; Karlen, 1995; Taubenberger et al., 2005a; Taubenberger et al., 2005b). The age/gender groups also have important implications for mobility because the model allows only adult males to travel. The third agent parameter, number of cliques, divides the agent population into smaller subgroups that are designed to represent the basic social unit of Cree society, the extended family hunting band. While the NHOHGL model retains the same basic agent design used in the Norway House agent-based model, it also allows for these three agent parameters to be independently altered for the agents at Norway House, Oxford House, and God's Lake, and thus includes three sets of agent parameters rather than one.

The design for the landscape in the NHOHGL model is also based upon its design in the original Norway House model. The Norway House model landscape is a

simplified version of the Norway House community around the time of the 1918-1919 influenza pandemic that was created using historical and ethnographic data. It consists of a central post, four outlying camps, and four paths connecting the camps with the post (see Figure 5.1). The fort represents the HBC trading post at Norway House while the camps represent the traditional hunting territories that were occupied by the Cree extended-family hunting bands in the wintertime. The paths represent the trails followed by the traders as they traveled between their camps and the HBC post. In reality, there would have been many more camps, one for each hunting band, and numerous paths linking these camps to the central post, but these numbers were reduced in the model for the sake of simplicity and feasibility of construction. Overall, this landscape is relatively simple and, at this time, does not simulate any physical environmental conditions, constraints, or topography, although this could potentially be added in the future.

The objectives of the NHOHGL model required that this basic landscape design be significantly expanded. The structure of the Norway House community in the original Norway House model has been retained in the NHOHGL model but has also been duplicated to produce two additional and structurally identical communities that represent Oxford House and God's Lake (see Figure 5.2). An effort was made to place the three communities at realistic distances from one another on the model landscape, but this was hindered by the fact that distance in the model is measured by the journey time for the agents (in days) rather than by geographic distance. Further, the time required to travel between any of the three communities would have varied by season and the only estimates for the duration of such a journey, from the HBC post journals, are for the trip between Norway House and Oxford House (Sattenspiel and Herring, 1998).

Figure 5.1: A visual model of the Norway House model landscape
Diagram shows the central post, the four outlying camps, and the paths connecting these camps to the post. Note that the four camps are set at varying distances from the central post. This basic community structure was retained in the NHOHGL model.

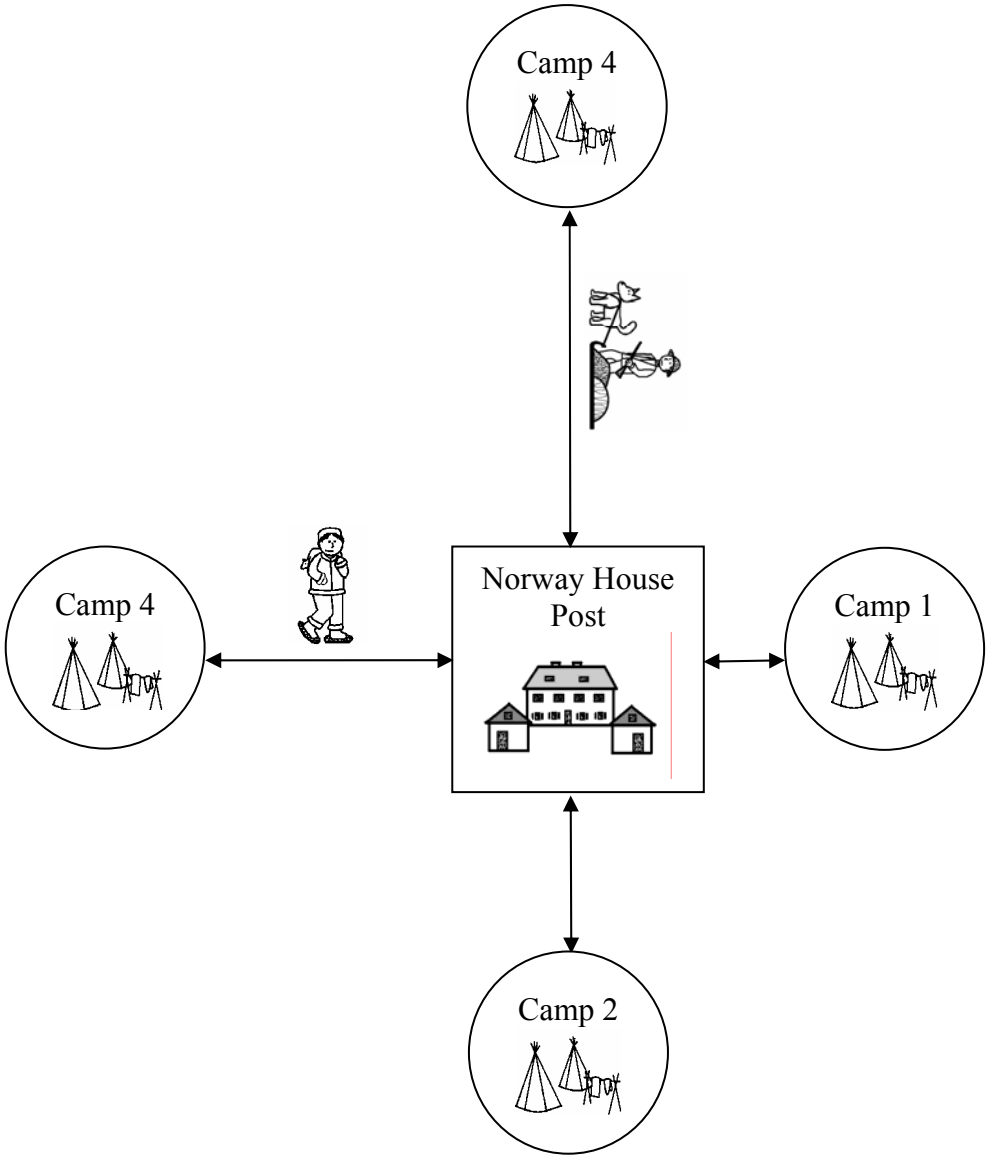
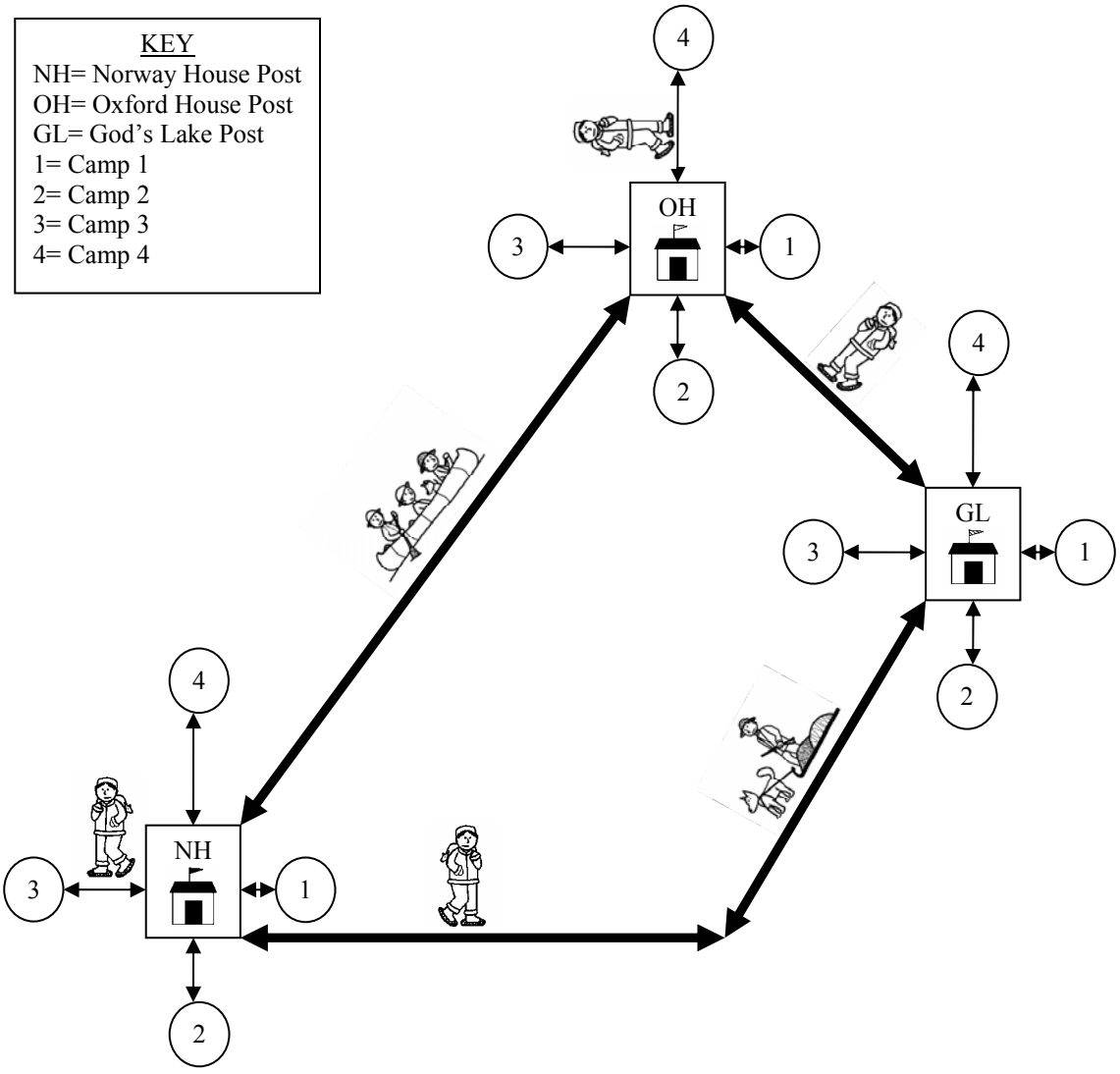


Figure 5.2: A visual model of the NHOHGL model landscape
Diagram shows the three communities, Norway House, Oxford House, and God's Lake, each with a central post and four camps, and the inter-community pathways that connect them to one another.



This same issue also proved to be problematic in the creation of the paths to connect the three communities. These inter-community paths were designed to represent the rivers and trails that linked the HBC posts at Norway House, Oxford House, and God's Lake. Because the length of the journey between each of the posts would have been seasonally variable, the design of these paths had to allow the length of the journey to be not only realistic, but also easily altered. For the paths within each community, travel time could be altered by simply lengthening or shortening the paths that connected the camps to the central post. This worked because only two points were being connected, but for inter-community paths, three points had to be connected. This meant that each time a path distance was to be changed, all three communities had to be moved and the other two paths had to be adjusted accordingly in order to allow all three communities to remain connected. Such changes would have been extraordinarily complicated and would likely have made changing the distances between the three communities impossible.

The solution to this dilemma came with the creation of a new parameter: steps-per-day. The steps-per-day parameter enables the agents to travel more than one space on the path per day, allowing them to cover the same distance on the model landscape at a range of speeds (see Table 5.1). This permits the rate of travel, and thereby distance, to be altered without requiring major revisions to the program each time a change is to be made. The steps-per-day parameter also helps to reduce user error by allowing all three inter-community paths to be changed simultaneously.

The design for the rules to govern agent action and interaction in the NHOHGL model, like the designs for the agents and the landscape, was built upon the foundation

Table 5.1: The steps-per-day parameter: its possible values and impact on travel times between the three communities

Value:	Travel Time (in days) between posts:		
	NH-OH	OH-GL	GL-NH
1	24	12	36
2	12	6	18
4	6	3	9
6	4	2	6
12	2	1	3

established in the original Norway House model described in Carpenter (2004). Both models include three distinct types of rules: mobility rules, rules for agent social interaction, and epidemic rules. Together, these rules control both the actions of the agents and the behavior of the simulated epidemic. All three types of rules had to be reevaluated and significantly extended in the NHOHGL model.

Although significant additions have been made to the mobility rules in the NHOHGL model, the rules controlling agent mobility within each community remain essentially identical to those in the original Norway House model. The key difference is that now these rules are applied to three communities rather than to just one. In both models, the movements of the agents within a community are controlled by two parameters that can be independently adjusted for the three communities in the NHOHGL model. The camp stay parameter defines the probability that the agents will stay at their home camp rather than travel to the post, and is essentially the opposite of a probability of movement. The probability of staying on the path parameter, in turn, defines the probability that the agents will follow the predetermined paths to the post rather than heading off in any direction.

Along with these parameters, a set of assumptions that are built into the model structure also guide the movements of the agents within each community. In both models, it is assumed that only the agents from the camps may travel between the camps and the post, that only adult males may make these journeys, that they travel alone, and that they stay at the post for just one day before returning home. These assumptions have been based upon ethnographic and historical data, but, as this is a model, are necessarily simplifications of historical reality. For example, although the trappers would have been the primary travelers, periodically journeying from their winter camps to the posts to trade or obtain supplies (Herring, 1994), it is quite possible that the year-round residents of the posts, who would have generally been HBC employees, may have occasionally journeyed to the camps. Further, even though ethnographic data suggest that the males were the primary travelers (Herring, 1994) it is also known that Cree women often took on male roles, especially in times of hardship or disease (Flannery, 1995) and thus, may have occasionally made the journey to the post. It is also quite clear that the males rarely traveled alone and, although their trips would generally have been brief because the survival of the family group often depended on their prompt return with supplies, some men may have stayed longer than one day at the post before heading back to their home camps. In evaluation of the data produced by the NHOHGL model, it is important to keep these simplifications in mind, as they may lead to inaccuracies in the predictions of the model.

While the model for mobility within a community in the NHOHGL model was adapted from the original Norway House model, a new model for inter-community mobility had to be developed. The post agents, who were unable to travel in the original

Norway House model, are the sole intercommunity travelers in the NHOHGL model and have the ability to travel from their home post to either of the other two posts. This is roughly consistent with ethnographic data which suggest that most of the movements between the three communities would have been by the European, native or Métis HBC employees who generally lived near the posts (Hanks, 1982). The model dictates that such travel only occurs along the specified pathways and can only be from an agent's home post and back. Travel to more than one community on a single trip is not allowed. Also, as for mobility within a community, only adult males travel, they travel alone, and they stay just one day at their destination before returning home. These rules are based upon historical data which indicate that most of the travel between communities would have been to move furs and supplies between the posts (Hanks, 1982), but, like the other model rules and assumptions, this structure is a simplification of a much more complex historical picture.

The addition of inter-community mobility in the NHOHGL model also required the creation of a new mobility function, as well as the addition of several new parameters. In the original Norway House model, the agents had only to decide whether or not to travel, but in the NHOHGL model, the post agents had to decide not only whether to travel, but also where to travel. A new mobility function was developed to enable them to make this choice, and a series of six parameters was added to the model in order to specify the probability of travel between each of the three posts: NH to OH move probability; NH to GL move probability; OH to NH move probability; OH to GL move probability; GL to NH move probability; and GL to OH move probability. These

parameters allow the frequency of travel to be independently altered for each of the possible journeys.

The second set of rules that had to be included in the design of the NHOHGL model were rules to govern interaction among the agents. These rules remain the same as in the original Norway House model and are based upon the distinctive social organization of the Cree/Métis people. Because Cree society revolved around the extended family hunting band, the model assumes that the probability of contact would have been higher for members of the same group and lower for members of different family groups. Thus, contact between the agents is specified by two parameters in the NHOHGL model: probability of contact within a family group and probability of contact outside of a family group. These parameters can be set at different values in order to replicate the differing contact rates within and outside of a family group.

The third and final set of rules designed for the NHOHGL model were rules to govern the spread of influenza among the agents. The epidemic sequence used in the NHOHGL model is simply an extended version of a basic SIR pattern like that used in Sattenspiel's mathematical models (see Figure 4.1). The NHOHGL model has added two additional disease states to the basic SIR sequence: exposed (E) and dead (D). In the resulting SEIRD epidemic sequence, exposed agents are those who have been infected with the flu but have not yet become infectious, and death provides an alternative to recovery as the final stage in the infection process.

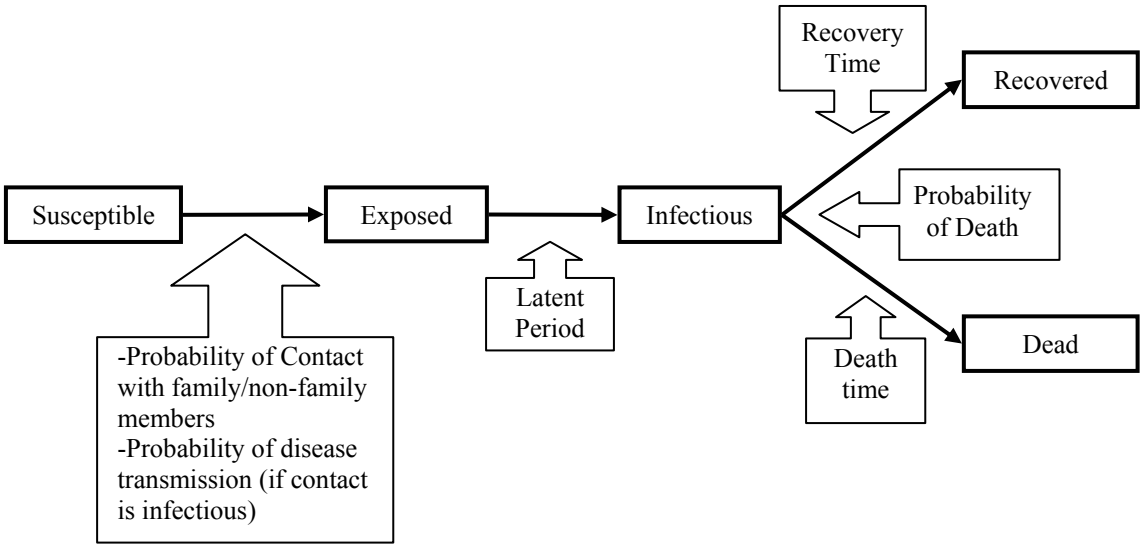
An exposed stage was first included in the original Norway House agent-based model and corresponds to the latent period for a disease. It has been set to last for one day in the NHOHGL model. This may be a low estimate for influenza, but the latent

period for any disease is extremely difficult to measure and one day appears to be a reasonable estimate for the 1918-1919 flu, which, in severe cases, was known to cause death in less than 48 hours (Crosby, 2003).

Death is a more recent addition to the model's epidemic sequence and was only added following the completion of Carpenter's thesis. The addition of death required the creation of two new parameters: probability of death (upon infection) and death time, or the average length of time that an agent is infected before death occurs. Death has been retained in the NHOHGL model, because it allows the simulation data to be more readily compared to historical data for the 1918-1919 influenza epidemic, which tend to record its impact in terms of the number of lives lost rather than the number of recovered individuals. However, death can also be disabled in the NHOHGL model by adjusting the death probability to zero, and simulations were run both with and without this parameter.

The transition of the agents from one disease state to another depends both upon their contacts with other agents and the values of a set of epidemic parameters. Following the SEIRD pattern (see Figure 5.3), the initially susceptible agents may become infected if they come into contact with an infectious agent. When such a contact occurs, the probability of infection parameter determines whether or not the flu is transmitted to the susceptible agent. Upon infection, an agent first becomes exposed and then, after the latent period has elapsed, infectious. Infectious agents may spread the flu to those they encounter as they traverse the model landscape, and the NHOHGL model assumes that agents still travel while infected. This assumption is based upon data which suggest that rates of travel in this region did not slow during the 1918-1919 flu epidemic

Figure 5.3: Plan for the SEIRD epidemic model used in the NHOHGL project
Diagram shows the transition from one stage of infection to the next, with the parameters controlling each stage change shown in the arrow boxes.



(Herring, 1994). Infectious agents may continue to spread the flu until they either recover or die. The length of the infectious period is determined by the value of the recovery time parameter. Infectious agents will automatically become recovered at the end of this period, assuming that they manage to survive, and will then be immune to reinfection. Death may occur during a specified portion of the infectious period, which is controlled by the death time parameter, and occurs only when triggered by the probability of death parameter. Dead agents are essentially made inactive and are unable to travel or to transmit the flu for the rest of the simulation.

In addition to the assumptions and parameters that have been included in the design of the agents, landscape, and epidemic, additional parameters have also been included in the NHOHGL model in order to allow the researcher control over the simulation process. Several of these parameters were also present in the original Norway

House model. These include a parameter to specify the number of days per run of the simulation, which was usually set at 200 to provide the epidemic with sufficient time to run its course, and another parameter which specifies the number of simulations to be run per parameter change. This is important because the model is stochastic, and thus each run of the model can be different even if the parameters are held at constant levels. In order to analyze a stochastic model such as this, several runs of the model at each set of parameters have to be performed and then the results averaged in order to provide a better sense of the behavior of the model (Kohler et al., 1999). In this project, the number of runs was usually set at 1000, an amount that was used in the original Norway House agent-based model (Carpenter, 2004).

Several new simulation control parameters were also added to the NHOHGL model to give the researcher greater control over the model output. Generally, as the model runs, the program is used to generate six different displays. One displays the landscape with the posts and camps and shows the movements of the agents as they travel about, another displays the running output of the model, and the others produce graphs of the epidemic displaying the number of susceptible, infected, recovered, and dead agents and updating this for each day of the simulation for the entire population, Norway House, Oxford House, and God's Lake, respectively. Although these displays can be entertaining and informative, they make the simulations slow significantly and hinder timely data collection. Three parameters: screen output, graphs, and world display, were added to the NHOHGL model to allow these displays to be turned off when displays are not needed, thus allowing the model to run faster during data collection.

The NHOHGL Model Program

Once the basic design for the agents, landscape, and rules for agent mobility, interaction, and the epidemic sequence had been established for the NHOHGL model, construction of the new model program could begin. Like its design, the programming code for the NHOHGL model was adapted from that of the original Norway House model. However, the addition of the two additional posts and intercommunity mobility in the NHOHGL model required that substantial additions be made to the original program code. The result of these alterations is the new NHOHGL model program, a copy of which can be found in appendix A.

The NHOHGL model is composed of three program segments known as classes. The first, the model class, is the main class for the NHOHGL model and stores all of the code for the set-up and operation of the model and for the data and display output. The second, the agents class, keeps track of the information for the agents including their age/gender group, family group, and, for the post agents, the location of their home post so that they will know which direction to travel when returning from a journey to another post. The landscape class, which was called the "paths" class in the original Norway House model, is the smallest class and merely sets up some of the model parameters. In addition to these program classes, the NHOHGL model also includes a text file which is linked to the program and allows the researcher to change the model parameters on this single form rather than having to locate them individually in the program code. This feature is a recent addition to model and has made it much more user friendly.

The NHOHGL model is written in the same computer language as the original Norway House model and utilizes the same agent-based modeling toolkit and the same

running environment. Although there are many computer languages, modeling toolkits, and running environments to choose from, these programs had performed reasonably well for the previous model and a change was thus deemed unnecessary at this time.

The program code for both of the models was written in the Java™ computer language, which was selected for its straightforwardness and its compatibility with agent-based simulation (Carpenter, 2004; Gilbert and Troitzsch, 1999). Java™ is a simple and dynamic computer language that is object-oriented and able to be run on several platforms (Campione and Walrath, 2005). It is powerful and easier to learn than many other computer languages; it allows the program code to be more concise and more quickly composed, facilitates debugging, and can be run on several different platforms (Campione and Walrath, 2005). These attributes make Java™ an ideal choice for agent-based modeling.

While the basic operating programs for both the original Norway House and the NHOHGL models were composed in Java™, an agent-based modeling toolkit known as RePast was used to create the basic framework for models, run the simulations, and collect and display the data (Carpenter, 2004). RePast (Recursive Porous Agent Simulation Toolkit) is an agent-based modeling toolkit that was developed by the social science research computing team at the University of Chicago. It is a free and open source toolkit that is designed to support and facilitate the construction of "highly flexible models of living social agents" (OSTG, 2005). It utilizes the Java™ computer language and like Java™, can be run on a variety of platforms. It is also more user-friendly than some of the other available agent-based modeling toolkits, making it a good choice for first-time modelers.

Although both Java™ and RePast are compatible with a number of running environments, both the original Norway House model and the NHOHGL model utilized the JBuilder platform. JBuilder is a leading Java IDE that allows for speed coding and debugging of Java™ based programs (Borland, 2005). It was selected because it is free to download and because it includes a powerful source code editor and debugger which flags errors in the code to allow problems in the program to be more easily identified and corrected. Although JBuilder takes some time to become familiar with, it proved to be sufficient for our needs at this time.

Stochasticity, one of the main advantages of the NHOHGL model over Sattenspiel's NHODE model, is replicated through the use of a combination of probability functions and a random number generator to determine the course of agent action. The model parameters serve as the probability functions for the model. While their value is set by the researcher, they specify only the probability that a certain action, choice, or event will occur under certain conditions. When these conditions arise, the actual "choice" of action is determined by drawing a number from a Java™ pseudo-random number generator (Carpenter, 2004). This number is then compared to the probability parameter to determine the agent's course of action. If the number drawn is less than the specified probability, the agent performs the action, if greater, no action is taken. This process lends a degree of randomness to the actions of the agents that was not possible in Sattenspiel's mathematical models.

This means of replicating stochasticity worked reasonably well for the Norway House model and for most of the parameters in the new NHOHGL model, but a problem arose with the addition of inter-community mobility. Unlike the other parameters, which

require that a choice be made between two alternatives (e.g., stay or travel; susceptible or infectious), the new inter-community mobility function required a choice of three options (Stay, travel to post A, or travel to post B). Modifications had to be made to the basic structure of the probability function in order to accommodate this additional choice. To do this, the two possible inter-community movement probabilities (home post to post A or home post to post B) were compounded in the NHOHGL model. If the random number generator selects a value that is greater than 0 and less than the first probability, the agent travels to the first location. If the number drawn is greater than the first probability, but less than the sum of the first and the second probabilities, then the agent travels to the second location. If the number drawn is greater than the sum of the two probabilities, the agent stays. This system allows a choice of direction to be made by the agent while preserving stochasticity.

Running the NHOHGL Model

In the NHOHGL model, the design of the agents, landscape, and rules governing agent mobility, agent interaction, and the epidemic model combine to produce a simulation of the 1918-1919 flu epidemic within and among the communities of Norway House, Oxford House, and God's Lake, Manitoba. A summary of the basic structure and assumptions for the final NHOHGL model can be found in Table 5.2. When the model is initialized, the computer generates the simulation landscape and places the agents upon it according to the values of the population parameters. Each agent is assigned to an age/gender group and to a family group and their initial disease state (susceptible) is recorded. On the first day of the simulation, the agents begin to move around the

Table 5.2: The basic structure and assumptions for the NHOHGL model

- Basic Structure:
  - 1000 simulation trials are run for each parameter change
  - Each simulation runs for 200 days
  - The flu is introduced to the NH post on the 20th day of the simulation
  - The total landscape consists of a 100x100 grid
  - The inputs for the simulation consist of the values chosen for the parameters
  - The output of the simulation consists of statistics for the simulated epidemic:
    - Duration of the epidemic
    - Peak day of the epidemic
    - Peak day for mortality during the epidemic
    - Peak number infected
    - Peak number dead
    - Total number of cases
    - Total number of deaths
    - Percent of the population that became infected
    - Percent of the population that died
    - Number of times the epidemic reached Oxford House, God's Lake, or all three communities

    ~These statistics are compiled for each community and for the population as a whole

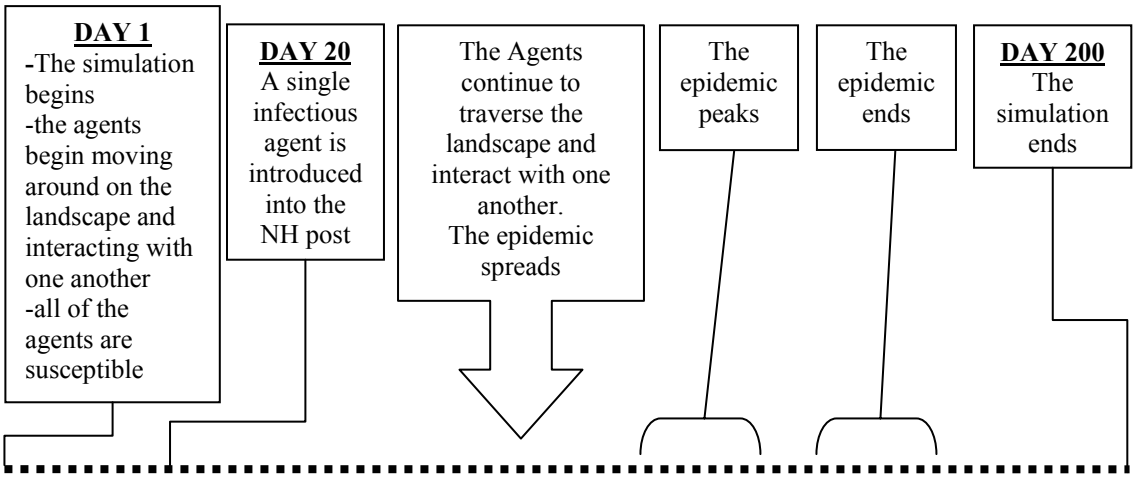
- Basic Assumptions:
  - No births
  - No deaths other than those caused by influenza
  - Each agent belongs to a hunting band or "clique"
  - Residents of the camps may periodically travel back and forth to their designated post
  - Residents of the posts may periodically travel from their home post to another post and then return. They can not travel to more than one post on a single trip
  - Only males can travel
  - Males travel alone and not in groups
  - When males travel to a new location, they stay for one day and then return
  - Agents continue to travel even if they are infected, but cease to travel upon death

landscape and to interact with one another according to the rules and parameters that govern agent mobility and contact. For each subsequent day, the computer updates and records both the location and disease status of each agent and determines his or her next course of action. On the twentieth day of the simulation, after the agents have had sufficient time to begin moving about, a single infectious agent is introduced into the Norway House post (this is where historical data suggest the 1918-1919 flu entered the Norway House region (Herring, 1994; Herring and Sattenspiel, 2003)). This infectious agent may potentially infect the susceptible agents it encounters, who may then become infectious and infect other susceptible agents, thus spreading the epidemic. The flu continues to spread until there are either no more infectious agents to spread the disease or until everyone has been infected and no more susceptibles remain. Each simulation is run for a period of 200 days in order to ensure that the simulated epidemic will have sufficient time to build, peak and decline. A basic visual model for this process may be found in Figure 5.4.

- **The Input Parameters**

In computer simulations like the NHOHGL model, the inputs for the model consist of the parameters that are needed to make it parallel a particular social setting, while the outputs are the behavior of the model through simulated time (Gilbert and Troitzsch, 1999). Insight into the phenomenon or system being modeled comes from observing the effects that varying the values of the initial parameters has on the outcome of the model. In the NHOHGL model, the inputs for the model consist of the values of the initial agent, landscape, mobility, epidemic, and contact parameters. For this project,

Figure 5.4: A basic visual plan/timeline for the spread of the influenza virus in the NHOHGL model



three sets of initial input parameters were developed for the purpose of testing and evaluating the new model: a set of standard parameters, a set of winter parameters, and a set of summer parameters. A list of the primary NHOHGL model input parameters with their standard, winter, and summer values can be found in Table 5.3.

The standard parameters were designed to serve as the default or control values for the model, and the simulated epidemic that they produced was used as a baseline for comparison with the data from the test simulations. Like the design for the model structure, the standard parameter values used in the NHOHGL model simulations are based, wherever possible, upon data from the historical, ethnographic, and epidemiological records. This was essential to the accuracy of the model in simulating the 1918-1919 influenza epidemic in this region, because a model is only as good as the data used to estimate its parameters (Sattenspiel, 2003).

Table 5.3: Summary of the basic parameter values used in the standard, winter, and summer NHOHGL model simulations

Parameter	Standard	Winter	Summer
Number of Runs	1000	1000	1000
Number of Time Steps/Days	200	200	200
Displays	0 (off)	0 (off)	0 (off)
Post Population Norway House	150	150	750
Post Population Oxford House	66	66	330
Post Population God's Lake	60	60	300
Camp Populations Norway House	150	150	0
Camp Populations Oxford House	66	66	0
Camp Populations God's Lake	60	60	0
Population Proportion Males *	0.25	0.25	0.25
Population Proportion Females *	0.25	0.25	0.25
Population Proportion Children/Elders *	0.50	0.50	0.50
Number of families per post or camp NH	10	10	50
Number of families per post or camp OH	5	5	25
Number of families per post or camp GL	5	5	25
Distance to Camp 1 (days)*	2	2	2
Distance to Camp 2 (days)*	3	3	3
Distance to Camp 3 (days)*	4	4	4
Distance to Camp 4 (days)*	5	5	5
Number of Steps Per Day	6	4	6
Probability of Staying on the camp paths	1.00	1.00	1.00
Probability of Staying at the Camp *	0.99	0.99	0.99
Probability of Staying on the post paths	1.00	1.00	1.00
NH to OH move probability	0.001	0.000006	0.0004
NH to GL move probability	0.001	0.00000	0.0001
OH to NH move probability	0.001	0.00010	0.0007
OH to GL move probability	0.001	0.00007	0.0002
GL to NH move probability	0.001	0.00004	0.0004
GL to OH move probability	0.001	0.00001	0.0002
Time of First Infection (day)	20	20	20
Recovery/Infectious period (days)	5	5	5
Probability of disease transmission **	0.2	0.2	0.2
Probability of Death upon infection **	0.2	0.04	0.04
Death Time (days)	2	2	2
Probability of Contact within a family	0.5	0.5	0.5
Probability of Contact outside a family	0.001	0.001	0.001

* Values can be independently adjusted for each community, but were left the same for these simulations.
** Values can be independently adjusted for each age/gender group, but were left the same for these simulations

The standard values selected for the agent parameters in the NHOHGL model simulations were based upon data from the historical archives. Population size was estimated from the 1917 Canadian census data as presented in Herring and Sattenspiel (2003) and was set at: Norway House = 750 (census count 749); Oxford House = 330 (census count 332); and God's Lake = 300 (census count 299). In each community, the population was divided evenly among the post and the four camps: 150 at each location in Norway House; 66 at each location in Oxford House; and 60 at each location in God's Lake. The age/gender composition of the agents was loosely based upon population data from the Norway House Cree Treaty Annuity Pay lists (Carpenter, 2004) and, although population composition can be independently adjusted for Norway House, Oxford House, and God's Lake, for the sake of consistency, the proportions were set at 25% adult males; 25% adult females; and 50% children/elderly for all three communities. The values for the parameters specifying the number of family groups in each community were based upon ethnographic data suggesting that the size of these groups tended to average about 16 people (Hallowell, 1992). They were set at: 10 groups per post/camp at Norway House (for a total of 50 families with 15 people in each); 5 groups per post/camp at Oxford House (for a total of 25 families with 13.2 people in each), and 5 groups per post/camp at God's Lake (for a total of 25 families with 12 people in each). The sizes of these groups do vary slightly from historical estimates, but an exact match was not possible within the constraints of the NHOHGL model structure.

The standard values for the landscape parameters have been estimated using data on the geography of the region and ethnographic information about the settlement patterns of the Cree people. Within each community, the distances between the central

post and its four associated camps were set at two, three, four, and five days away, respectively. These values are identical to those used in the original Norway House model (Carpenter, 2004) and allow the model to somewhat approximate the varying distances that the winter hunting territories would have been from the nearest HBC post. Although the NHOHGL model allows these distances to be independently adjusted for each community, they were left the same for this project as they represent a settlement pattern that would have been consistent at all three communities.

At the inter-community level, the steps-per-day parameter, which determines the distances between the three communities, was generally set at 6 for the simulations. This makes the distance between Norway House and Oxford House 4 days, the distance between Oxford House and God's Lake 2 days, and the distance between God's Lake and Norway House 6 days. These travel times ensure that the simulated epidemic can periodically reach Oxford House and God's Lake under standard conditions, making the effects of the other input parameters on the spread of the simulated epidemic between communities easier to discern. Additionally, these travel times appear to coincide with estimates of the actual travel times that have been gleaned from the HBC post journals; the limited available data suggest that the journey from Norway House to Oxford House (downstream) would have taken 3 days in the spring and 6 in the winter, while the journey between Oxford House and Norway House (upstream) would have taken 7 days in the fall (Sattenspiel and Herring, 1998). Thus, when steps-per-day is set at 6, the travel time between Norway House and Oxford House is similar to the time it took to make this journey in the summer. The length of the journey between the other communities

remains only an estimate and the length of the journey can not be varied for the departure trip and the return trip.

The standard values for the parameters governing agent mobility in the NHOHGL model are based upon the values used in the original Norway House model and on archival data from the HBC post journals. For travel within a community, the probability of staying at the camp has been set at 0.99 because it is known that travel between the camps and the post would have been relatively infrequent (Herring, 1994). For inter-community travel, the sequence of six parameters, which govern each possible journey, have been set at a constant rate, 0.001, in order to allow these new parameters to be more effectively tested and evaluated. For travel both within and between communities, the probability of the agents staying on the paths, rather than heading off in any direction, has been set at 1.0 because when the agents stray too far from the designated paths, the computer has difficulty keeping track of their positions, causing the model to malfunction and the data to be adversely affected.

The standard values for the epidemic parameters are based upon the available epidemiological data, as well as the values used in the original Norway House model and in Sattenspiel's mathematical models. Recovery time in the NHOHGL model, as in the original Norway House model, has been set to last for 5 days, a value consistent with estimates that the total period of infection with influenza usually lasts about a week (WHO, 2003). Death time, in turn, has been set to 2 days, meaning that an agent is infected for two days before he/she can die. This value is consistent with the estimate for severe cases in the 1918-1919 pandemic of between 1 and 3 days (Crosby, 2003). The probability of infection upon contact with an infectious agent and the probability of death

upon infection are more difficult to measure, especially for historical epidemics, and thus their standard values in the NHOHGL are purely speculative. For this project, the probability of infection was alternately set at 0.2 to mimic its value in the original Norway House model and at 0.5 to approximate its value in Sattenspiel's mathematical models. The probability of death, in turn, was set at either 0.2, the value used in the original Norway House Mortality model, or 0.04, a value which appeared to elicit a 15-20% mortality rate among the infected agents, a rate equal to that at Norway House during the 1918-1919 flu epidemic (Herring, 1994). Although both the probability of infection and the probability of death can be independently adjusted for each of the age/gender groups in the model, the values were kept the same for all three in an effort to better replicate the patterns of the 1918-1919 flu epidemic, which tended to target all ages due to an unusually high severity among young adults (Crawford, 2000; Crosby, 2003; Karlen, 1995; Taubenberger et al., 2005a; Taubenberger et al., 2005b).

The standard values for the parameters controlling agent interaction are also speculative, but are based upon ethnographic data concerning Cree social organization. While it is clear from these data that contact between members of the same family group would have been much more common than contact between members of different family groups, especially during the winter months, actual historical interaction rates for the Cree and Métis of this region are impossible to determine accurately. The probability of contact within a family group has thus been set at 0.5 and the probability of contact outside of a family group has been set at 0.01 in order to preserve the historical pattern in the absence of any evidence for historical contact rates.

This set of standard parameter values provided a means for testing the new model, but also served as the basis for the construction of two additional model scenarios: a summer and a winter scenario. Separate summer and winter scenarios were developed for the NHOHGL model because the traditional seasonal residence pattern that was practiced by the Cree and Métis at the time of the 1918-1919 flu epidemic led to marked seasonal differences in population composition and mobility that could have had a strong influence on the spread of infectious disease in this region. A summary of these differences, as incorporated into the NHOHGL winter and summer scenarios can be found in Table 5.4.

In the winter scenario, the agent population at each community is divided evenly between the post and the four outlying camps. This is meant to replicate the dispersal of the family hunting groups to their traditional winter hunting territories. The agents move randomly within their home areas. Adult males from the camps are allowed to travel back and forth to their home post, while adult males from the posts are allowed to travel to the other posts. Travel between the three communities is known to have been slower during the winter months, when the primary mode of transportation would have been by snowshoe or dogsled (Herring and Sattenspiel, 2003). For this reason, the steps-per-day parameter has been set at 4 for the winter simulations. This makes the journey between Norway House and Oxford House last 6 days, a value equal to an estimate taken from the HBC post journals (Sattenspiel and Herring, 1998). Travel would also have been less frequent in the winter, and the values used for the probability of travel in the winter scenario are: NH to OH = 0.000006; NH to GL = 0.00000; OH to NH = 0.00010; OH to GL = 0.00007; GL to NH = 0.00004; and GL to OH = 0.00001. These values have been

Table 5.4: A comparison of the assumptions and parameter values for the winter and summer model scenarios

- **Winter Scenario**:
  - In each community, the agents are evenly divided between the posts and the four camps
    - Each post population = 1/5 total community population
      - NH = 150 at each location (750 total)
      - OH = 66 at each location (330 total)
      - GL = 60 at each location (300 total)
    - Each camp population = 1/5 total community population
  - The agents move randomly within their home areas
  - Travel between the posts is slower
    - Steps–per-day set at 4
  - Travel between the posts is less frequent
    - Probability of travel from NH to OH = 0.000006
    - Probability of travel from NH to GL = 0.00000
    - Probability of travel from OH to NH = 0.00010
    - Probability of travel from OH to GL = 0.00007
    - Probability of travel from GL to NH = 0.00004
    - Probability of travel from GL to OH = 0.00001

- **Summer Scenario:**
  - All the agents congregate at their community post
    - Each post population = total community population
      - NH = 750
      - OH = 330
      - GL = 300
    - Each camp population = 0
  - The agents move randomly at the posts
  - Travel between the posts is faster
    - Steps-per-day set at 6
  - Travel between the posts is more frequent
    - Probability of travel from NH to OH = 0.0004
    - Probability of travel from NH to GL = 0.0001
    - Probability of travel from OH to NH = 0.0007
    - Probability of travel from OH to GL = 0.0002
    - Probability of travel from GL to NH = 0.0004
    - Probability of travel from GL to OH = 0.0002

calculated using data on the frequency of departures from each post and the distribution of travel by post as recorded in the HBC post journals (Herring and Sattenspiel, 2003). Movement between the posts was more limited in the wintertime than it was during the summer months, and most travel consisted of one or more trappers leaving their winter camps to hunt or to collect furs from their trap lines or, less often, to return to the fort to pick up supplies and trade their furs (Herring, 1994).

In the summer scenario, the agent population at each community is placed entirely at the post while the outlying camps are left empty. This is meant to replicate the aggregation of the family hunting groups near the HBC posts during the summer months. The agents move randomly at the posts, and although contact rates are left at their standard values in both the summer and the winter scenarios, because the family hunting band was still the basic unit of society regardless of season (Hallowell, 1992), the concentration of the agents at the posts in the summer scenario promotes increased interaction, just as it would have done historically. In addition to interacting, adult males are allowed to travel to the other posts; however, no travel between the camps and the post occurs in the summer simulation, as all the agents are assumed to be gathered at the posts. Travel between the three communities would have been faster during the summer months, when canoe travel was possible (Herring and Sattenspiel, 2003). For this reason, the steps-per-day parameter is set at 6 in the summer scenario, making the journey between Norway House and God's Lake last 4 days as in the standard scenario. The probabilities of travel between each of the three communities in the summer scenario have been calculated in the same fashion as the values for the winter scenario using data from the HBC post journals as provided by Sattenspiel (personal communication

8/17/05). These probabilities have been calculated to be: NH to OH = 0.0004; NH to GL = 0.0001; OH to NH = 0.0007; OH to GL = 0.0002; GL to NH = 0.0004; and GL to OH = 0.0002.

- **The Model Output**

While the inputs for the model consist of the values chosen for the initial parameters, the output consists of the statistics that are generated by the simulated epidemic. The statistics that are generated by the NHOHGL simulations are similar to those that would be of interest for a real epidemic and include the number of susceptibles (S), exposed (E), infectious (I), recovered (R), and dead (D) agents at each day of the simulation. These data were collected for the total population, for each community, and for each post, and camp. The model has been programmed to display these data in numerical form as well as to produce running graphs illustrating the number of susceptible, infected, recovered, and dead per day of the simulation in the total population and at each community.

Because the NHOHGL model is stochastic, each run is unique and a single run may not be informative on its own. For this reason, the results of several runs are often averaged in order to provide more representative and informative statistics (Gilbert and Troitzsch, 1999). In this model, as in the original Norway House model, the number of runs per parameter change was made into a parameter so that it could be easily altered by the researcher. For the testing of the NHOHGL model, as for the testing of the original Norway House model, 1000 runs at each set of parameter values were averaged to obtain a better picture of how the model behaved. The results of the model simulations, which

will be presented in the following chapter, are thus the averages of 1000 runs, rather than the results from only one 200-day simulation.

Once a set of 1000 runs had been completed, the basic epidemic data were used to calculate epidemic statistics for each community and for the population as a whole. These include: the duration of the epidemic; the time of peak infection (the day of the epidemic on which the most agents are infected); the peak number infected; the length of the epidemic (the period between the first and last associated death); the time of the peak number dead; the peak number dead per day; the total number infected (recovered + dead); the percent of the population infected; and the percent of the population that died. For each set of runs, a count was also made of the number of times that the epidemic reached Oxford House, God's Lake, or both of the additional posts from Norway House. All calculations were performed by a Microsoft Excel macro program that was created by Nathaniel Green and written in Visual Basic (a copy of this program can be found in Appendix B). The NHOHGL macro takes the raw data produced from the simulations, averages the values for all 1,000 runs, and deposits these averages, as well as a count of the number of times the epidemic reached a location other than NH, into an Excel spreadsheet. The original data were always saved in case details on any individual run were needed or further calculations had to be made.

The data from the simulated epidemic were also used to create several graphs to provide a visual representation of the behavior of the model. For each run, Microsoft Excel was used to create several graphs using the averaged data as compiled by the NHOHGL Excel macro. These include: a graph of the number of susceptibles, infected, recovered, and dead agents per day of the simulation for the total population; a similar

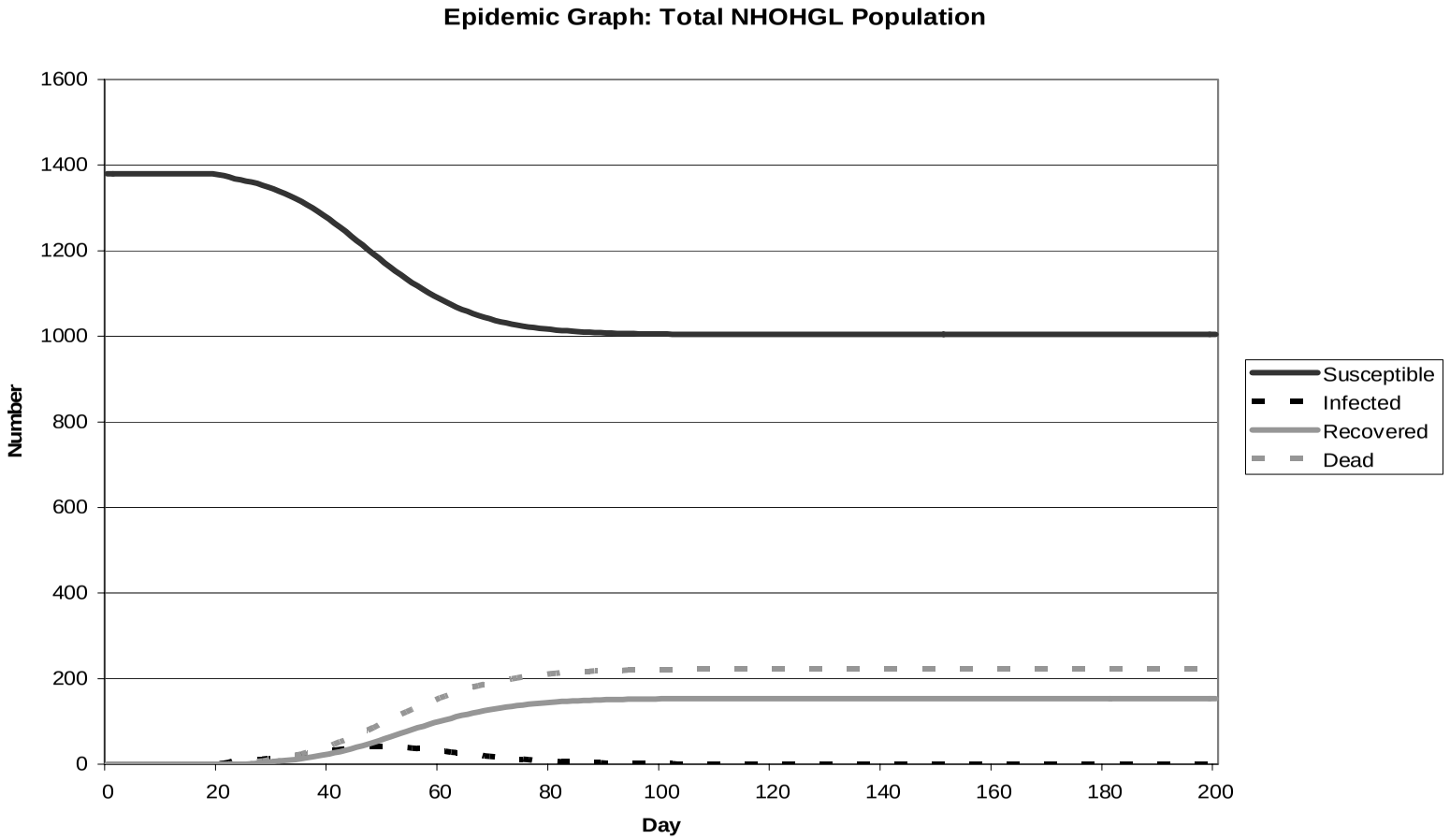
graph for each of the three communities; a graph comparing the number of infected agents per day at each community; and a graph comparing the number of dead agents per day at each community. Samples of these graphs may be found in Figure 5.5.

Advantages of the New NHOHGL Model

Although many challenges were faced in the design and construction of the NHOHGL model, the benefits of this model have made it well worth the effort. The expanded framework of this model offers an opportunity to study the 1918-1919 flu epidemic on a regional rather than a community level and allows it to provide insight into the impact that the traditional Cree seasonal population movements may have had upon the spread and severity of the 1918-1919 flu epidemic between communities as well as within them. Further, the three-community format of the NHOHGL model enables it to address questions that could not be addressed with the previous Norway House model. These include: why the 1918-1919 influenza epidemic failed to reach Oxford House and God's Lake and how the results of the agent-based simulations compare to those of Sattenspiel's mathematical models of the 1918-1919 flu epidemic. Although these questions were eventually investigated using the NHOHGL model (see chapter 7), the new model first had to be thoroughly tested in order to ensure that it was working properly. These tests will be the subject of the following chapter.

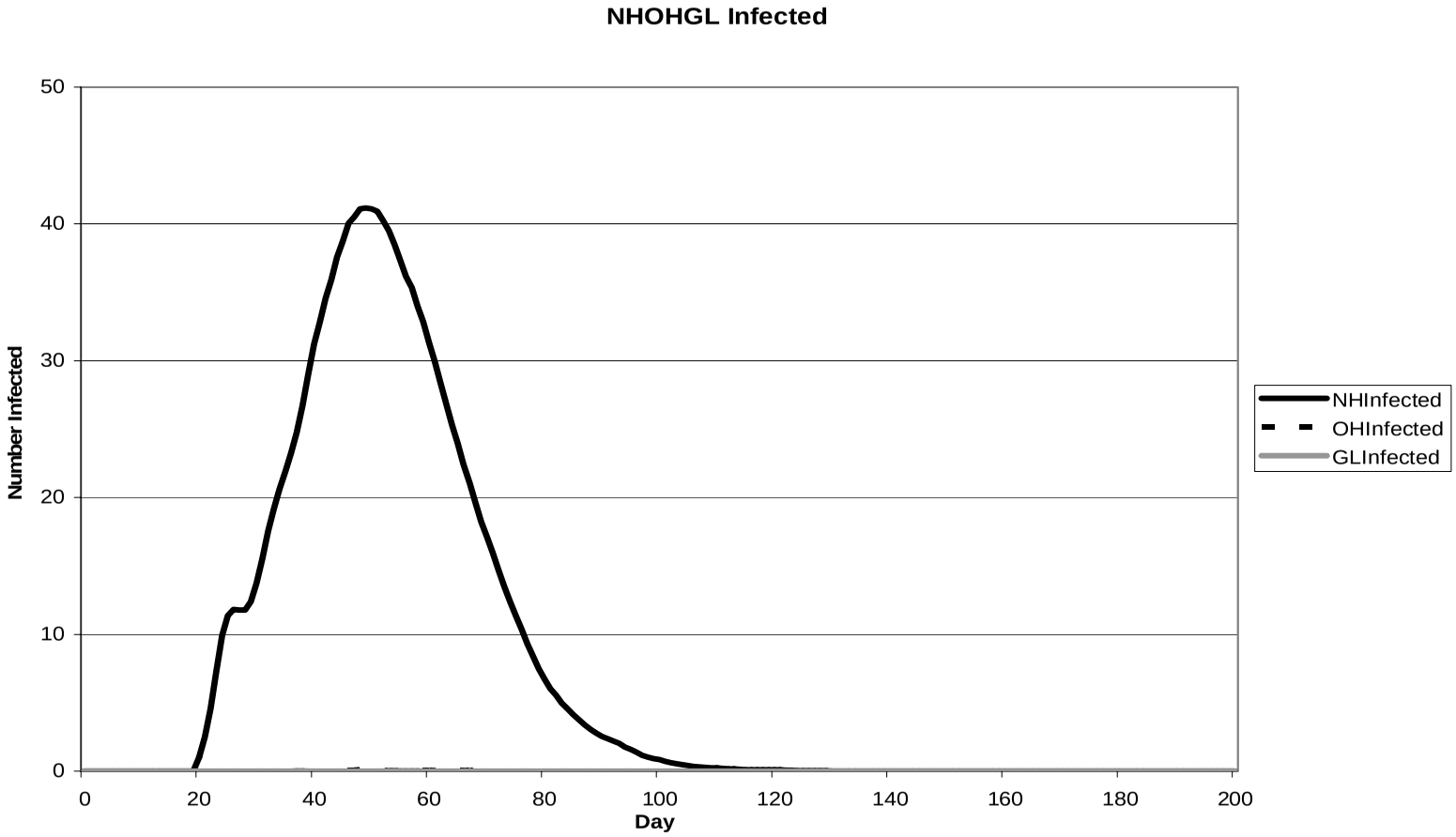
Figure 5.5: Sample graphical output for the NHOHGL model simulations (run#071305- standard parameter values)

a. The number of susceptibles, infected, recovered, and dead agents per day of the simulation for the total population (similar graphs were also produced for each community)



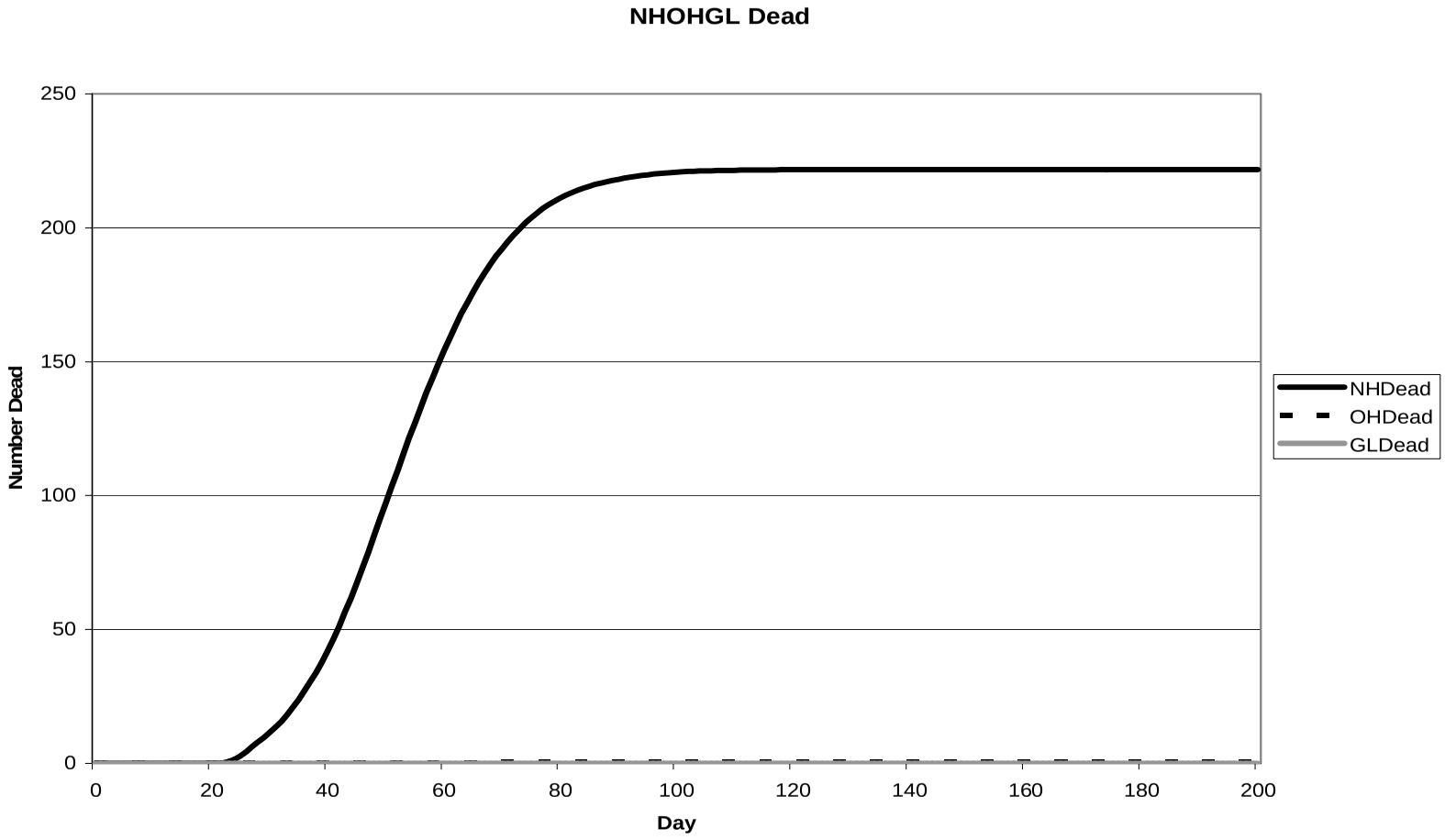
b. A comparison of the number of infected agents per day at each community

In this graph, the largest epidemic is seen at Norway House and there is no discernable epidemic peak at Oxford House or God's Lake. This is because, in the majority of the 1000 runs, the flu failed to reach either of these communities.



c. A comparison of the number of dead agents per day at each community

In this graph, the largest death toll is seen at Norway House and there are no discernable deaths at Oxford House or God's Lake. This is because, in the majority of the 1000 runs, the flu failed to reach either of these communities.



Chapter 6: Results of the NHOHGL Test Simulations

Once the NHOHGL model had been constructed, it had to be thoroughly tested before it could be used to investigate the research questions that were outlined for this project. The first step in the testing of any new computer simulation is known as verification. Verification is the process of confirming that the computer simulation is performing appropriately and as designed. It allows the researcher to identify errors in the model design or program and correct them before the model is used to investigate any research questions. The process generally includes three steps: debugging the computer program; running the model several times at the same set of parameters; and testing the model at extreme parameter values where the outcome can be predicted (Gilbert and Troitzsch, 1999). For the NHOHGL model, debugging was performed as part of the model construction process and will not be profiled here. The results of the other two verification tests, the replication study and the sensitivity analysis, will be presented in this chapter.

The Replication Study

The replication study compared the averaged data for 10 to 20 sets of 1000 runs of the NHOHGL model using the same set of parameter values. Tests were performed for each of the three input scenarios: the standard parameters, the winter parameters, and the summer parameters (see Table 5.3. for a review of the values used in each scenario). The goals of this study were twofold: to verify whether the new model was behaving in a consistent manner and to establish a control range for the simulated epidemics.

Establishing a control range is an important part of the analysis of any new stochastic model. Unlike deterministic models, where each run of the simulation at the same set of parameter values generates the same set of data, each time a stochastic model is run, slightly different data are produced. This is why 1000 runs of the NHOHGL model were performed for each parameter change in this project. Yet, even the averaged data for 1000 runs of a stochastic model may differ from the averaged data for another set of 1000 runs at the same set of parameter values. This randomness is the primary advantage of stochastic models, but it can make data analysis problematic.

To aid in the analysis of stochastic data, a control range at a set of standard parameter values can be used to establish the normal range of variation for the model. The data resulting from a change in parameter values can then be compared to this range in order to establish whether any observed differences are the result of the parameter change or due to normal stochastic fluctuations. The control ranges established in this study for the standard, winter, and summer model scenarios were used in this fashion throughout the project in order to analyze the effects of each parameter change. While the results for these tests will be presented later in this chapter, the results of the replication tests and the control ranges that were established will be presented in the following paragraphs.

- **Replication Test for the Standard Parameter Scenario**

As noted in the previous chapter, the standard parameters were designed to serve as the default or control values for the sensitivity tests, and the epidemic that they elicited was to serve as a baseline for comparison in these tests. For this reason, it was important

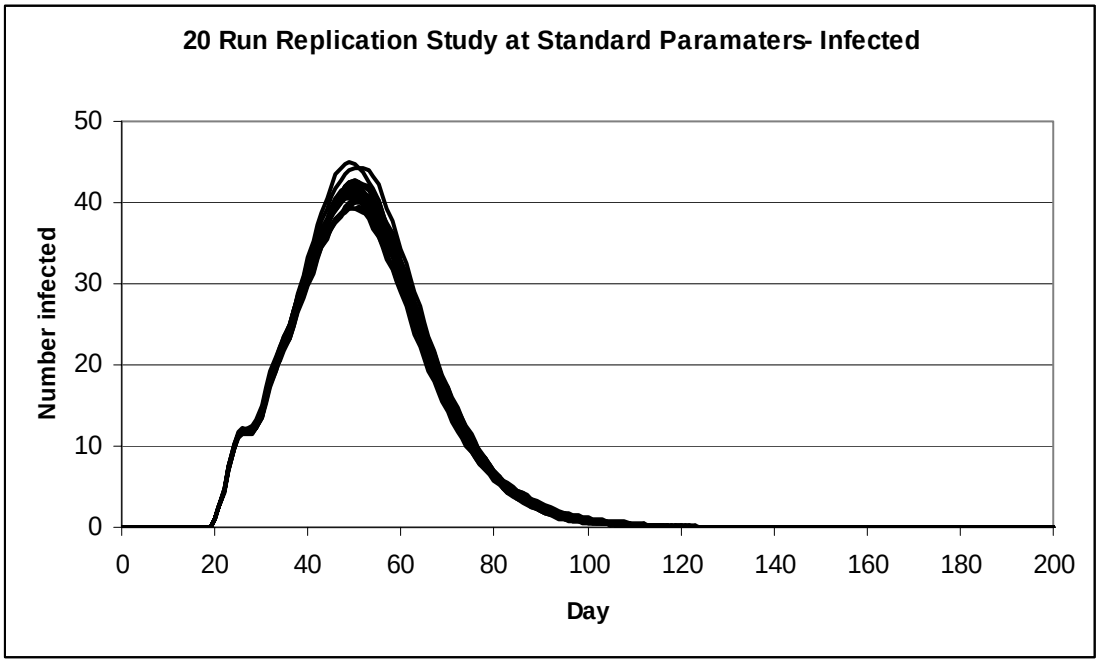
that the normal range of variation for the standard epidemic scenario be established. A replication test was performed in order to establish this control range. The test compared 20 sets of 1000 runs of the NHOHGL model using the standard parameter values and assuming a probability of infection of 0.2. A graphical comparison of the averaged data from each set of 1000 runs may be found in Figure 6.1.

As would be expected for a stochastic model, the graphs indicate that the averaged data for each set of 1000 runs were somewhat variable, even though the same set of input parameters was used throughout the replication test. The impact of stochasticity is evident in both the number of infected agents per day of the simulation (Figure 6.1a) and in the number of dead agents per day (Figure 6.1b). While the overall shape of these morbidity and mortality curves are essentially the same for each of the 20 sets of runs, the size of the epidemic peak and the overall death toll are somewhat more variable.

Means and variances were calculated to measure the extent of this variability and to establish the control range for the standard parameter scenario (see Table 6.1). The control range reveals variability both in the duration of infection and in the length of the epidemic. In both cases, the variability fell within a range of 103-133 days and had a standard deviation of about 8 days. In contrast to this variability, the peak time of infection remained relatively stable, occurring only within a four-day period. The peak number of infected agents was also relatively stable and ranged from 39.40 to 44.94 with a standard deviation of 1.42. As for the epidemic totals, the averaged final death toll for a set of 1000 runs ranged from 210.71 to 232.33 with a standard deviation of 5.37 and the total number infected ranged from 56.46 to 393.45 with a standard deviation of 9.11.

Figure 6.1: A comparison of the averaged data for 20 sets of 1000 runs of the NHOHGL model at the standard parameters showing the number of infected agents per day of the simulation (a) and the number of dead agents per day of the simulation (b)

a.



b.

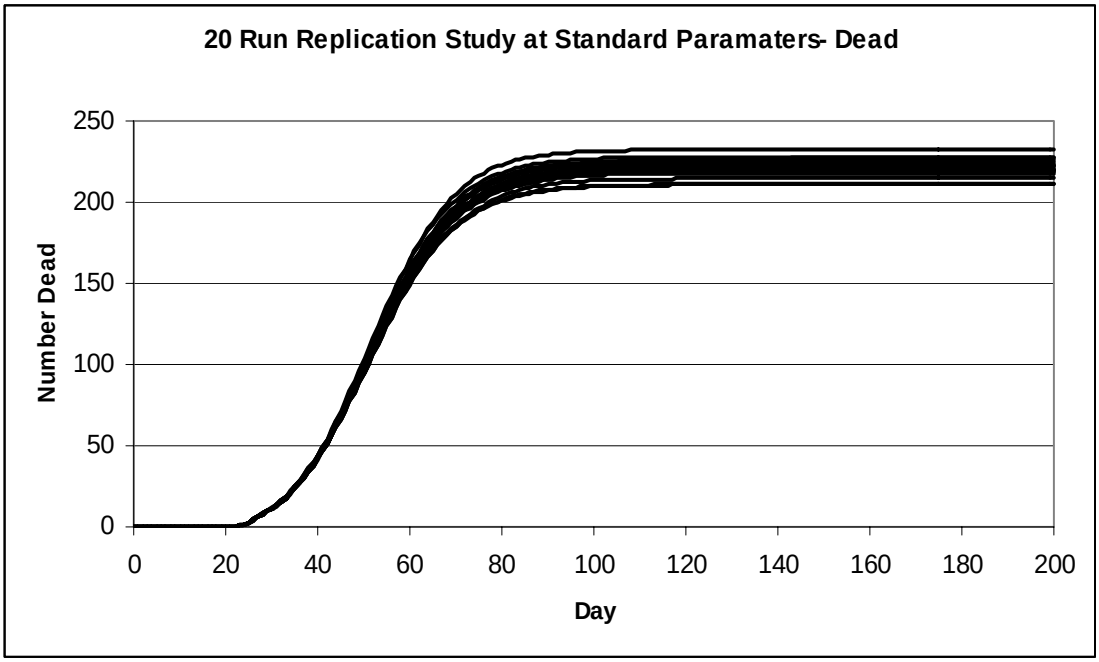


Table 6.1: Control range for the standard parameter scenario
Table shows the epidemic data for the total model population and for the populations at each community.

Standard Control Range	Maximum	Minimum	Mean	Std. Dev.
Duration of Infection (days)	133	105	123.45	8.12
Length of the Epidemic (days)	132	103	121.50	8.33
Peak time of Infection (day)	32	28	29.70	0.92
Peak Number Infected	44.94	39.40	41.62	1.42
Total Number Recovered	161.13	145.75	153.41	3.76
Total Number Dead	232.33	210.71	221.19	5.37
Total Number Infected	393.45	356.46	374.60	9.11
% Infected	28.51%	25.83%	27.14%	0.01
% Dead	16.84%	15.27%	16.03%	0.00
Dead/Infected	0.59	0.59	0.59	0.00
Epidemics at OH	27	8	16.65	5.31
Epidemics at GL	13	2	4.50	2.61
Epidemics at OH and GL	1	0	0.15	0.37
NH Number Dead	231.55	210.37	220.58	5.27
NH Number Infected	392.17	355.90	373.57	8.95
NH % Dead	30.87%	28.05%	29.41%	0.01
NH % Infected	52.29%	47.45%	49.81%	0.01
OH Number Dead	0.85	0.21	0.52	0.20
OH Number Infected	1.41	0.37	0.88	0.33
OH % Dead	0.26%	0.06%	0.16%	0.00
OH % Infected	0.43%	0.11%	0.27%	0.00
GL Number Dead	0.18	0.01	0.09	0.05
GL Number Infected	0.31	0.02	0.15	0.09
GL % Dead	0.06%	0.00%	0.03%	0.00
GL % Infected	0.10%	0.01%	0.05%	0.00

Although these differences appear to be larger than those observed for the other epidemic statistics, they are actually quite small in relation to the size of the model population: the range in death toll represents only a 1.6% difference and the range in the number infected represents only a 2.7% difference.

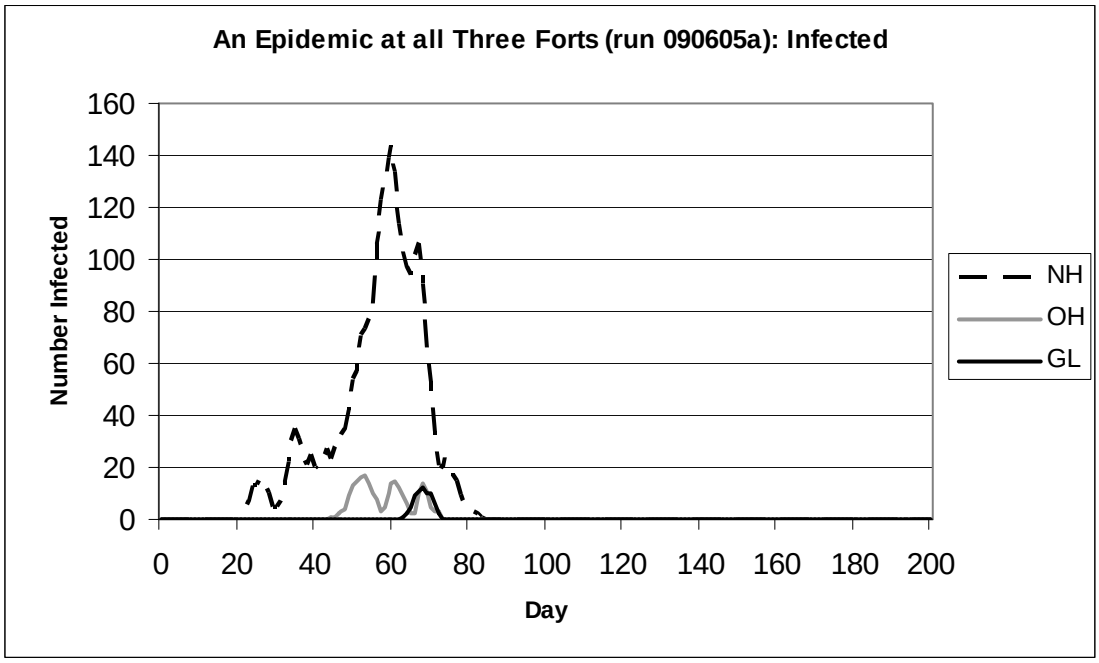
The statistics also reveal variability in the number of times that the epidemic managed to spread from Norway House to either of the two additional communities. Within any set of 1000 runs, the epidemic managed to reach Oxford House between 8 and 27 times (standard deviation 5.31) and to reach God's Lake between 2 and 13 times (standard deviation 2.61). In three of the sets of runs, the epidemic managed to reach all three posts in a single run, however; in each case this occurred during just one of the 1000 simulations. Graphs showing the morbidity and mortality data for one of these runs are shown in Figure 6.2. These graphs show that, even when the flu did manage to reach all three posts, most of the epidemic activity was still at Norway House. This observation is supported by the control range statistics, which indicate that most of the variability in the morbidity and mortality figures relate to the epidemics at Norway House and show only minor epidemic activity at Oxford House and God's Lake. This would be expected because the epidemic always occurred at Norway House and only rarely reached either or both of the other communities.

- **Replication Tests for the Winter and Summer Scenarios**

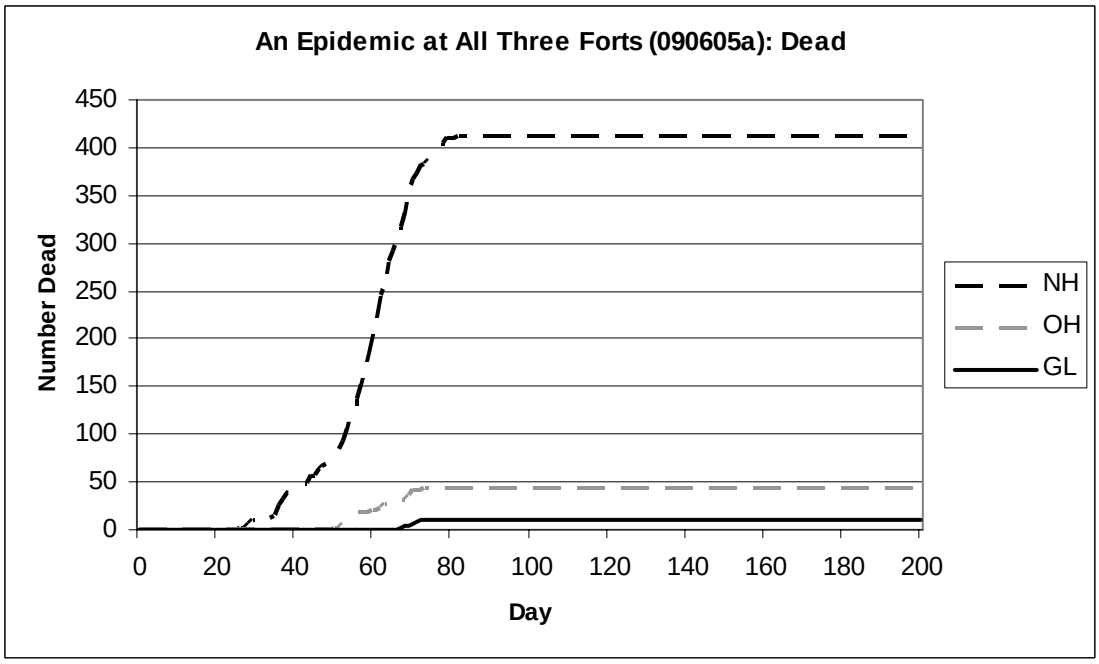
In addition to the replication test for the standard parameters, replication tests were also performed for both the summer and winter parameters in order to establish a control range for each scenario. These replication tests each compared the averaged data

Figure 6.2: Graphical output for an epidemic that reached all 3 posts showing the number of infected agents per day (a) and the number of dead agents per day (b). Note the larger epidemic at Norway House as compared to at Oxford House and God's Lake. Data is from a single run in which the epidemic reached all three posts during the standard parameter replication test number 090605a

a.



b.



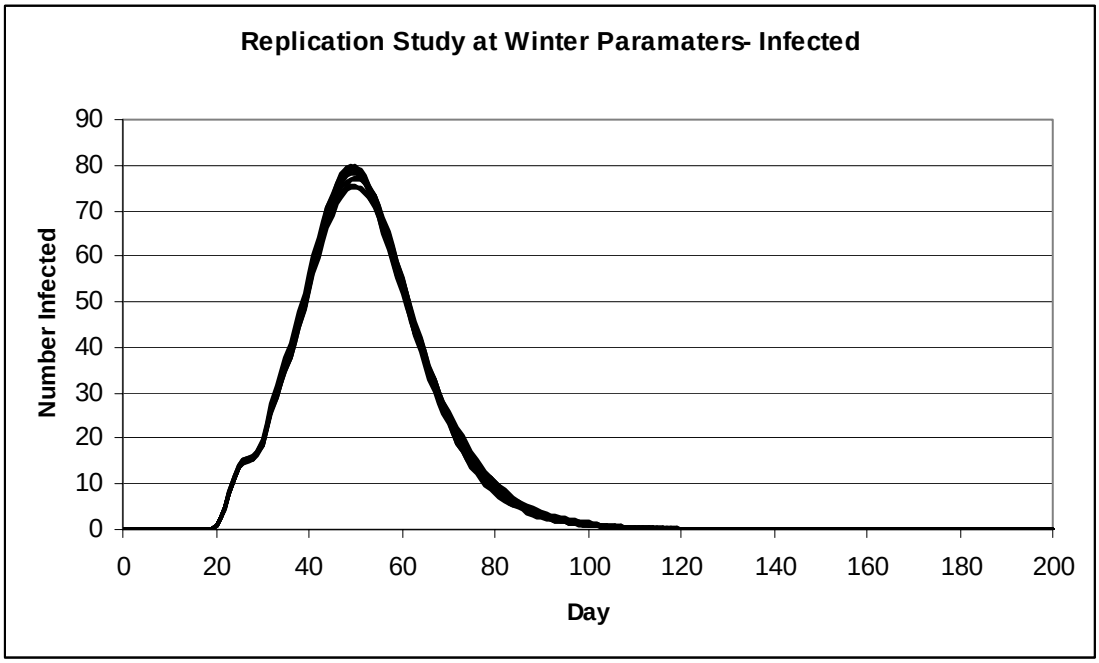
from 10 sets of 1000 runs of the NHOHGL model using the same set of parameter values. For both tests, the probability of infection was set at 0.2 and the probability of death was set at 0.04. A graphical comparison of the averaged data for the winter runs may be found in Figure 6.3 while the graphs for the summer runs may be found in Figure 6.4. The control ranges established by the winter and summer replication tests are presented in Table 6.2 (winter) and Table 6.3 (summer).

In the winter scenario, the averaged duration of infection for the 10 run sets varied between 109 and 148 days with a standard deviation of 11.0. The length of the epidemic was also variable, ranging from 104 to 144 days with a standard deviation of 10.94. The peak time of infection was less variable, and only fell upon the 29th and 30th day of the epidemic (standard deviation 0.48). The peak number infected was also quite stable, ranging between 75.18 and 79.85 for the 10 sets of runs (standard deviation 1.78). The epidemic totals were somewhat variable, but less so than for the standard parameter replication test. The death toll ranged from 76.57 to 79.49 (standard deviation 0.91) while the total number of infected agents ranged from 510.71 to 532.28 (standard deviation of 6.03). For each of the 10 sets of 1000 runs, the simulated epidemic managed to reach Oxford House between 8 and 10 times (standard deviation 3.08) but did not reach God's Lake in any of the 10,000 runs.

In the summer scenario, the duration of infection lasted between 75 and 108 days with a standard deviation of 10.97 while the length of the epidemic ranged between 72 and 104 days with a standard deviation of 10.71. The timing and extent of the epidemic peak in the summer scenario are a bit more stable than for the winter scenario, with the average time of peak infection for all 10 sets of runs falling only upon day 17 of the

Figure 6.3: A Comparison of the averaged data for each of the 10 sets of 1000 runs of the NHOHGL model at the winter parameter values showing the number of infected agents per day of the simulation (a) and the number of dead agents per day of the simulation (b)

a.



b.

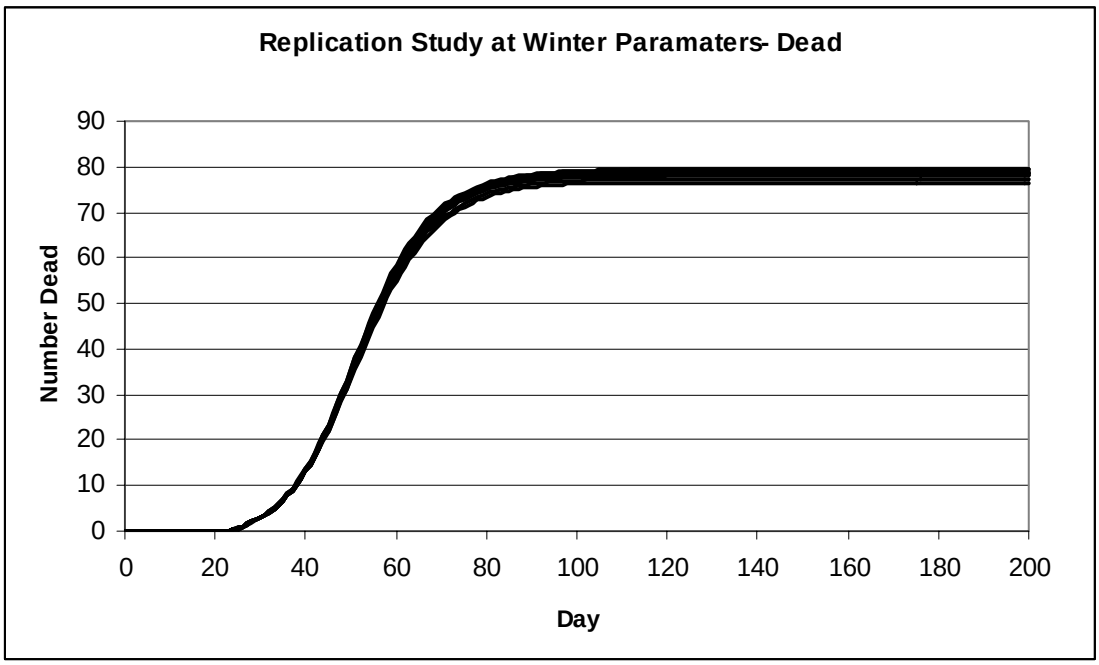
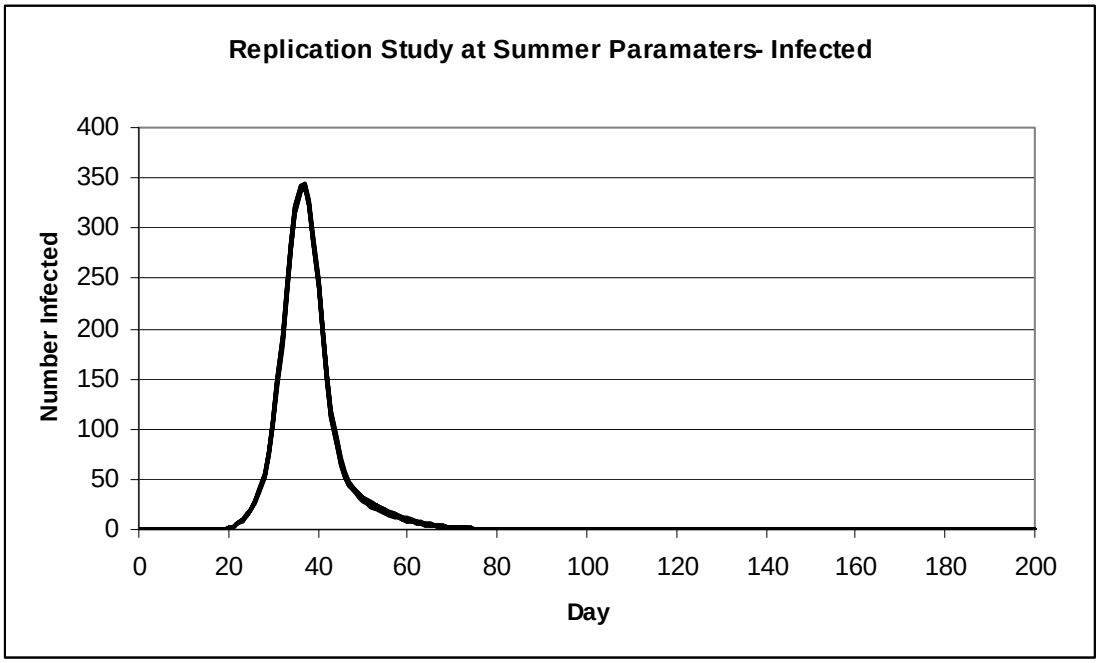


Figure 6.4: A Comparison of the averaged data for each of the 10 sets of 1000 runs of the NHOHGL model at the summer parameter values showing the number of infected agents per day of the simulation (a) and the number of dead agents per day of the simulation (b)

a.



b.

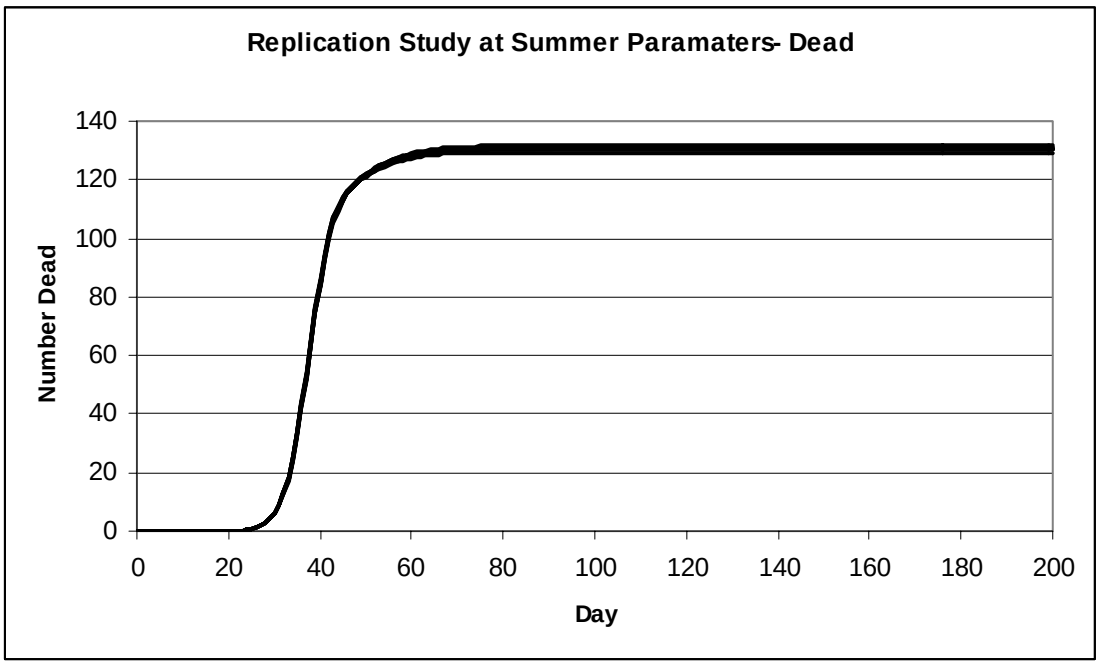


Table 6.2: Control range for the winter parameter scenario

Winter Control Range	Maximum	Minimum	Mean	Std. Dev.
Duration of Infection	148	109	126.50	11.00
Length of the Epidemic	144	104	119.70	10.94
Peak time of Infection	30	29	29.30	0.48
Peak Number Infected	79.85	75.18	78.06	1.78
Total Number Recovered	452.79	434.14	442.08	5.20
Total Number Dead	79.49	76.57	78.33	0.91
Total Number Infected	532.28	510.71	520.40	6.03
% Infected	38.57%	37.01%	37.71%	0.00
% Dead	5.76%	5.55%	5.68%	0.00
Epidemics at OH	18	8	13.20	3.08
Epidemics at GL	0	0	0.00	0.00
Epidemics at OH and GL	0	0	0.00	0.00

Epidemic data are combined totals for the three communities.

Table 6.3: Control range for the summer parameter scenario

Summer Control Range	Maximum	Minimum	Mean	Std. Dev.
Duration of Infection	108	75	86.30	10.97
Length of the Epidemic	104	72	82.00	10.71
Peak time of Infection	17	17	17.00	0.00
Peak Number Infected	344.13	339.55	342.62	1.21
Total Number Recovered	740.50	728.52	734.68	3.20
Total Number Dead	131.67	129.35	130.55	0.57
Total Number Infected	872.17	857.86	865.23	3.66
% Infected	63.20%	62.16%	62.70%	0.00
% Dead	9.54%	9.37%	9.46%	0.00
Epidemics at OH	390	347	374.40	11.94
Epidemics at GL	0	0	0.00	0.00
Epidemics at OH and GL	0	0	0.00	0.00

Epidemic data are combined totals for the three communities.

epidemic and the peak number infected ranging only from 339.55 to 344.13 (standard deviation of 1.21). The epidemic statistics were, however, similar in variability with the summer death toll ranging between 129.35 and 131.67 with a standard deviation of 0.57 and the total number infected ranging from 857.86 to 872.17 with a standard deviation of 3.66. In the summer replication test, out of each set of 1000 runs, the flu managed to make it to Oxford House an average of 347 to 390 times. However, as in the winter replication test, the flu never reached God's Lake.

While the winter and summer control ranges appear to show similar patterns of variability, the epidemic data for the two scenarios exhibit significant and consistent differences. The two scenarios differ in the duration of infection and the length of the epidemic, in the timing and size of the epidemic peak, in the total numbers of agents that were infected or died, and in the frequency with which the simulated epidemic managed to spread to the additional communities of Oxford House and God's Lake. These differences are expected due to the differences in population density and travel for the two seasonal scenarios and they will be examined further in the following chapter.

The Sensitivity Analysis

In the sensitivity analysis, each of the primary input parameters in the NHOHGL model was tested over a wide range of values while the other parameters were held constant at their standard control values. The goal of this analysis was to determine whether each parameter was behaving appropriately and as designed. During these tests, special attention was paid to the newly created inter-community mobility parameters, which were previously untested. Less attention was paid to other parameters, e.g., camp

distance, camp stay, and probability of staying on the paths, which retained the same form as in the original Norway House model, where they had been thoroughly evaluated (see Carpenter, 2004). The following are the results of the sensitivity tests for each of the primary NHOHGL model parameters.

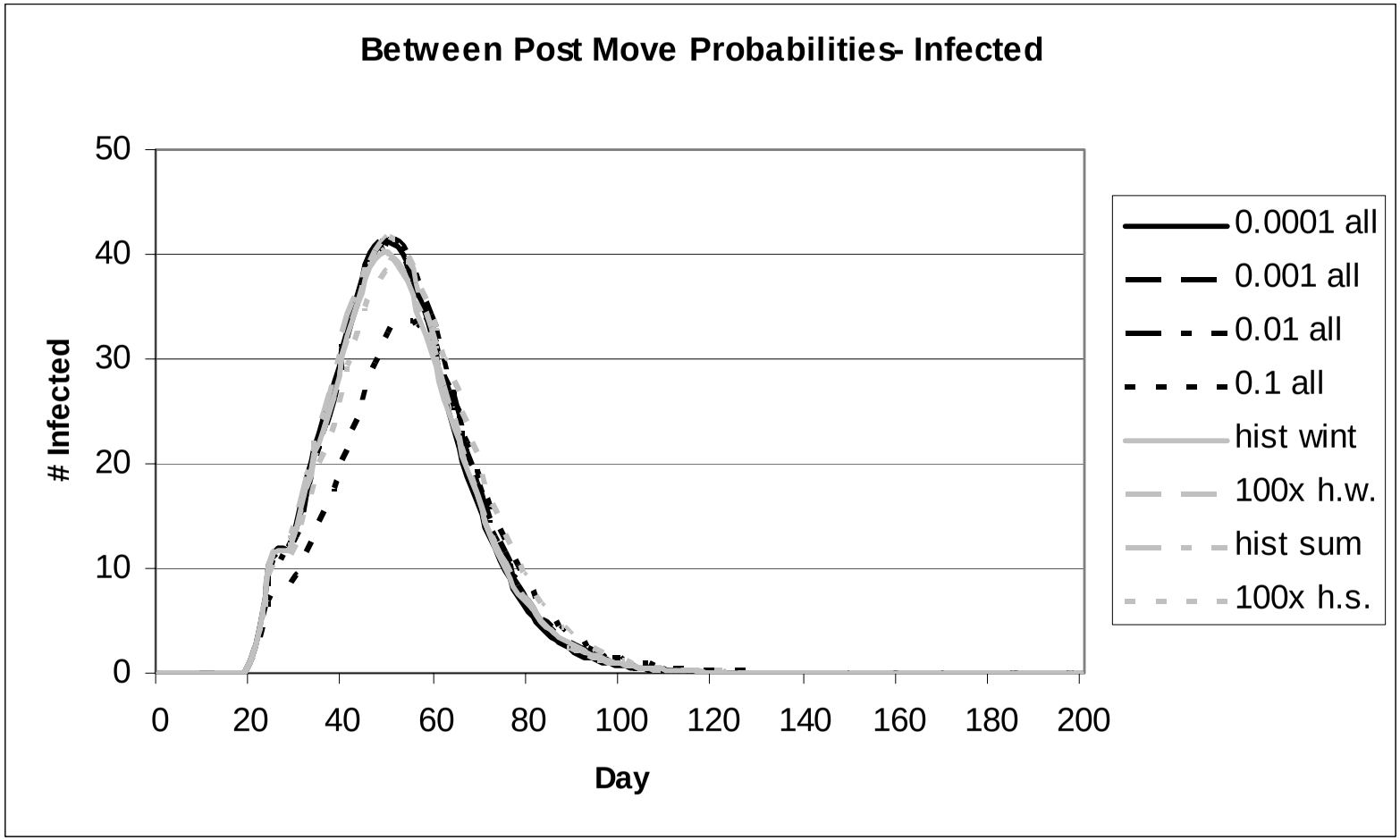
- **Between Post Move Probabilities**

The between post move probabilities are a series of six parameters which specify the probability that an agent from one post would travel to another post on any given day of the simulation. Each of the parameters specifies a different potential journey: Norway House to Oxford House; Norway House to God's Lake; Oxford House to Norway House; Oxford House to God's Lake; God's Lake to Norway House; or God's Lake to Oxford House. The standard control value for all of these parameters is 0.001; however, historical estimates from the HBC post journals were used in both the winter and summer parameter scenarios. Because the between post move probabilities control the flow of potentially infected individuals from one community to another, they are important factors to consider in the inter-community spread of the simulated epidemic.

The between post move probability parameters were tested at a range of values in order to ensure that they were functioning properly. The results of this sensitivity test are shown in Table 6.4 and a graphical comparison of the infection and mortality data for the test runs is provided in Figure 6.5. Analysis of the chart reveals some interesting features that seem to indicate that these new parameters are indeed performing in an appropriate manner. As the probability of moving between the forts was increased, the epidemic more often reached Oxford House and God's Lake, falling outside of the standard control

Figure 6.5: Results of the between post move probabilities sensitivity test

a. The number of infected agents per day of the simulation at different values for the between post move probabilities



b. The number of dead agents per day of the simulation at different values for the between post move probabilities

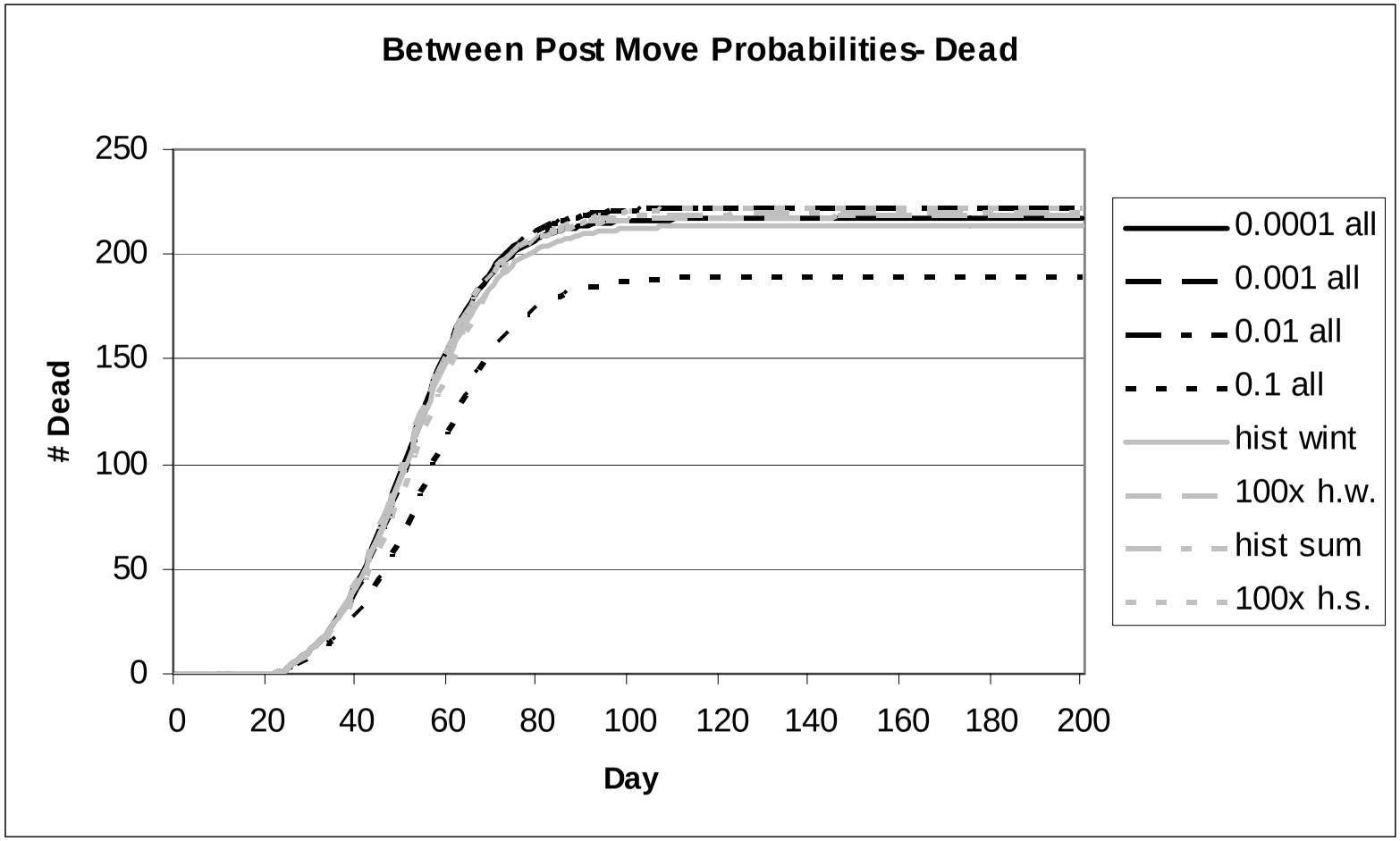


Table 6.4: Epidemic characteristics for the between post move probabilities

Between Post Move Probs.	All 0.0001	All 0.001*	All 0.01	All 0.1	Hist. Wint.	100x Winter	Hist. Sum.	100x Sum.
Duration of Infection	130	123	112	131	125	133	113	143 ↑
Length of the Epidemic	127	122	108	130	121	113	112	140 ↑
Peak Time of Infection	29	30	31	33 ↑	30	30	28	32
Peak Number Infected	41.24	41.62	41.42	33.81 ↓	40.31	41.60	40.37	39.64
Total Recovered	150.41	153.41	153.69	132.53	148.44	151.42	151.53	154.54
Total Dead	217.05	221.19	222.57	189.78↓	213.69	217.85	219.13	222.51
Total Infected	367.45	374.60	376.26	322.31↓	362.12	369.26	370.66	377.04
% Infected	26.6%	27.1%	27.3%	23.4% ↓	26.2%	26.8%	26.9%	27.3%
% Dead	15.7%	16.0%	16.1%	13.8% ↓	15.5%	15.8%	15.9%	16.1%
Epidemics at OH	10	17	138 ↑	368 ↑	17	91 ↑	16	328 ↑
Epidemics at GL	0 ↓	5	39 ↑	199 ↑	0 ↓	0 ↓	0 ↓	41 ↑
Epidemics at OH&GL	0	0	14 ↑	154 ↑	0	0	0	26 ↑

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

range at both the 0.01 and 0.1 input values. A similar increase also occurred when the historical travel rates were multiplied by 100. This would be expected because increased travel rates would increase the chance that an infectious traveler would carry the flu from Norway House to one of the other communities. Increasing the between post move probabilities to 0.1 also led to a slightly delayed and smaller epidemic peak as well as fewer infections and deaths. These values all fell outside of the control range for the standard parameter scenario and this may have been due to the fact that many of the infected agents at Norway House were off traveling, making it more difficult for the flu to spread within the community while also increasing its chance of spreading outside of

the community. Interestingly, increasing the between post move probabilities did not appear to significantly impact the duration of infection or mortality, except when the summer travel values were increased. It may have been that, because the populations in both the standard and winter scenarios were divided among the post and the camps, the population was too dispersed to be able to support a longer epidemic.

- **Steps-Per-Day**

The steps-per-day parameter was created especially for the NHOHGL model in order to control the speed with which the agents traveled between the three communities. Speed of travel can strongly impact the likelihood that a disease will spread from one community to another due to the dynamics of disease transmission. In order for a disease to be transmitted between communities, an individual first has to contract the disease in a community where it is present and then travel to a new, uninfected, community and transmit the disease to a local susceptible before either dying or recovering. When the length of the recovery/infectious period and the distance between the two communities are held constant, an epidemic would be more likely to reach the new community as the speed of travel is increased. This is because more of the infected individuals would be able to reach the new community before the end of the recovery period.

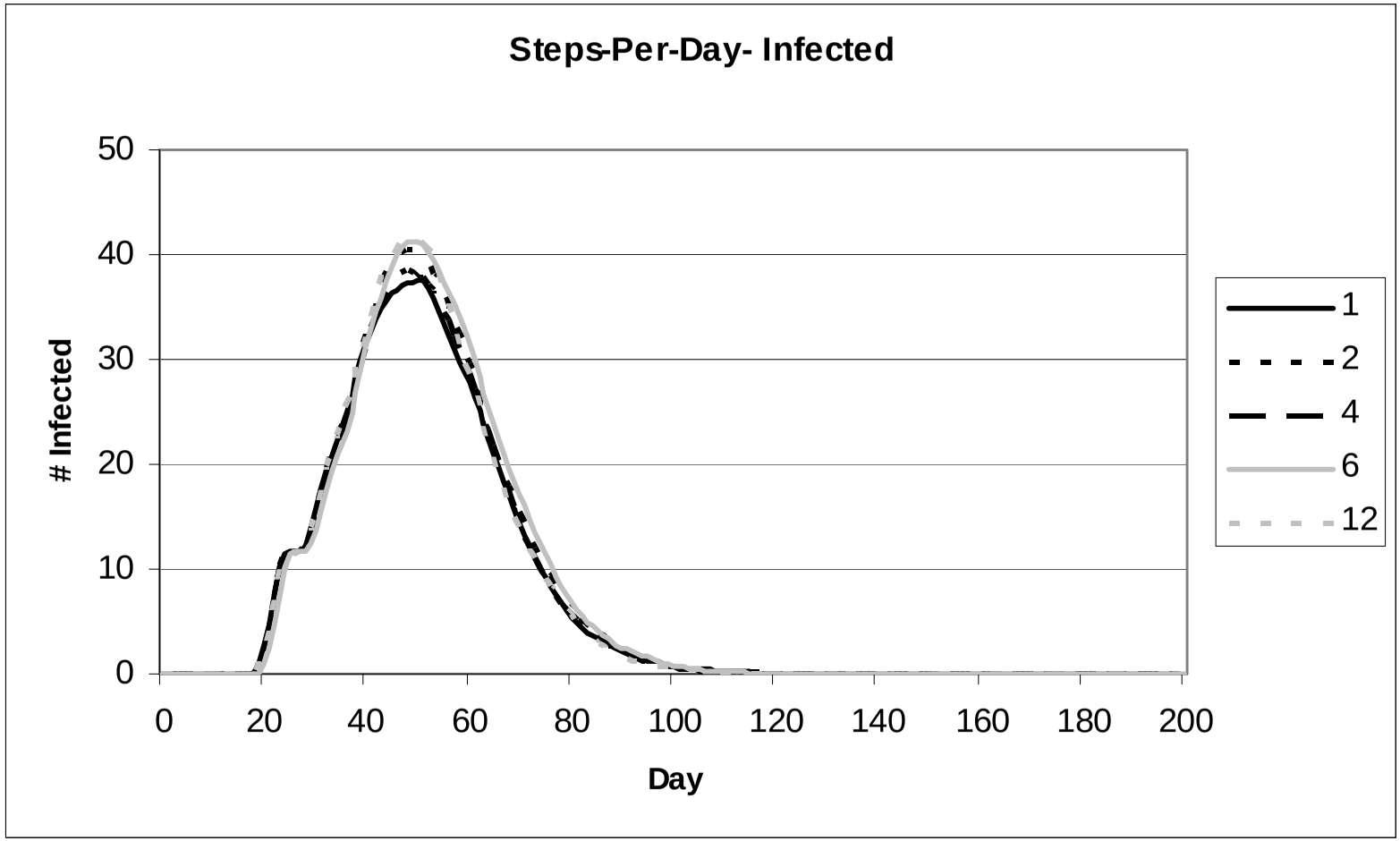
The standard value for the steps-per-day parameter is 6. This makes the journey between Norway House and Oxford House 4 days, the journey between Oxford House and God's Lake 2 days, and the journey between God's Lake and Norway House 6 days (a review of the other values for this parameter may be found in Table 5.1). Given that the standard length of the recovery period is 5 days, it should therefore be possible for the

flu to travel to Oxford House, and from there, be passed on to God's Lake. The flu would be expected to spread to Oxford House and God's Lake more frequently when steps-per-day is increased to 12. Conversely, it would be expected to reach these other communities only rarely when steps-per-day is set at 4, 2, or 1 because at these values, the length of the journey becomes longer than the length of the recovery period. Under these circumstances, the flu could only reach Oxford House or God's Lake if an infectious agent transmitted the flu to a susceptible agent that he met along the path, and then this newly infected agent carried the flu to the new community.

The steps-per-day parameter was tested at all 5 of its possible values while the other NHOHGL parameters were held constant at their standard values. The graphical results for these tests are provided in Figure 6.6 and the numerical results are provided in Table 6.5. The data display a pattern similar to that outlined in the previous paragraph. When steps-per-day is set at 6, its standard value, the flu does manage to occasionally spread to both Oxford House and God's Lake. As predicted, the flu reaches these other communities a bit more frequently when steps-per-day is set at 12 and almost never reaches them when steps-per-day is set at 1, 2, or 4. However, since the flu did make it to Oxford House twice and God's Lake once, even when steps-per-day was set at 1, it is clear that an infectious agent traveling along the path may indeed pass the flu on to a susceptible traveler, who may in turn carry it to another community, although, as expected, this is exceedingly rare. Altering the value of the steps-per-day parameter also appears to influence the timing and size of the epidemic peak, as well as the number infected and dead; however, the data only fell outside of the control range when steps-per-day was set at 1 or 4, and even then were only just below it. It is thus clear that the

Figure 6.6: Results of the steps-per-day sensitivity test

a. The number of infected agents per day of the simulation at different values for the steps-per-day parameter



b. The number of dead agents per day of the simulation at different values for the steps-per-day parameter

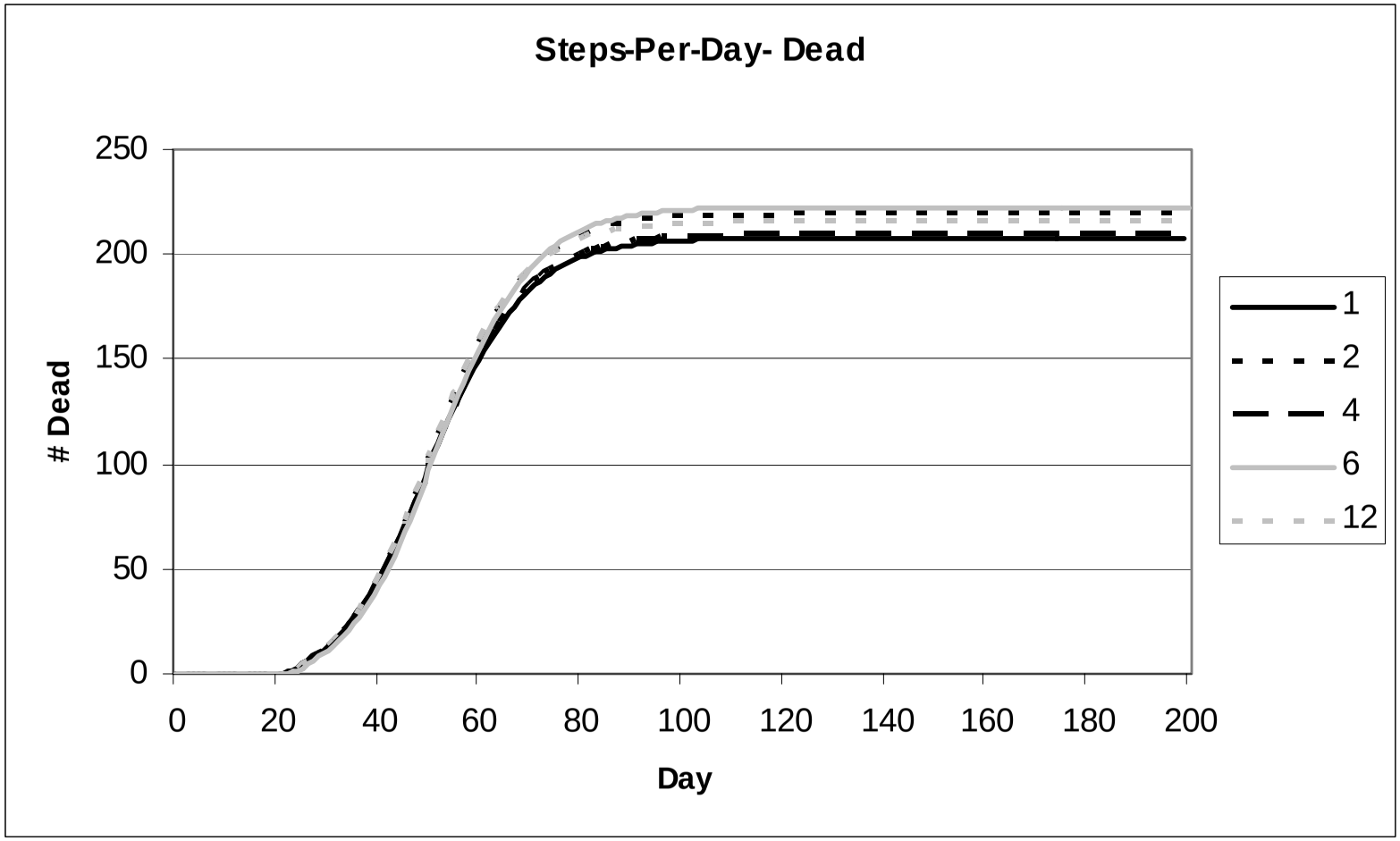


Table 6.5: Epidemic characteristics for the steps-per-day parameter

Steps Per Day	1	2	4	6*	12
Duration of Infection	115	116	132	123	118
Length of the Epidemic	112	115	129	122	117
Peak Time of Infection	32	30	29	30	31
Peak Number Infected	37.61 ↓	40.53	38.56 ↓	41.62	41.38
Total Recovered	144.37↓	151.06	145.50 ↓	153.41	149.71
Total Dead	207.70 ↓	219.13	209.93 ↓	221.19	215.67
Total Infected	352.07 ↓	370.19	355.43 ↓	374.60	365.38
% Infected	25.5% ↓	26.8%	25.8%↓	27.1%	26.5%
% Dead	15.1% ↓	15.9%	15.2% ↓	16.0%	15.6%
Epidemics at OH	2 ↓	0 ↓	4 ↓	17	27
Epidemics at GL	1 ↓	0 ↓	0 ↓	5	23 ↑
Epidemics at OH&GL	0	0	0	0	1

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

steps-per-day parameter has a much stronger impact on the frequency with which the simulated epidemic spreads between the three communities than on its epidemiological characteristics within each community.

- **Camp Stay**

The camp stay parameter controls agent mobility within each of the three communities. It specifies the probability that the agents from the camps will stay at the camps, rather than travel to their designated post, on any given day of the simulation. The camp stay parameter was the primary mobility function in the original Norway House model and was thoroughly tested during its evaluation (Carpenter, 2004). Although the form of the parameter in the NHOHGL model remains essentially identical

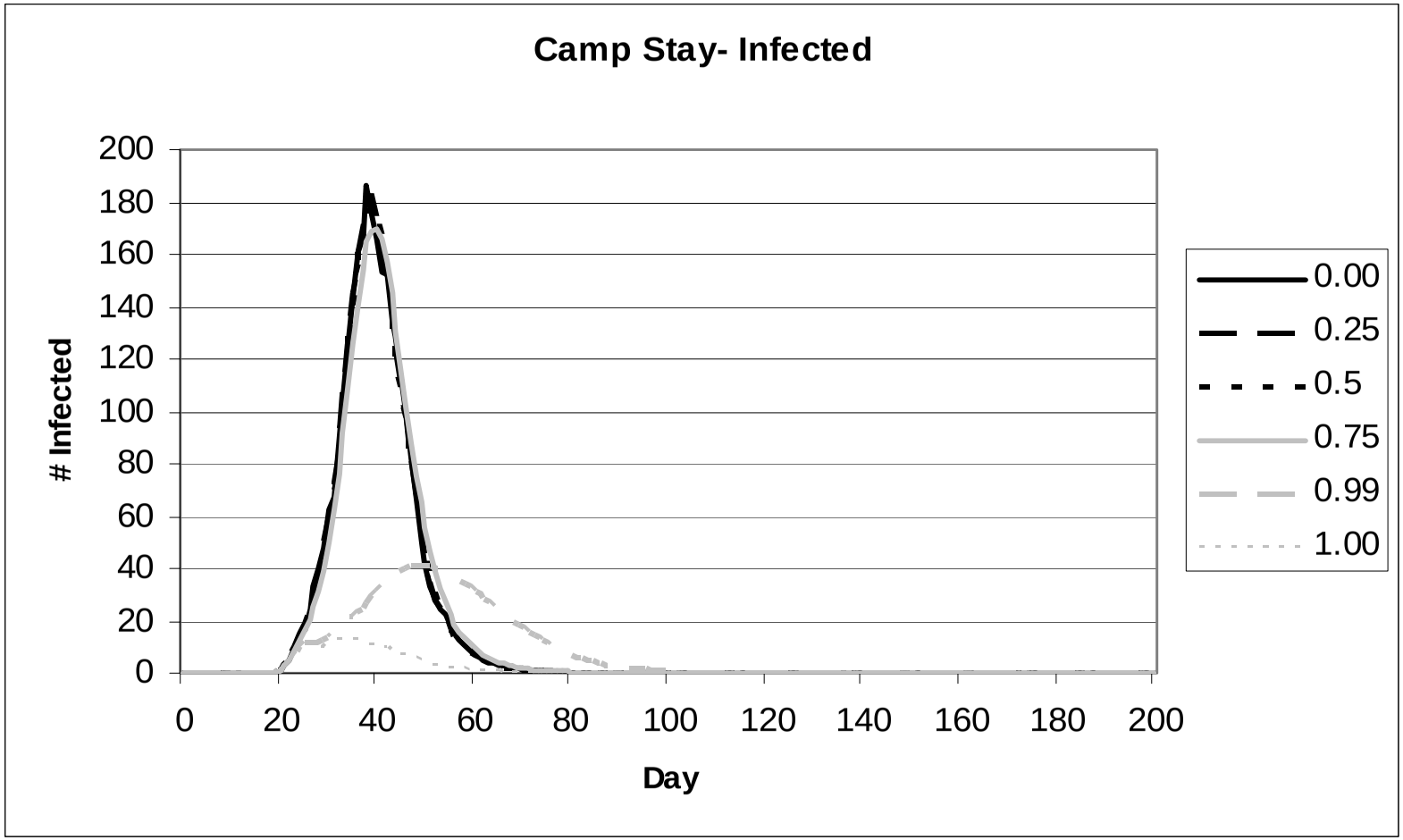
as in the original Norway House model, camp stay was tested again here due to the potentially complicating effects caused by the addition of intercommunity mobility. camp stay was tested at 6 values: 0.0; 0.25; 0.5; 0.75; 0.99 (the standard value); and 1.0. The graphical results for this sensitivity test may be viewed in Figure 6.7 while the numerical data may be found in Table 6.6.

When camp stay is set at 0.0, and travel between the camps and the central posts is nearly constant, the duration of the epidemic and peak time were shorter than the control values, whereas the height of the epidemic peak, and the numbers infected, dead, and recovered are higher. This occurs because the increased rate of travel increases the effective population size at each of the posts, allowing the agents greater opportunity to interact with one another and increasing the probability that the infectious agents will spread the flu to the susceptible agents. Further, increased travel between the camps and the posts works to ensure that the flu will nearly always reach all of the hunting camps, thus ensuring its presence in all segments of the NHOHGL population. The result is a shorter, but more intense epidemic as the flu quickly spreads to all the available susceptibles and promptly burns itself out. This same pattern was observed in the sensitivity test for the camp stay parameter in the original Norway House model (Carpenter, 2004), confirming that the original function of this parameter has been retained.

As camp stay is increased to 0.25, 0.5, and even 0.75, the same pattern of shorter and more intense epidemics is preserved but as the value of camp stay increases, the duration of the epidemic and its associated mortality lengthens, the timing of the epidemic peak increases, and the epidemic totals begin to gradually decrease. This

Figure 6.7: Results of the camp stay sensitivity test

a. The number of infected agents per day of the simulation at different values for the camp stay parameter



b. The number of dead agents per day of the simulation at different values for the camp stay parameter

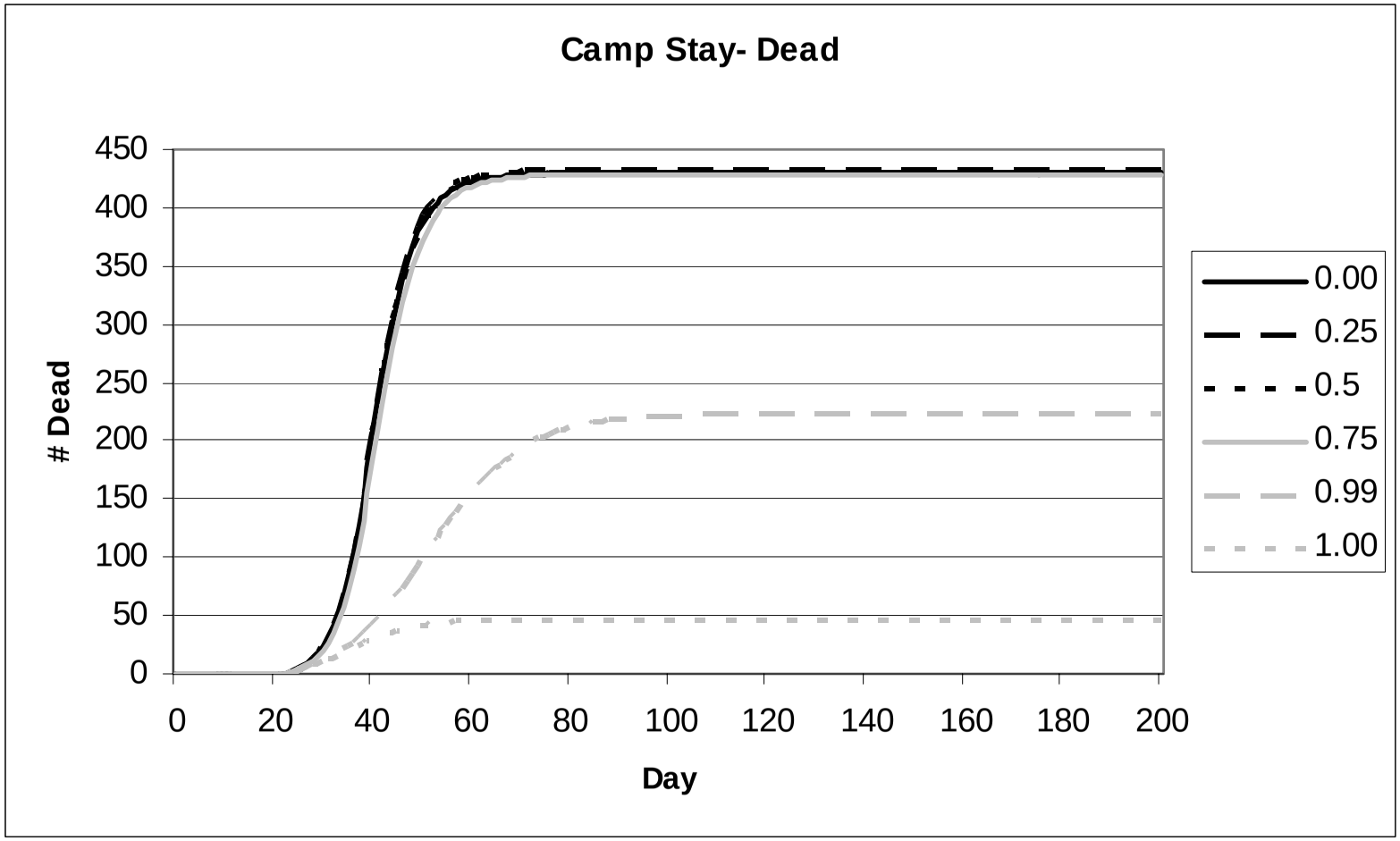


Table 6.6: Epidemic characteristics for the camp stay parameter

Camp Stay Probability	0.00	0.25	0.5	0.75	0.99*	1.00
Duration of Infection	80 ↓	79 ↓	86 ↓	95 ↓	123	76 ↓
Length of the Epidemic	79 ↓	78 ↓	82 ↓	94 ↓	122	75 ↓
Peak Time of Infection	18 ↓	19 ↓	19 ↓	20 ↓	30	14 ↓
Peak Number Infected	185.90 ↑	183.88 ↑	179.63 ↑	169.74 ↑	41.62	13.09 ↓
Total Recovered	299.09 ↑	300.29 ↑	298.28 ↑	296.93 ↑	153.41	31.99 ↓
Total Dead	429.88 ↑	432.76 ↑	430.03 ↑	428.77 ↑	221.19	46.80 ↓
Total Infected	615.78 ↑	733.05 ↑	728.31 ↑	725.70 ↑	374.60	78.79 ↓
% Infected	44.6% ↑	53.1% ↑	52.8% ↑	52.6% ↑	27.1%	5.7% ↓
% Dead	31.2% ↑	31.4% ↑	31.2% ↑	31.1% ↑	16.0%	3.4% ↓
Epidemics at OH	48 ↑	46 ↑	53 ↑	48 ↑	17	3 ↓
Epidemics at GL	12	20 ↑	14 ↑	14 ↑	5	0 ↓
Epidemics at OH & GL	1	0	0	0	0	0

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

occurs because mobility between the camps and the posts is gradually decreasing, causing interactions among the agents to decrease and allowing the flu less opportunity to spread to the camps. This slows down the epidemic and causes the total number of casualties to decrease. When camp stay is set at 1.0, travel between the camps and the posts halts entirely and the pattern changes: the duration and peak time plunge to their lowest values, far below the control range, as do the totals for the number infected, recovered, and dead. This occurs because the agents from the camps are now unable to reach the posts and the epidemic is thus restricted to the post populations which are too small to support a larger or longer epidemic on their own.

A final insight revealed by the camp stay sensitivity test was somewhat surprising. During construction of the NHOHGL model, it was assumed that only the

post agents could travel between the three communities and that the camp agents could travel only between their home camp and the central post. For this reason, camp stay, which controls these movements between the camps and the post, would be expected to influence the spread of the epidemic within a community but not between communities. However, the data from this sensitivity test indicate that the value of camp stay does in fact influence the frequency with which the flu reaches Oxford House and God's Lake. When camp stay was set at 0 (constant travel), the flu reached Oxford House 48 times, a value double the control range maximum, and reached God's Lake a total of 12 times, a value on the high end of the control range. Conversely, when camp stay was set at 1.0 (no travel), the flu reached Oxford House only 3 times and never reached God's Lake. While the reason for this observation is unclear, it does not necessarily imply that the camp agents are traveling between the communities. It is also possible that when camp stay is increased, the larger number of camp agents traveling between the camps and the central post simply makes it more likely for a post agent to become infected and then to carry the flu to a new community. This is an area for future investigation.

- **Population Size and Geographic Distribution**

As noted in the previous chapter, the size of the NHOHGL agent population and the distribution of the agents across the model landscape are determined by a series of six parameters which specify the post populations and the camp populations at each of the three communities. In the standard model scenario, these population parameters were set at the following values: at Norway House, 150 agents at the post and at each camp; at Oxford House, 66 agents at the post and at each camp; at God's Lake, 60 agents at the

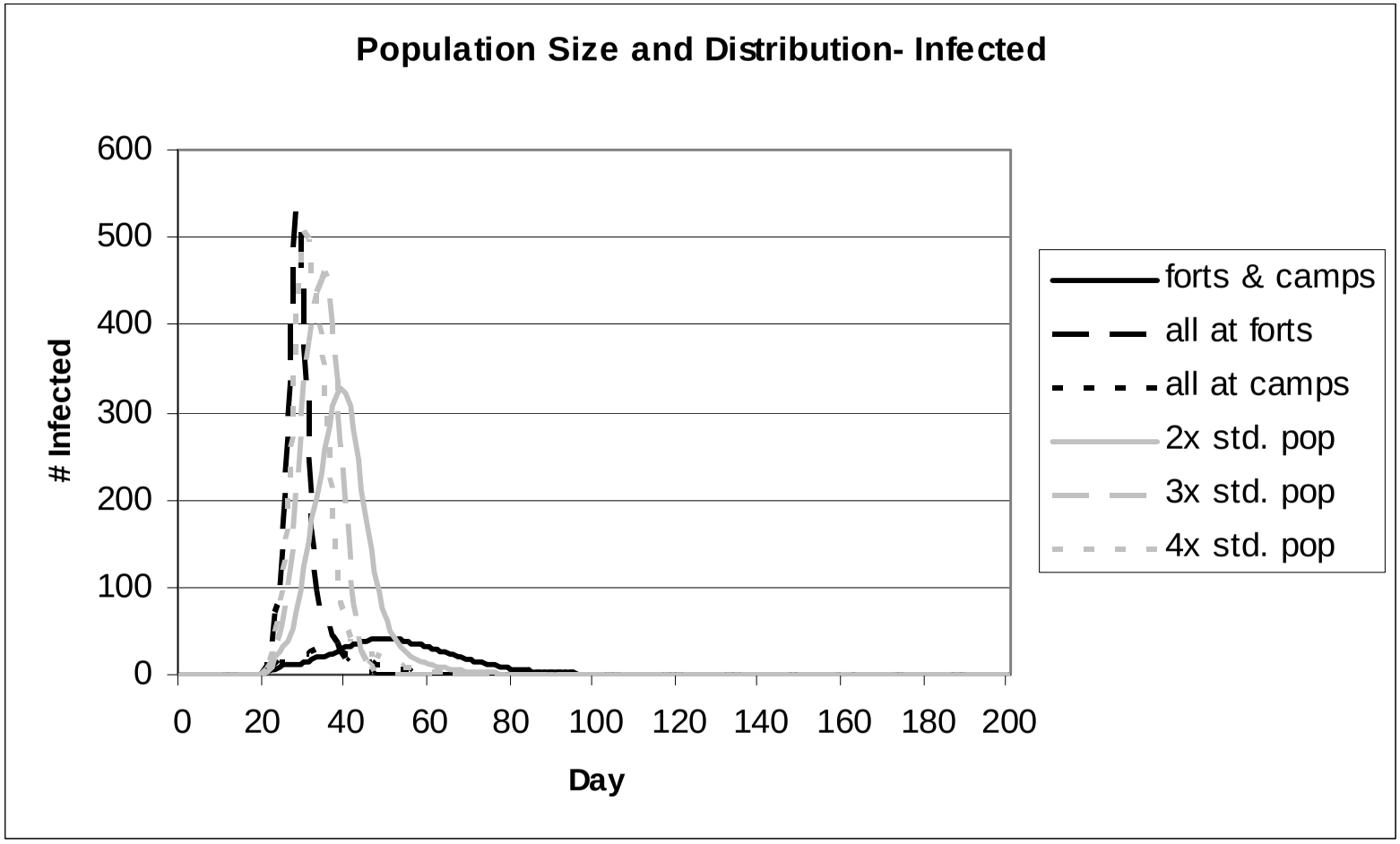
post and at each camp. The result is a total of 750 agents at Norway House, 330 at Oxford House, and 300 at God's Lake, numbers which approximate historical population estimates for these communities. When the six population parameters are altered, the number of agents and their distribution can be adjusted. For this sensitivity test, five different population scenarios were tested, in addition to the standard scenario. The results of these test simulations are shown in Figure 6.8 and Table 6.7.

In the first two test scenarios, the standard population sizes for the three communities were retained, but the distribution of the agents within each community was altered by placing the agents at only the posts or at only the camps. Placing all of the agents at the posts caused the duration of infection, length of the epidemic, and timing of the epidemic peak to fall far below the standard control range, but caused the peak number infected and the total number of infections and deaths to far exceed the control range maximums. This occurred because the more concentrated population at the posts facilitated contact among the agents, allowing the flu to spread rapidly throughout this larger population until there were no more susceptibles left to infect. Having more agents at the posts also increased the pool of intercommunity-travelers, allowing the flu to spread to Oxford House and God's Lake much more frequently than would normally be expected.

Placing all of the agents at the camps, like the previous test, also caused the duration of infection, length of the epidemic, and timing of the epidemic peak to fall below the standard control range, however the difference here was less remarkable and is explained not by an increased concentration of agents but rather by a smaller effective population size. Under the standard conditions, the NHOHGL model introduces the first

Figure 6.8: Results of the population size and distribution sensitivity test

a. The number of infected agents per day of the simulation at different values for population size and distribution parameters



b. The number of dead agents per day of the simulation at different values for the population size and distribution parameters

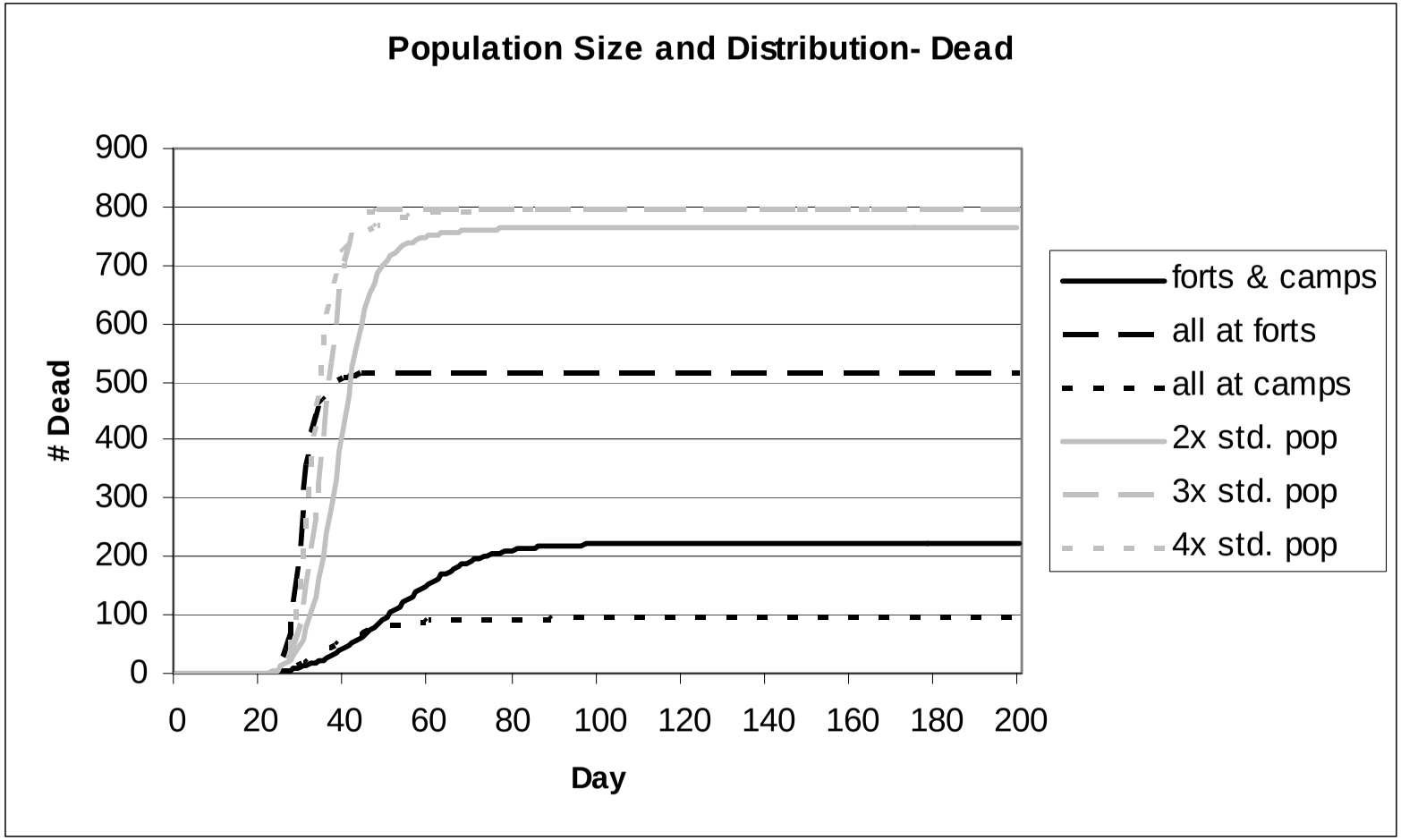


Table 6.7: Epidemic characteristics for population size and geographic distribution

Population Size and Distribution	posts & camps *	all at posts	all at camps	2x std. pop	3x std. pop	4x std. pop
Duration of Infection	123	34 ↓	97 ↓	94 ↓	38 ↓	76 ↓
Length of the Epidemic	122	32 ↓	96 ↓	90 ↓	35 ↓	74 ↓
Peak Time of Infection	30	18 ↓	14 ↓	19↓	15 ↓	10 ↓
Peak Number Infected	41.62	528.40 ↑	28.63 ↓	327.01 ↑	458.11 ↑	506.05 ↑
Total Recovered	153.41	358.91 ↑	65.50 ↓	529.87 ↑	550.73 ↑	550.83 ↑
Total Dead	221.19	516.51 ↑	94.16 ↓	764.09 ↑	794.51 ↑	793.75 ↑
Total Infected	374.60	875.41 ↑	159.67 ↓	1293.96↑	1345.24↑	1344.57↑
% Infected	27.1%	63.4% ↑	11.6% ↓	46.9% ↑	32.5% ↑	24.4% ↓
% Dead	16.0%	37.4% ↑	6.8% ↓	27.7% ↑	19.2% ↑	14.4% ↓
Epidemics at OH	17	309 ↑	0 ↓	0 ↓	0 ↓	0 ↓
Epidemics at GL	5	89 ↑	0 ↓	0 ↓	0 ↓	0 ↓
Epidemics at OH&GL	0	30 ↑	0	0	0	0

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

infectious agent into the Norway House post, but when the NH post is empty, as in this scenario, the flu is instead introduced into Norway House Camp 1. The flu is generally confined to this small population because, with no agents at the post, the only way for it to spread to the other camps would be for an infectious agent from Camp 1 to encounter an agent from another camp while both were visiting the post. Further, the lack of a post population also makes it impossible for the flu to spread to Oxford House or God's Lake because the post agents are normally the sole inter-community travelers. Thus limited to the small population of Camp 1, the flu quickly spreads to all of the available susceptibles and, having no means to spread elsewhere, promptly burns itself out. This explains not only the shorter epidemic and earlier epidemic peak, but also why the height of the

epidemic peak and the epidemic totals for this scenario all fall far below the standard control range.

While the first two population parameter tests retained the standard population size but altered the distribution of the agents, the other three tests retained the standard distribution pattern but altered the size of the agent population. This was achieved by doubling, tripling, or quadrupling the post and camp population parameters for each community. A test where the size of the population would be reduced by half had also been planned, but was unable to be performed due to a population check feature that had been built into the model program. This feature could possibly be removed if future applications of the NHOHGL program necessitate a population smaller than 750, but time did not permit such an adjustment to be made for this project.

As the number of agents was doubled, tripled, and quadrupled, some interesting trends arose in the data. First, when the size of the agent population was doubled, the duration of the epidemic dropped below the standard control range and when the size of the agent population was tripled, it became even shorter. Strangely, when the size of the agent population was quadrupled, rather than decreasing further, the length of the epidemic was instead intermediate between that of the 2x and 3x runs. The reason for this is unclear but, since a re-run of these trials produced the same result, it is likely not due to stochasticity and may instead indicate that a population threshold for the simulated epidemic had been reached during the 3x run and then breached during the 4x run.

A second notable trend was that increasing the size of the agent population elicited an epidemic peak that was both earlier and larger than the standard control range. As the size of the population increased, the epidemic peak came increasingly earlier and

the peak number infected grew increasingly large. This occurred because the growing population of agents provided the flu with a larger base of susceptibles, while placing all of these additional agents at the same locations increased population density and thereby increased the likelihood of interpersonal contact. This larger and more concentrated population allowed for more agents to become infected at the same time and, because the flu spreads exponentially, caused the epidemic to rise faster, peak earlier, and then fade away more quickly.

This same explanation may also underlie some additional interesting features noted in the population size trials. First, while the total number infected, recovered and dead was higher than the standard control range for all three tests, these statistics remained at similar values throughout the 2x, 3x and 4x trials. Because the size of the population expanded with each trial, this caused the percent infected and percent dead to progressively decrease as the size of the population increased. Further, because increasing the size of the agent population also resulted in increasingly larger epidemic peaks, it is clear that a higher and higher percentage of those who became infected were becoming infected on this single day: 24.3% for the 2x trial; 34.1% for the 3x trial; and 37.6% for the 4x trial. These observations add further support to the conclusion that a larger and more concentrated population of susceptibles caused the epidemic to spread more quickly and then burn out before reaching the entire model population.

Another interesting observation was that in all three of these population size tests, the epidemic never managed to reach Oxford House or God's Lake. This is somewhat surprising given that a larger population at the posts would be expected to increase the population of travelers and thereby increase the probability that the flu would reach

Oxford House and God's Lake. It is possible that this discrepancy indicates a malfunction of the intercommunity mobility function at high values. Alternatively, it is possible that this too results from the increased speed of the epidemic, which would reduced the window of time during which travelers leaving Norway House would be potentially infectious.

- **Proportion of Males, Females, and Children/Elders**

This series of parameters divides the agent population into three age/gender categories: adult males aged 20-50; adult females aged 20-50; and a group including all agents under the age of 20 and over the age of 50. These groups were differentiated in the NHOHGL model because the traditional division of labor in Cree society would have led to behavioral differences among these groups which might have influenced the course of the epidemic in these communities, most notably, mobility patterns. In traditional Cree society, the men would have been the primary hunters, trappers, and traders while the women, children and elders would have been in charge of the domestic household. Following this pattern, the NHOHGL model allows the adult males to travel but assumes that the adult females and children/elders remain at their home area throughout the entire simulation. The adult males are thus the only agents who are able to spread the flu across the model landscape.

In the standard model scenario, the male group makes up 25% of the population, the females make up another 25% and the children/elders make up the remaining half. These proportions are similar to the known population proportions at Norway House during the time of the 1918-1919 flu epidemic, but they can be altered by adjusting the

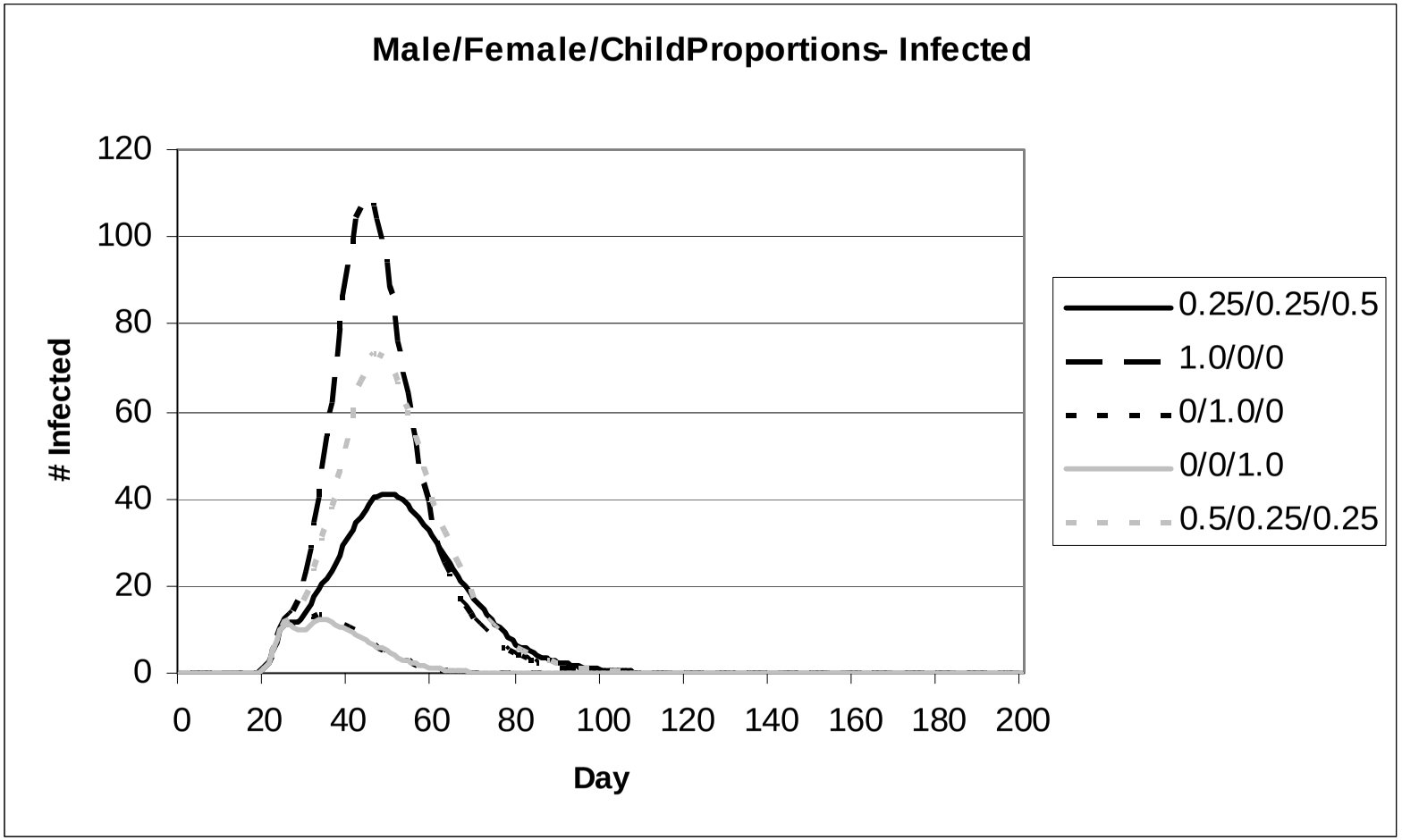
male and female proportion parameters for each post (the children/elders are simply the remaining segment of the population). For the replication test, these parameters were varied in order to produce four alternative test scenarios: one where the population consisted of only males; one where it consisted of only females; one where it consisted of only children/elders; and one where males made up half of the agent population and the females and children/elders each made up one quarter of the population. The results for these test simulations may be found in Figure 6.9 and Table 6.8.

When the population parameters were altered so that the agent population consisted entirely of adult males, the duration of infection and length of the epidemic fell within the limits of the standard control range. The peak time of infection, however, fell below the control range minima while the peak number infected, and total number infected, recovered, and dead all were above the control range maximums. Under this scenario, the flu also made it to Oxford House and God's Lake at rates well above the standard control range. This is explained by the fact that when the population of males is increased, there are more agents moving within and between the three communities. This allows the flu a greater opportunity to spread more quickly, promoting the early peak time and higher epidemic totals.

When the population was made to consist entirely of adult females or of children and elders, the result was very similar. In both cases, the duration of infection, length of the epidemic, peak time of infection, peak number infected, total number recovered, total number dead, and total number infected all fell well below the standard control range and there were no epidemics at Oxford House or at God's Lake. This occurred because, in both cases, the entire agent population was made to consist of non-travelers. With no

Figure 6.9: Results of the proportion of males, females and children/elders sensitivity test

a. The number of infected agents per day of the simulation using different proportions of males/females/children and elders



b. The number of dead agents per day of the simulation using different proportions of males/females/children and elders

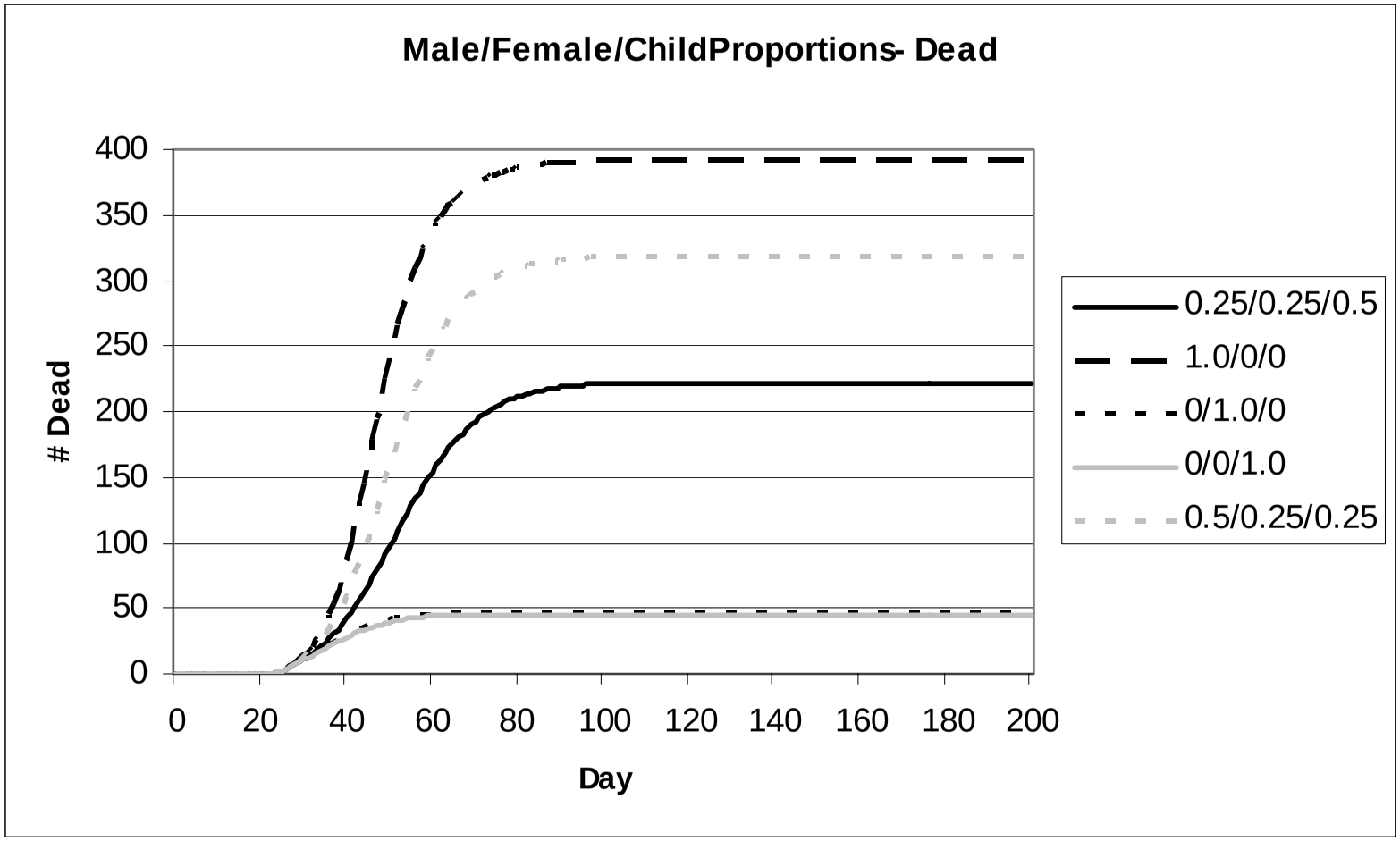


Table 6.8: Epidemic characteristics for the proportions of males, females, and children/elders

Proportion Adult Males	**0.25 ***	**1.00**	**0.00**	**0.00**	**0.5**
Proportion Adult Females	**0.25 ***	**0.00**	**1.00**	**0.00**	**0.25**
Proportion Children/Elders	**0.5 ***	**0.00**	**0.00**	**1.00**	**0.25**
Duration of Infection	123	121	72 ↓	69 ↓	106
Length of the Epidemic	122	120	70 ↓	67 ↓	105
Peak Time of Infection	30	25 ↓	13 ↓	14 ↓	27 ↓
Peak Number Infected	41.62	109.18 ↑	13.20 ↓	12.30 ↓	73.32 ↑
Total Recovered	153.41	272.09 ↑	31.75 ↓	31.51 ↓	221.75 ↑
Total Dead	221.19	392.26 ↑	46.54 ↓	45.20 ↓	318.98 ↑
Total Infected	374.60	664.35 ↑	78.29 ↓	76.71 ↓	540.73 ↑
% Infected	27.1%	48.1% ↑	5.7% ↓	5.6% ↓	39.2% ↑
% Dead	16.0%	28.4% ↑	3.4% ↓	3.3% ↓	23.1% ↑
Epidemics at OH	17	156 ↑	0 ↓	0 ↓	54 ↑
Epidemics at GL	5	32 ↑	0 ↓	0 ↓	10
Epidemics at OH & GL	0	5 ↑	0	0	0

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

agents moving between the camps and the posts or between the three posts, the flu was unable to escape the Norway House post and was effectively confined to this small population of 150 agents. These results confirm that it is indeed only the adult males who are moving within and between the three communities and indicates that the model is functioning as designed.

In the final test, half of the agents were designated adult males while one quarter were adult females and the remaining quarter were placed into the children/elders group. This effectively gave half of the agent population the ability to travel, but left the rest stationary. Under this scenario, the duration of infection and length of the epidemic both fell within the standard control range, although they remained at its lower end, the peak

number infected was slightly below the standard control range, and the peak number infected and epidemic totals all were above the control range maximums. The number of epidemics at Oxford House also exceeded the control range, but the instances at God's Lake fell within the control range, although at its higher limits. The earlier epidemic peak and higher epidemic totals are explained by the fact that the higher proportion of males in this scenario produced more travelers and thus facilitated the spread of the simulated epidemic. As would be expected, the data for this simulation fall between that of the standard scenario, where 25% of the agent population was male, and the test where 100% of the population was male. This provides further evidence that the model is functioning according to design.

- **Number of Family Groups**

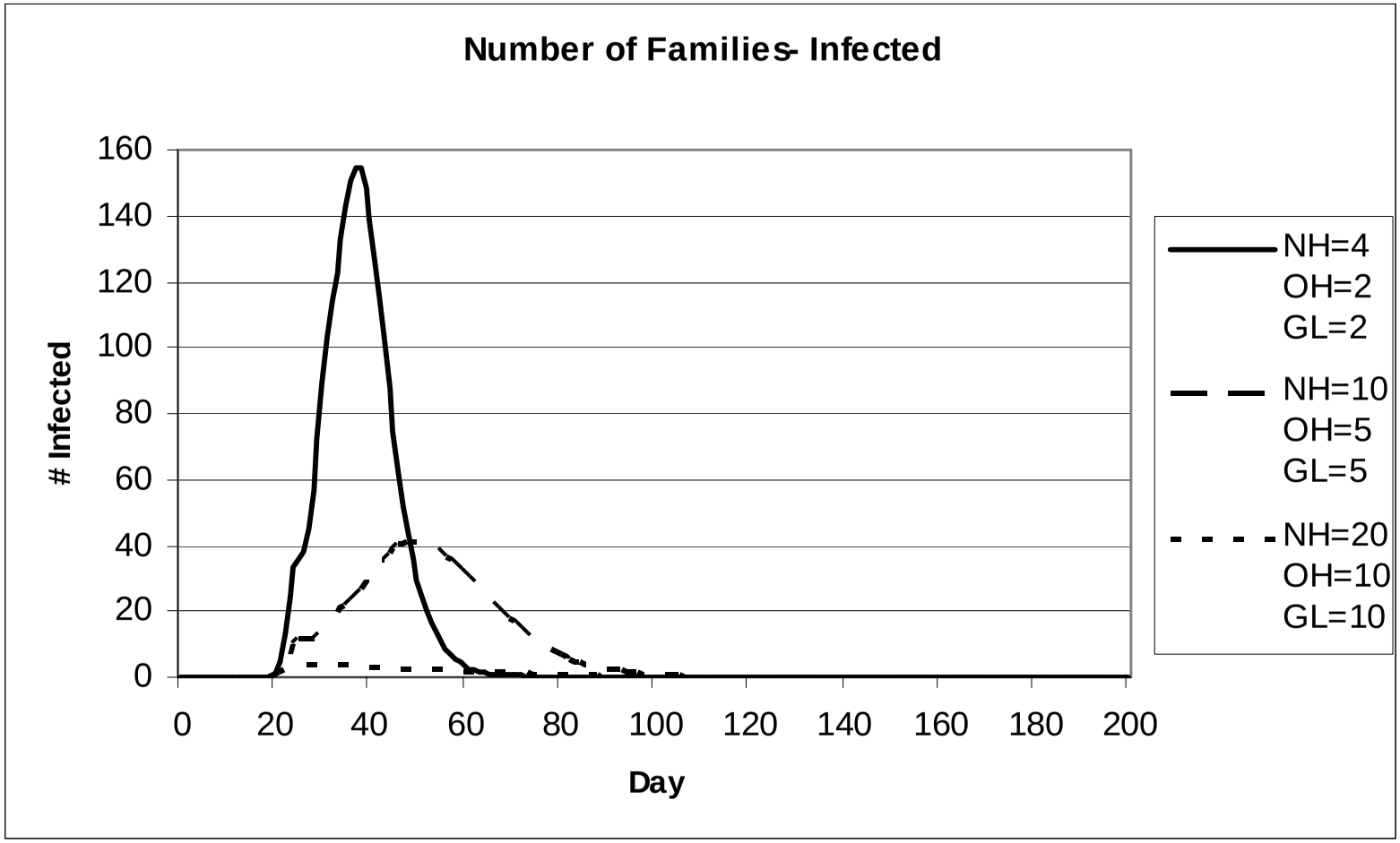
The basic building block of Cree society was the extended family hunting band and for this reason, each agent in the NHOHGL model is assigned to a family group that is meant to represent this fundamental social unit. The Number of Family Groups parameters specify the number of families per area (post and camps) for each of the three communities and, when population size remains constant, also establish the size of the family groups. The standard values for the Number of Family Groups parameters were chosen because they allow family size in the model to approximate the average size of a Cree extended family hunting band: 10 families at each area of Norway House (15 agents per family); 5 families at each area of Oxford House (13.2 agents per family); and 5 families at each area of God's Lake (12 agents per family). For the replication test, two alternative scenarios were also performed: one where the number of family groups

was increased and one where the number of family groups was decreased. The result of the standard and test simulations may be found in Figure 6.10 and Table 6.9.

Because the model assumes that the contact rate for members of the same hunting band is higher than the contact rate for members of different hunting bands, the size of the family group can have a strong impact on the simulated epidemic. In the replication test, decreasing the number of family groups, and thus increasing the size of each group, caused the duration of infection, length of the epidemic, and peak time of infection to drop below the standard control range; caused the peak number infected and total number recovered, dead, and infected to soar above it; and caused the epidemics to reach Oxford House more often and God's Lake at high levels. This is because increasing the size of the family groups facilitated contact among the agents, allowing the flu to spread faster and to a larger proportion of the NHOHGL population. Increasing the number of family groups, and thus decreasing the size of each group, had the opposite effect. This caused the duration of infection and length of the epidemic to reach the highest end of the standard control range; caused the peak number infected, and the total number recovered, dead, and infected to drop well below the control range; and led to a reduction in the incidence of the epidemic at Oxford House and God's Lake. Like the previous test, however, it also caused the peak time of infection to fall below the control range minima. This occurred because decreasing the size of the family groups reduced agent interaction, making it more difficult for the flu to spread both within Norway House and to the other communities.

Figure 6.10: Results of the number of family groups sensitivity test

a. The number of infected agents per day of the simulation using different numbers of family groups
In this graph, the numbers listed in the key indicate the number of families per post or camp at each community.



b. The number of dead agents per day of the simulation using different numbers of family groups

In this graph, the numbers listed in the key indicate the number of families per post or camp at each community.

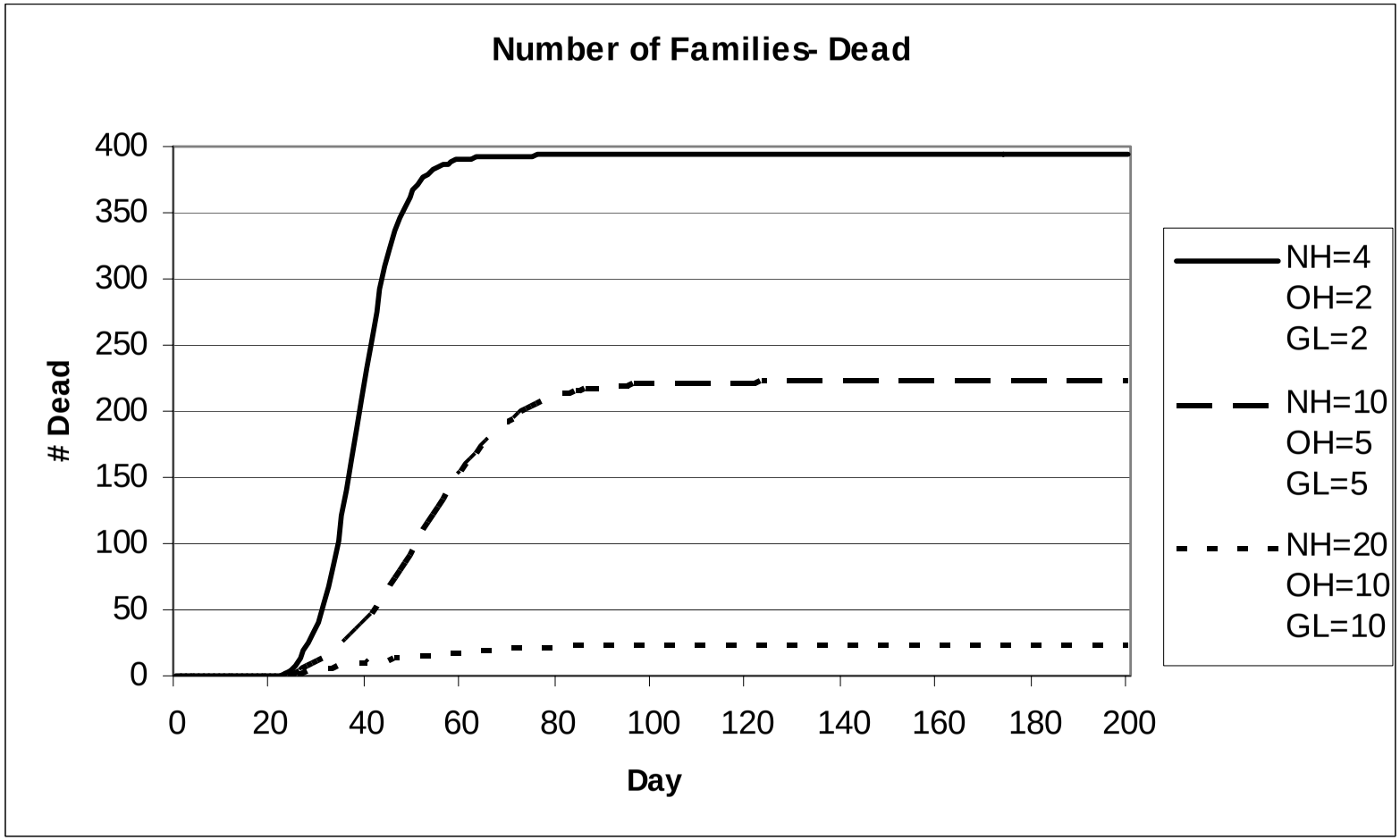


Table 6.9: Epidemic characteristics for the number of family groups

NH Families per area	4	10 *	20
OH Families per area	2	5 *	10
GL Families per area	2	5 *	10
Relative Family Size	**Large**	**Normal**	**Small**
Duration of Infection	70 ↓	123	132
Length of the Epidemic	69 ↓	122	127
Peak Time of Infection	17 ↓	30	6 ↓
Peak Number Infected	154.63 ↑	41.62	4.48 ↓
Total Recovered	273.43 ↑	153.41	16.00 ↓
Total Dead	393.56 ↑	221.19	23.40 ↓
Total Infected	667.00 ↑	374.60	39.40 ↓
% Infected	48.3% ↑	27.1%	2.9% ↓
% Dead	28.5% ↑	16.0%	1.7% ↓
Epidemics at OH	42 ↑	17	2 ↓
Epidemics at GL	10	5	0 ↓
Epidemics at OH & GL	0	0	0

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

- **Probability of Contact Within a Family Group**

The probability of contact within a family group parameter determines the probability that one family member would be in contact with another member of the same family group. As noted in the previous section, the probability of contact between members of the same family group was set higher than the probability of contact between members of different family groups, based upon the assumption that members of the same extended family hunting band would have been in closer contact with one another than with others in the community. In the standard parameter scenario, the probability of contact within a family group was set at 0.5. For this replication test, four other values

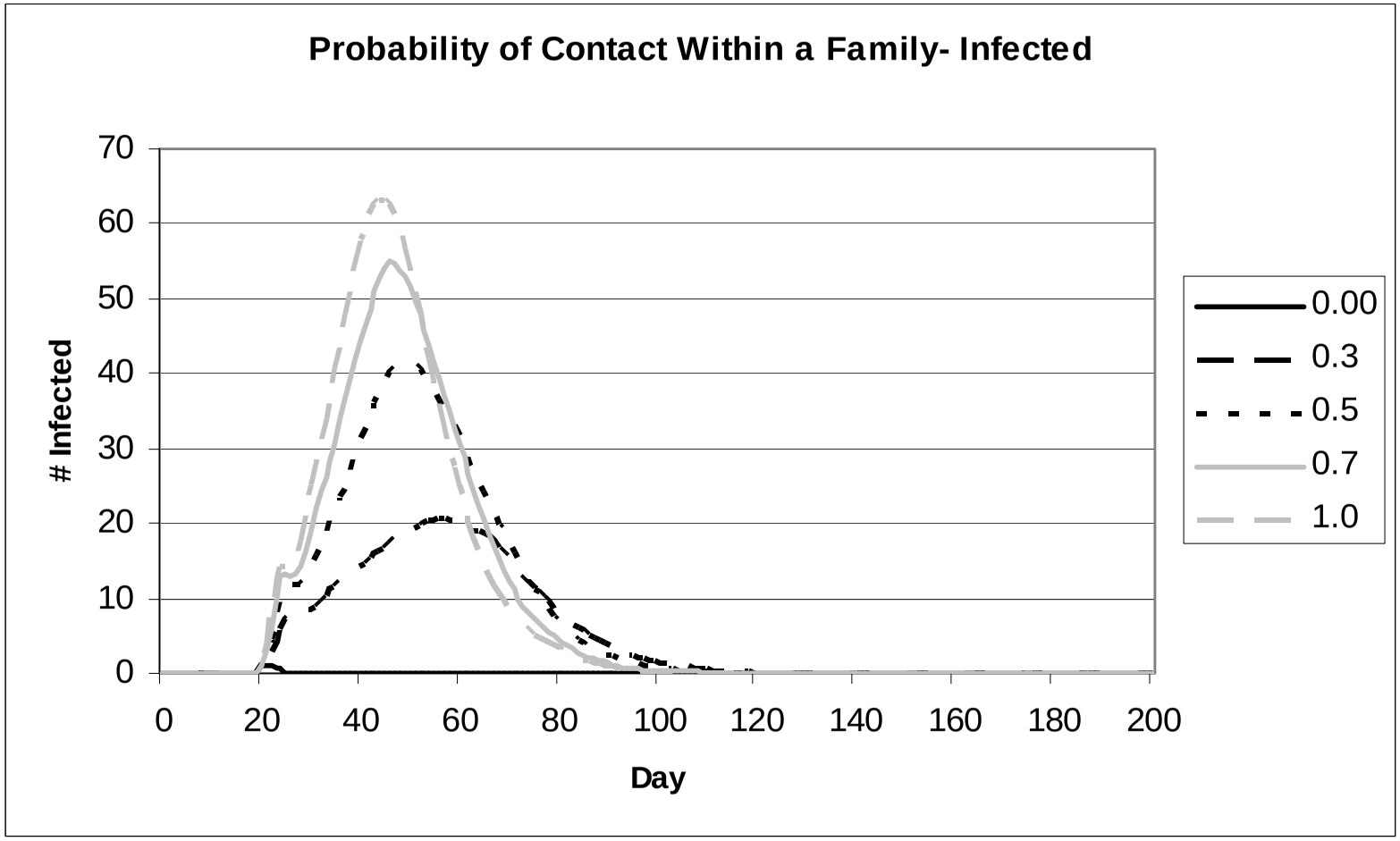
were tested: 0.0, 0.3, 0.7, and 1.0. The graphical results of this test may be found in Figure 6.11 and the numerical results may be found in Table 6.10.

Decreasing the probability of contact within a family group from 0.5 to 0.3 caused the duration of infection and length of the epidemic to reach the higher levels of the standard control range, caused the peak time of infection to occur later, and caused the total number recovered, dead, and infected to fall below the control range minims. This slower and milder epidemic occurred because when the probability of contact was decreased, the probability that an infectious agent would encounter a susceptible family member and successfully transmit the disease also decreased. Little change, however, was observed in the number of instances of the flu at Oxford House and God's Lake, which stayed well within the standard control range. These statistics were probably less affected because the spread of the flu from one community to another depends primarily upon contact between agents from different communities, and thus from different family groups.

A more dramatic result was obtained when the probability of contact within a family was set at 0. With no contact among members of the same family, the epidemic died out almost immediately: the duration of infection and length of the epidemic were both extremely short; the epidemic peak occurred just one day after the epidemic began; and the number recovered, dead, and infected were all well below the control range minima, accounting for little more than the first, introduced, infectious agent. Further, there were no epidemics at either Oxford House or God's Lake. Thus, it is clear that with no contact within the family groups, the epidemic has difficulty spreading at all. This is

Figure 6.11: Results of the probability of contact within a family group sensitivity test

a. The number of infected agents per day of the simulation at different values for the probability of contact within a family group parameter



b. The number of dead agents per day of the simulation at different values for the probability of contact within a family group parameter

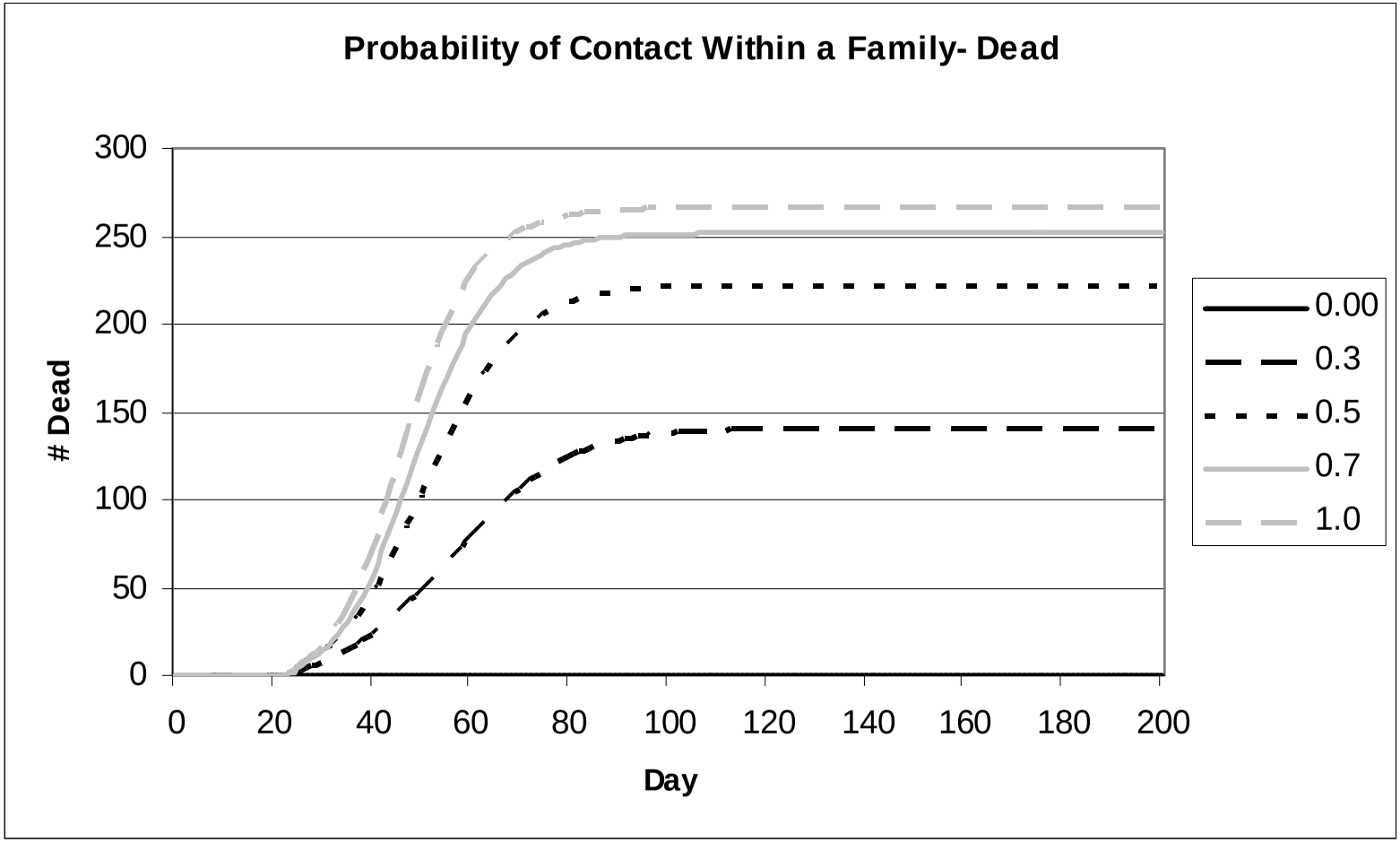


Table 6.10: Epidemic characteristics for the probability of contact within a family group

Probability of Contact Within a Family	0.0	0.3	0.5 *	0.7	1.0
Duration of Infection	14 ↓	125	123	105	109
Length of the Epidemic	12 ↓	124	122	103	108
Peak Time of Infection	1 ↓	36 ↑	30	26 ↓	25 ↓
Peak Number Infected	1.03 ↓	20.58 ↓	41.62	54.99 ↑	63.26 ↑
Total Recovered	0.46 ↓	97.05 ↓	153.41	174.46 ↑	184.41 ↑
Total Dead	0.67 ↓	140.37 ↓	221.19	251.78 ↑	266.45 ↑
Total Infected	1.13 ↓	237.42 ↓	374.60	426.24 ↑	450.86 ↑
% Infected	0.1% ↓	17.2% ↓	27.1%	30.9% ↑	32.7% ↑
% Dead	0.0% ↓	10.2% ↓	16.0%	18.2% ↑	19.3% ↑
Epidemics at OH	0 ↓	13	17	26	27
Epidemics at GL	0 ↓	3	5	8	8
Epidemics at OH & GL	0	0	0	0	0

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

because the contact rate within families is such a large portion of the total contact rates in the model.

When the probability of contact within a family group was increased, different results were observed. Increasing the value of this parameter to 0.7 caused the duration of infection and length of the epidemic to drop to the lower end of the control range; caused the epidemic peak to occur early; and caused the number recovered, dead, and infected to rise well above control range maximums. This occurred because increasing the probability of contact increased the probability of meeting an infectious agent, allowing the flu to spread more quickly to more agents and to more quickly burn through the available supply of susceptibles. Increasing the probability of contact within a family

group to 1.0 produced similar results in all areas, but caused the number of infected, recovered, and dead to rise even higher above the control range maximums. This is due to the even higher contact rates, which further facilitated the spread of the flu among the family groups. Interestingly, as when the probability of contact within a family was decreased, increasing this parameter also appeared to have very little effect on the incidence of the flu at Oxford House and God's Lake. This further supports the conclusion that the spread of the epidemic between communities is more dependent on the contact rate between members of different families than between members of the same family.

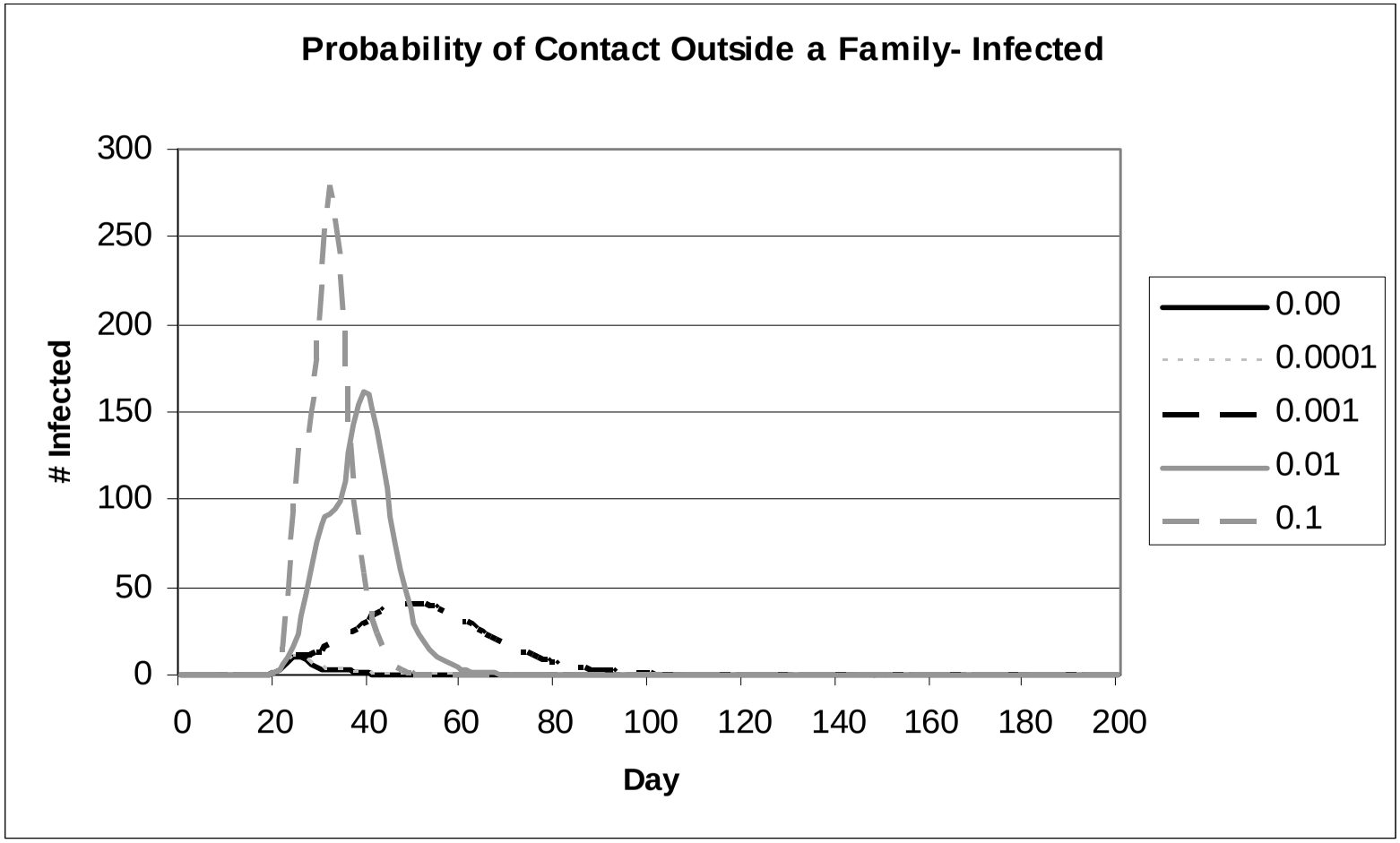
- **Probability of Contact Outside of a Family Group**

The probability of contact outside a family group parameter determines the probability that an agent from one family group will come into contact with an agent from another family group. In the NHOHGL model, this probability was lower than the probability of contact within a family group because it was assumed that there would be more frequent contact among members of the same extended family hunting band. In the standard parameter scenario, the probability of contact between non-family members was set at 0.001. For the replication test, this parameter was tested at four other values: 0.0, 0.0001, 0.01, and 0.1. The results of this test are shown in Figure 6.12 and Table 6.11.

Decreasing the probability of contact outside of a family group resulted in a decrease in the duration of infection and length of the epidemic; an earlier epidemic peak; higher numbers of recovered, infected, and dead agents; and a lack of epidemics at both Oxford House and God's Lake. This was because reducing the probability of contact

Figure 6.12: Results of the probability of contact outside of a family group sensitivity test

a. The number of infected agents per day of the simulation at different values for the probability of contact outside of a family group parameter



b. The number of dead agents per day of the simulation at different values for the probability of contact outside of a family group parameter

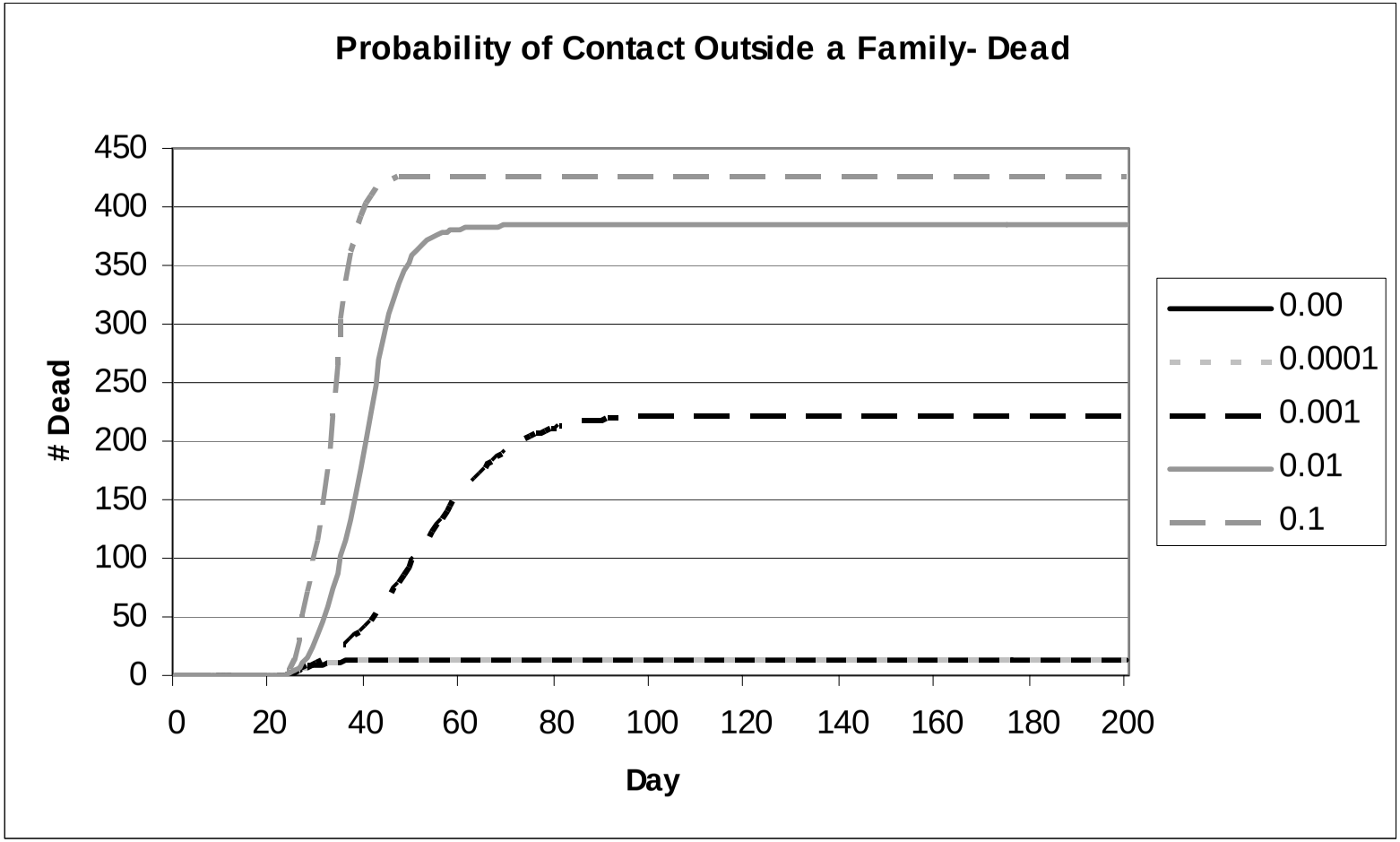


Table 6.11: Epidemic characteristics for the probability of contact outside of a family group

Probability of Contact Outside a Family	0.00	0.0001	0.001 *	0.01	0.1
Duration of Infection	34 ↓	32 ↓	123	64 ↓	41 ↓
Length of the Epidemic	32 ↓	30 ↓	122	60 ↓	35 ↓
Peak Time of Infection	5 ↓	7 ↓	30	19 ↓	12 ↓
Peak Number Infected	10.33 ↓	10.46 ↓	41.62	161.29 ↑	279.51 ↑
Total Recovered	9.46 ↓	9.56 ↓	153.41	267.45 ↑	297.10 ↑
Total Dead	13.53 ↓	14.02 ↓	221.19	384.09 ↑	427.15 ↑
Total Infected	23.00 ↓	23.58 ↓	374.60	651.53 ↑	724.25 ↑
% Infected	1.7% ↓	1.7% ↓	27.1%	47.2% ↑	52.5% ↑
% Dead	1.0% ↓	1.0% ↓	16.0%	27.8% ↑	31.0% ↑
Epidemics at OH	0 ↓	0 ↓	17	65 ↑	78 ↑
Epidemics at GL	0 ↓	0 ↓	5	15 ↑	23 ↑
Epidemics at OH & GL	0	0	0	3 ↑	2 ↑

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

between agents from different family groups reduces the probability that the flu will be transmitted from one family to another. This localizes the epidemic in one family and makes it difficult for the epidemic to spread within a community or to reach Oxford House or God's Lake. The result was the same when the parameter was reduced to 0.0001 and when contact between family groups was halted at 0.0. This indicates that any reduction in the probability of contact outside of a family group below the standard control value is likely to inhibit the spread of the simulated epidemic.

Increasing the probability of contact outside of a family group also caused the duration of infection and length of the epidemic to fall below the standard control range and led to an earlier epidemic peak. Here, however, the reason for the accelerated

epidemic was not the failure of the flu to spread but rather an increased rate of spread due to the increased rate of contact among the agents. This caused the flu to spread rapidly to all the available susceptibles and then quickly burn out when there were no more to infect. This also led to a higher epidemic peak, greater numbers of recovered, dead, and infected agents, and an increased incidence of the flu at Oxford House and God's Lake. These numbers exceeded the control range maximums when the parameter value was raised to 0.01 and were even higher when the value was raised to 0.1. This indicates that the flu is indeed able to spread more easily when the contact rate between agents from different families is increased and shows that this parameter is functioning as designed.

- **Time of First Infection**

The time of first infection parameter specifies the day of the simulation on which the first infectious individual will be introduced into the Norway House post, thereby initiating the simulated epidemic. In the standard parameter scenario, the flu was introduced on the twentieth day of the simulation, the same value used in the original Norway House model. This value was originally chosen because it would allow the agents time to begin moving about the landscape before the initiation of the epidemic. Although the sensitivity test of this parameter in the original Norway House model showed that it actually made no significant difference if the flu was introduced on day 20 or on day 1, this value was retained in the NHOHGL model scenarios because it made the graphical data more clear and because it allowed sufficient time for the simulated epidemic to run its course before the end of the simulation period. This parameter has maintained the same form in the NHOHGL model as in the original Norway House

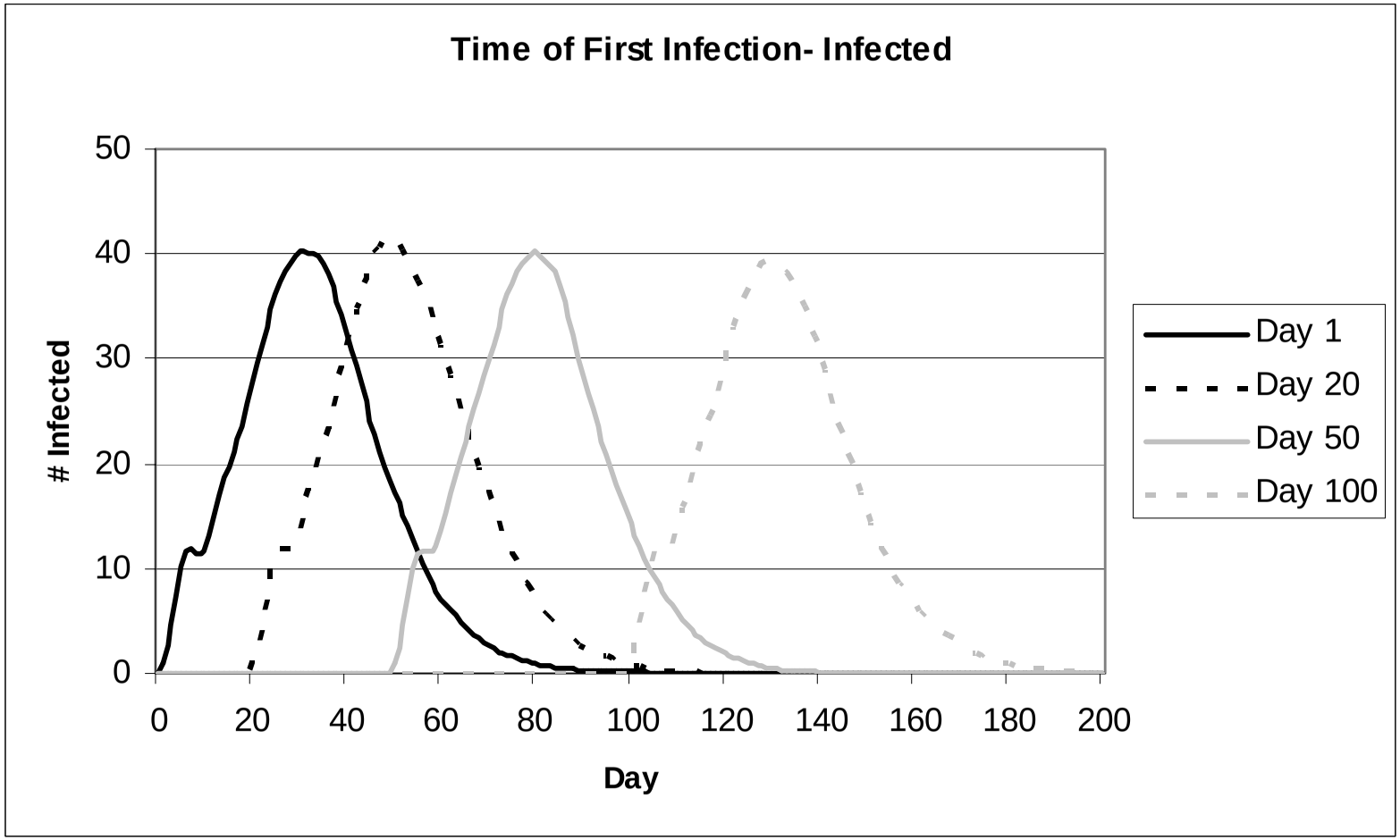
model, but was tested again here in order to establish that day 20 was an appropriate value for this parameter in the newly extended model. The results of this test can be found in Figure 6.13 and Table 6.12.

The graphs and statistics for this sensitivity test indicate that, as in the original Norway House model, the time of first infection in the NHOHGL model has very little impact on the course of the simulated epidemic as long as there is enough time for the epidemic to finish before the end of the simulation, even though the timing of the epidemics is clearly variable. The graphs in Figure 6.13 show that the overall size and shape of the morbidity and mortality curves is nearly identical for all four trials. Further, the morbidity graph helps to verify that the epidemic is initiated on different days in each trial, thus indicating that the time of first infection parameter is functioning properly.

The epidemic statistics listed in Table 6.12 reveal a very similar picture. When the epidemic is introduced on the first day of the simulation, the resulting epidemic data all fall within the standard control range. When the epidemic is introduced on the fiftieth day of the simulation, most of the epidemic data again fall within the standard control range except for the epidemic totals, which are slightly below the standard control range values. These lower values may be the result of an especially mild epidemic or they may be an indication that the simulated epidemics did not always have sufficient time to finish before the end of the simulation on day 200, which could have led to an underestimation of the total number of recovered, dead, and infected agents at the end of the simulation. When the epidemic is introduced on day 100 of the simulation, the lack of sufficient time for the epidemic to run its course is more obvious. Although the timing and size of the epidemic peak and the epidemic totals all fall within the standard control range (albeit in

Figure 6.13: Results of the time of first infection sensitivity test showing

a. The number of infected agents per day of the simulation at different values for the time of first infection parameter



b. The number of dead agents per day of the simulation at different values for the time of first infection parameter

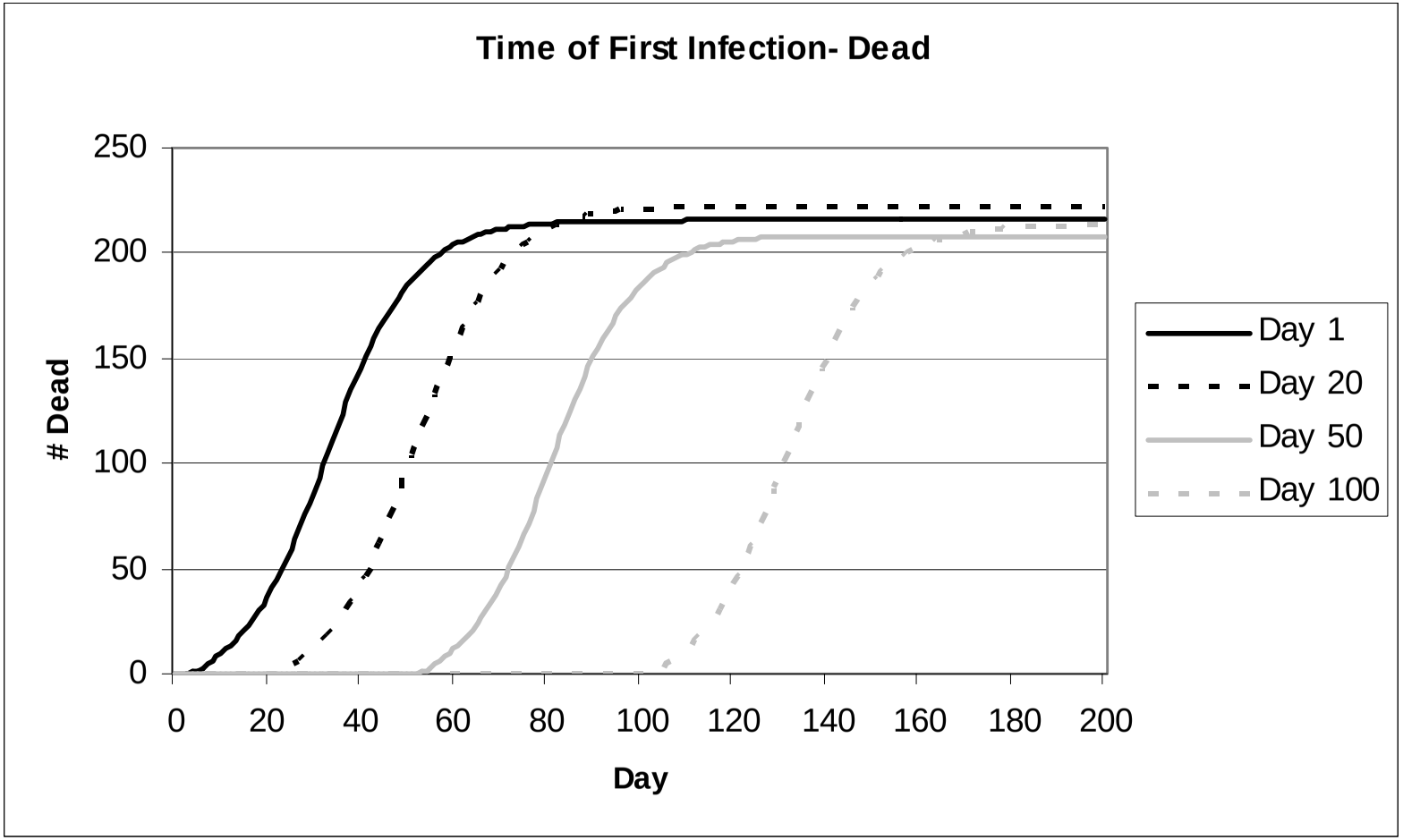


Table 6.12: Epidemic characteristics for the time of first infection

Time of First Infection	1	20 *	50	100
Duration of Infection	125	123	114	101 ↓
Length of the Epidemic	124	122	113	99 ↓
Peak Time of Infection	30	30	30	30
Peak Number Infected	40.36	41.62	40.22	39.61
Total Recovered	149.45	153.41	143.35 ↓	148.89
Total Dead	215.64	221.19	208.13 ↓	213.31
Total Infected	365.09	374.60	351.48 ↓	362.20
% Infected	26.5%	27.1%	25.5% ↓	26.2%
% Dead	15.6%	16.0%	15.1% ↓	15.5%
Epidemics at OH	18	17	21	11
Epidemics at GL	5	5	2	4
Epidemics at OH & GL	1	0	0	0

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

its lower end), the duration of infection and length of the epidemic fall below the control range minima, indicating an epidemic that was halted when the simulation ended rather than an epidemic that ended naturally. This indicates that the simulated epidemic did not have enough time to reach its peak before the simulation ended. The data from any such truncated run would be severely compromised; thus, this test illustrates the importance of including enough time steps in the simulation process in order to allow for the simulated epidemic to run its course.

- **Location of First Infection**

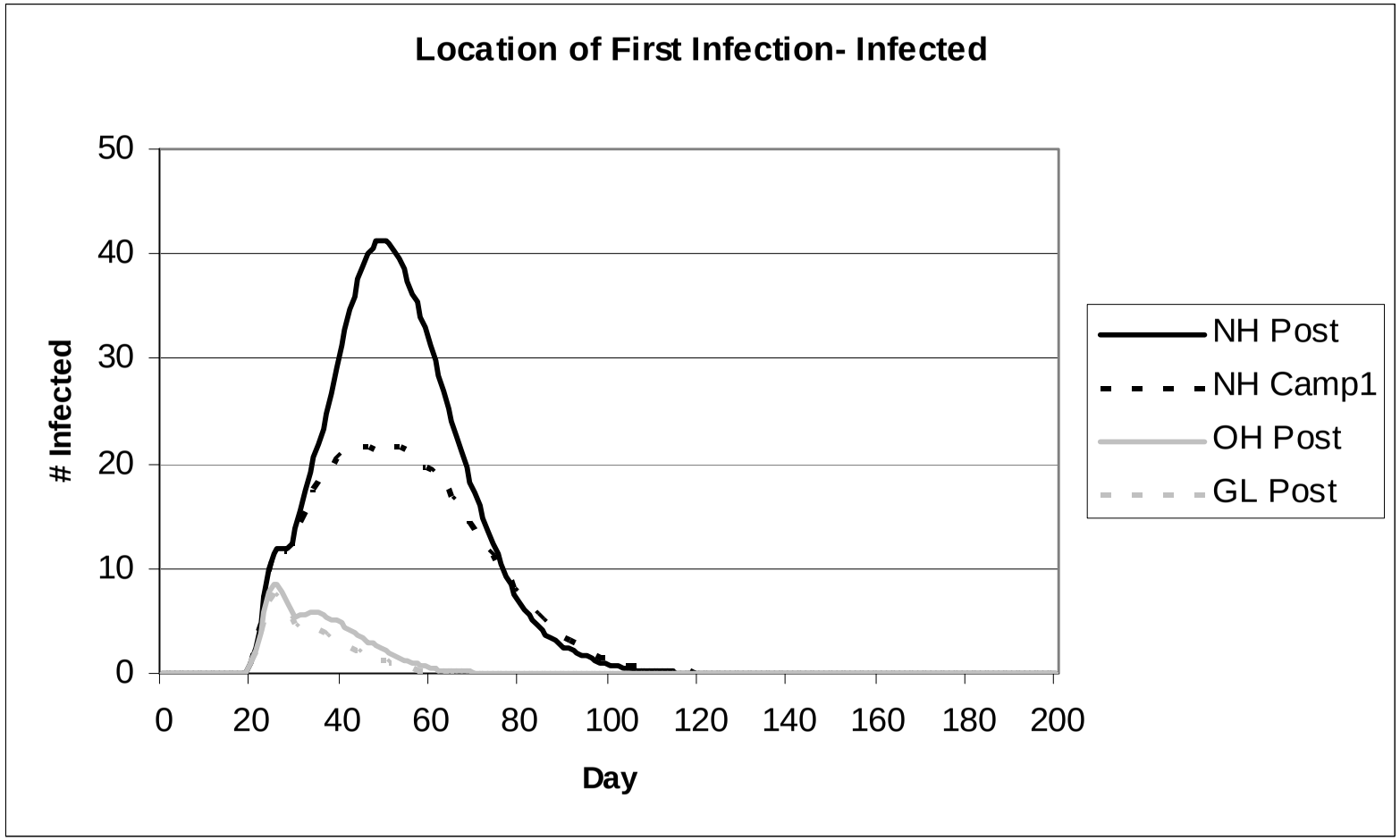
The location of first infection is simply the site at which the first infectious agent is introduced into the NHOHGL model. In the standard model scenario, the flu is always

introduced into the Norway House Post. This is because historical data from the HBC post journals indicate that this was where the 1918-1919 flu epidemic first entered the Norway House community (see chapter 3). Although the original Norway House model also introduced the flu into the Norway House post, no test was performed to determine the effect of varying the site of the flu's initial introduction. Such a test was deemed important for the NHOHGL model because this model offered more, and more realistic, alternative sites of introduction and because varying the site of disease introduction would provide insight into inter-community epidemic spread. The data for this replication test may be found in Figure 6.14 and Table 6.13.

When the first infectious agent was introduced into Norway House Camp 1 rather than the post, the resulting epidemic was noticeably milder. The duration of infection and length of the epidemic both fell within the standard control range, as did the number of epidemics at Oxford House and God's Lake, but the peak time of infection was slightly early, the peak number infected was below the control range minima, and the total numbers recovered, infected, and dead were also below the standard control range. This would be the expected result of introducing the epidemic into a less centralized and more remote location. Because the camp agents may only travel to and from the central post and are unable to visit other camps, when the flu is introduced into Camp 1, an agent from Camp 1 must carry the flu to the central post and successfully transmit the flu to the post agents if the epidemic is to spread to the other Norway House camps or to Oxford House or God's Lake. This makes it more difficult for the flu to spread, reducing the number of infected agents and producing an early epidemic peak when the flu quickly

Figure 6.14: Results of the location of first infection sensitivity test

a. The number of infected agents per day of the simulation at different values for the location of first infection parameter



b. The number of dead agents per day of the simulation at different values for the location of first infection parameter

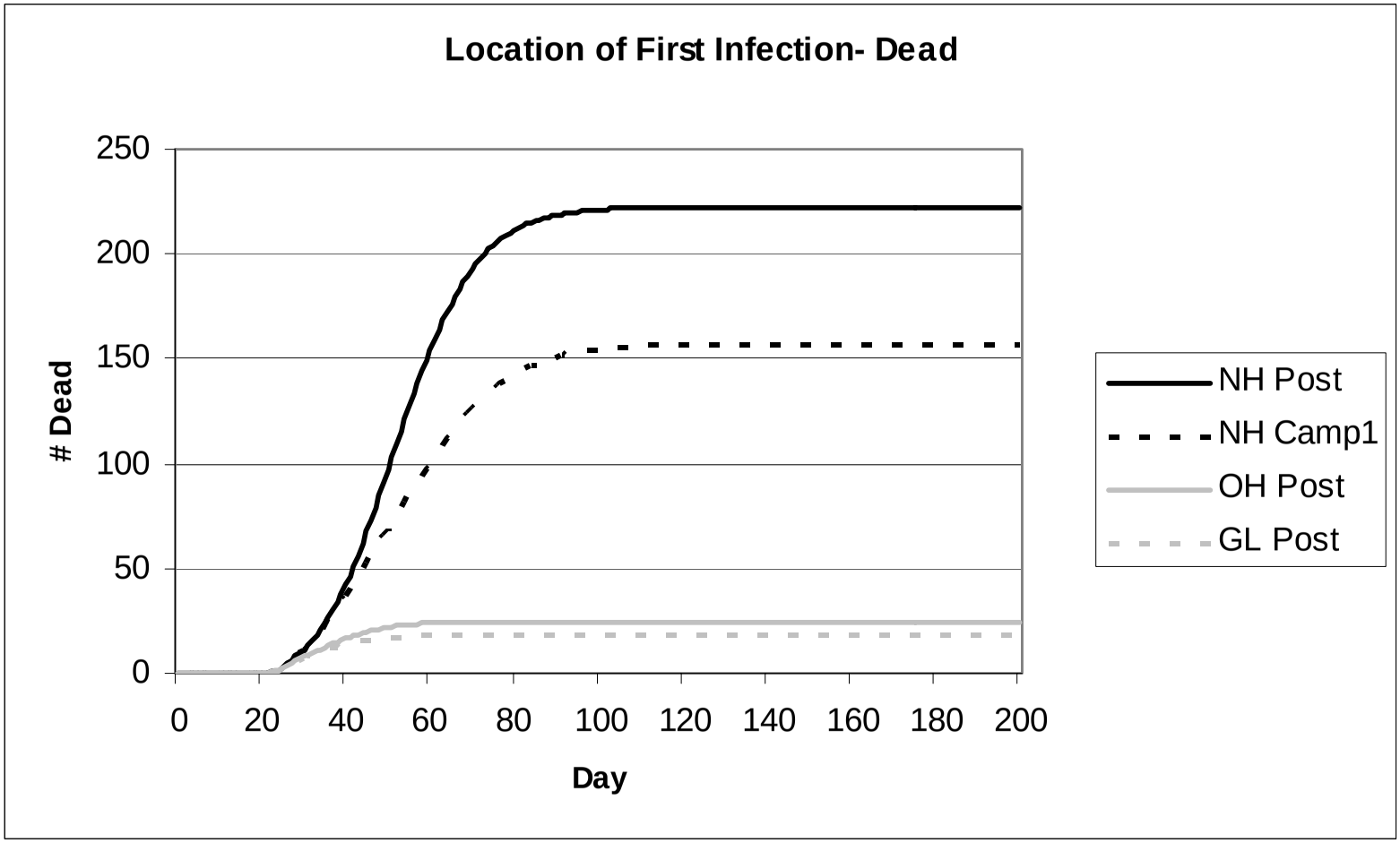


Table 6.13: Epidemic characteristics for the location of first infection

Location First Infection	NH Post*	NH C1	OH Post	GL Post
Duration of Infection	123	126	68 ↓	76 ↓
Length of the Epidemic	122	118	65 ↓	75 ↓
Peak Time of Infection	30	25 ↓	13 ↓	13 ↓
Peak Number Infected	41.62	21.55 ↓	5.86 ↓	4.41 ↓
Total Recovered	153.41	108.34 ↓	17.07 ↓	12.44 ↓
Total Dead	221.19	156.52 ↓	24.57 ↓	18.13 ↓
Total Infected	374.60	264.86 ↓	41.64 ↓	30.56 ↓
% Infected	27.1%	19.2% ↓	3.0% ↓	2.2% ↓
% Dead	16.0%	11.3% ↓	1.8% ↓	1.3% ↓
Epidemics at NH	1000	1000	0 ↓	0 ↓
Epidemics at OH	17	16	1000 ↑	0 ↓
Epidemics at GL	5	4	0 ↓	1000 ↑
Epidemics at OH & GL	0	0	0	0
Epidemics at NH & GL	5	4	0 ↓	0 ↓
Epidemics at NH & OH	17	16	0 ↓	0 ↓

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

infects all the available Camp 1 susceptibles and then has difficulty finding a fresh supply.

When the flu was introduced to the Oxford House post instead of the Norway House post, a shorter and milder epidemic ensued: the duration of infection, length of the epidemic; peak time of infection; peak number infected; and total number recovered, dead, and infected were all well below the control range minima. Further, the flu failed to spread outside of the Oxford House community. Because the population at Oxford house was nearly half that of Norway House, when the epidemic was initiated and confined to this community, it was unable to infect a large number of agents and quickly ran through the available supply of susceptibles. This led to an earlier epidemic

peak and a shorter epidemic duration. The fact that the flu never spread to Norway House or nearby God's Lake reflects the smaller population at Oxford House, which limited the pool of available intercommunity travelers and thus reduced the probability that the flu would successfully reach the other communities.

When the epidemic was introduced to the God's Lake post, the results were quite similar to when the epidemic was introduced at Oxford House. This also resulted in a shorter duration of infection and mortality, an earlier time of peak infection, a lower epidemic peak, and epidemic statistics that were well below the control range minima. The reasons here were the same as for Oxford House; the population was smaller and the flu was unable to spread to the other communities due to the smaller pool of travelers.

These tests indicate that the location of introduction can have a large impact on the extent of disease diffusion in this region of Canada. If a disease is introduced to Norway House rather than to less central Oxford House or even more remote God's Lake, it will be more likely to reach the other communities than if it were introduced elsewhere. Further, if a disease is introduced to a central post rather than at one of the camps, it would be more likely to spread within the community. These observations could have important implications for disease control, but it is important to note that the qualities that make the Norway House post the most important focus of local epidemics, namely its larger size and more central location, also make it the most likely site of disease introduction. It is thus likely that most epidemics in this region of Manitoba were initiated at the Norway House Post.

- **Disease Probability**

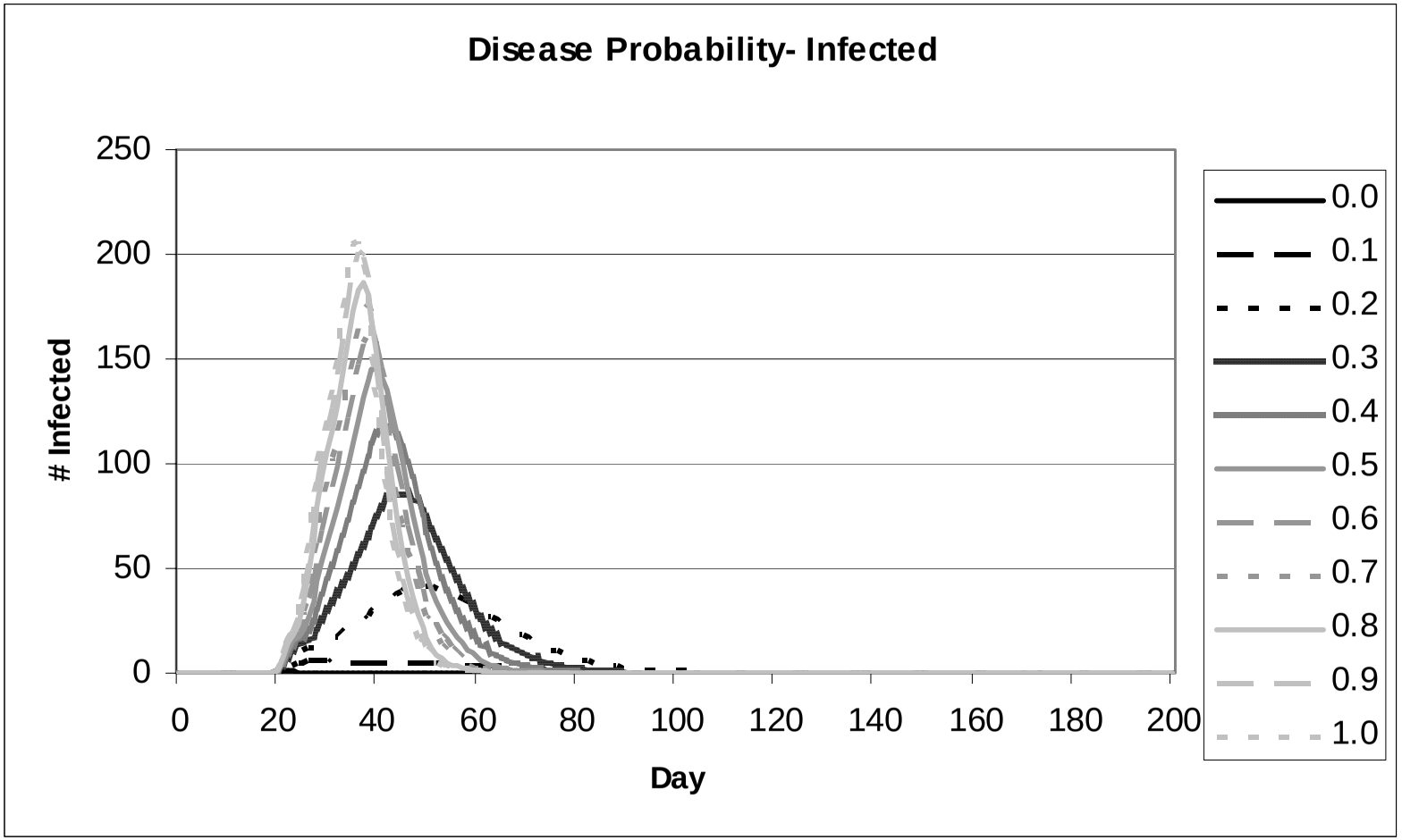
The disease probability parameter specifies the probability that a susceptible agent will become infected upon coming into contact with an infectious agent. It thus would be expected to play a major role in the spread of the simulated epidemic. Although the value of this parameter can be varied across the three age/gender groups, this feature was not used in this project. Instead, the standard value for disease probability has been set at 0.2 for all three age/gender groups, which was the value used in the original Norway House model. In the replication test, the disease probability parameter was tested at a wide range of values. The results of these test simulations can be viewed in Figure 6.15 and Table 6.14.

When the probability of disease was reduced from its standard value of 0.2 to 0.1, the course of the simulated epidemic was clearly affected. Although the duration of infection and length of the epidemic remained well within the standard control range, the peak time of infection dropped below control range minima, and so did the number of recovered, dead, and infected agents and the number of epidemics at Oxford House and God's Lake. The reduced disease probability meant that fewer of the contacts between infectious and susceptible agents would result in disease transmission. This made it more difficult for the flu to spread, causing it to infect fewer agents, peak earlier, and stay confined to the Norway House community.

When the disease probability was decreased even further to a value of 0, none of the contacts between infectious agents and susceptible agents would be expected to result in transmission of the flu. The resulting epidemic has an extremely short duration which is equal to the length of the recovery period, an early epidemic peak that occurs just one

Figure 6.15: Results of the disease probability sensitivity test

a. The number of infected agents per day of the simulation at different disease probabilities



b. The number of dead agents per day of the simulation at different disease probabilities

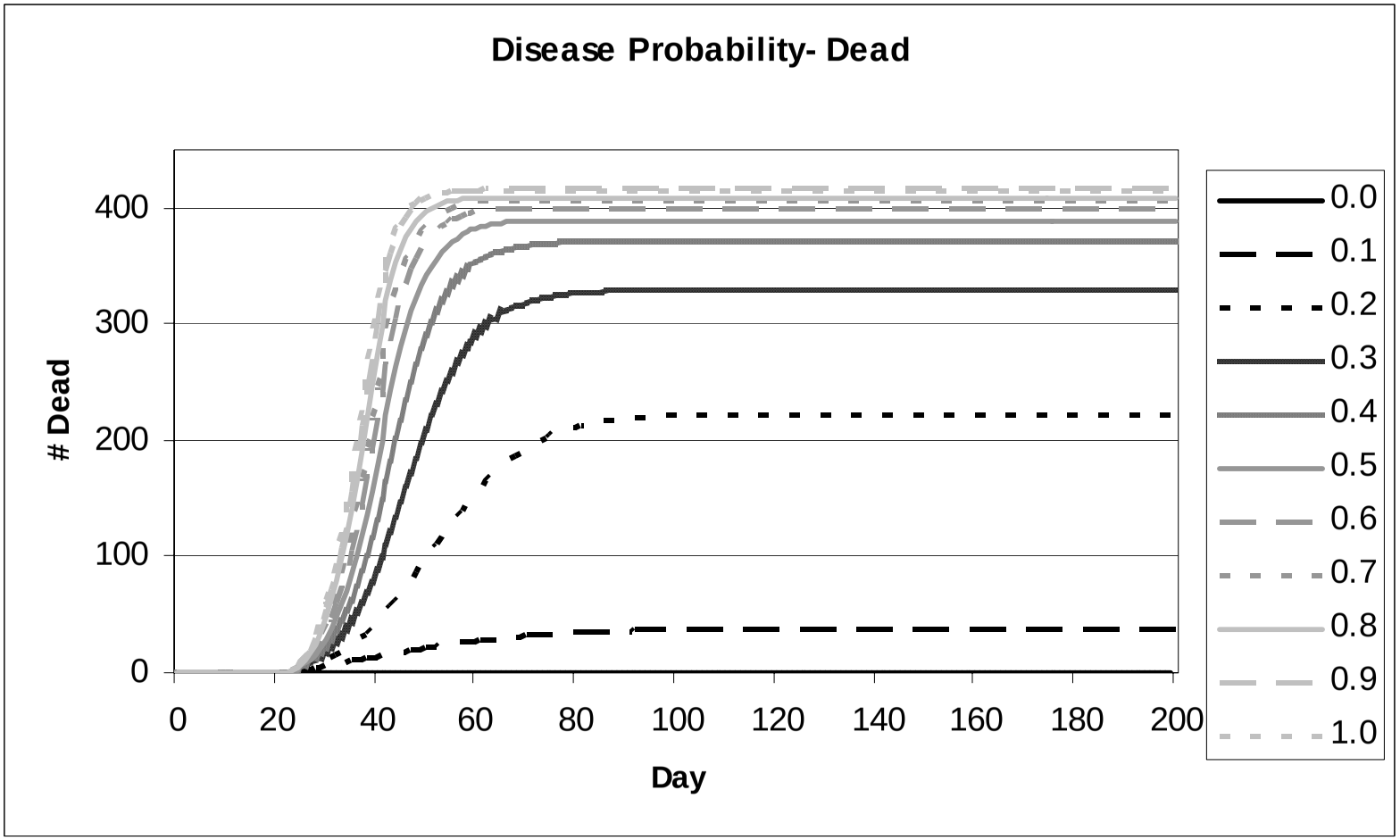


Table 6.14: Epidemic characteristics for disease probability

Disease Probability	0.0	0.1	0.2 *	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Duration of Infection	5 ↓	122	123	95 ↓	79 ↓	74 ↓	70 ↓	61 ↓	65 ↓	57 ↓	54 ↓
Length of the Epidemic	4 ↓	121	122	93 ↓	77 ↓	72 ↓	68 ↓	59 ↓	64 ↓	55 ↓	49 ↓
Peak Time of Infection	1 ↓	7 ↓	30	25 ↓	22 ↓	20 ↓	19 ↓	18 ↓	17 ↓	16 ↓	16 ↓
Peak Number Infected	1.00 ↓	6.43 ↓	41.62	85.68 ↑	118.29 ↑	145.38 ↑	163.18↑	175.39 ↑	186.21↑	200.66 ↑	206.60↑
Total Recovered	0.37 ↓	26.31 ↓	153.41	228.34 ↑	258.36 ↑	270.21 ↑	277.27 ↑	281.66 ↑	285.02 ↑	288.73 ↑	287.61 ↑
Total Dead	0.63 ↓	37.38 ↓	221.19	329.36 ↑	371.01 ↑	389.45 ↑	399.92 ↑	406.37 ↑	409.29 ↑	416.32 ↑	415.58 ↑
Total Infected	1.00 ↓	63.69 ↓	374.60	557.71 ↑	629.37 ↑	659.66 ↑	677.19 ↑	688.03 ↑	694.30 ↑	705.05 ↑	703.19 ↑
% Infected	0.1% ↓	4.6% ↓	27.1%	40.4% ↑	45.6% ↑	47.8% ↑	49.1% ↑	49.9% ↑	50.3% ↑	51.1% ↑	51.0%↑
% Dead	0.0% ↓	2.7% ↓	16.0%	23.9% ↑	26.9% ↑	28.2% ↑	29.0% ↑	29.4% ↑	29.7% ↑	30.2% ↑	30.1%↑
Dead/Infected	0.63 ↑	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59
Epidemics at OH	0 ↓	0 ↓	17	36 ↑	54 ↑	54 ↑	71 ↑	54 ↑	57 ↑	70 ↑	73 ↑
Epidemics at GL	0 ↓	0 ↓	5	11	15 ↑	14 ↑	9	18 ↑	21 ↑	18 ↑	24 ↑
Epidemics at OH & GL	0	0	0	0	0	1	0	3 ↑	0	0	1

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

day after the flu is introduced, and low epidemic totals which account only for the single infectious individual that was introduced to the Norway House post. This confirms that, as expected, the epidemic was unable to spread when the disease probability was set at zero and indicates that this parameter is functioning as designed.

When disease probability was raised above its standard value of 0.2, a different trend emerged: the duration of infection and length of the epidemic dropped below the standard control range and grew progressively shorter as disease probability was increased; the epidemic peak occurred increasingly earlier; the peak number of infected continued to rise above the control range maximums; the epidemic totals continued to grow larger; and the incidence of the flu at Oxford House and God's Lake became increasingly common. This occurred because, as disease probability was increased, a larger proportion of the contacts between infectious agents and susceptible agents resulted in disease transmission. This allowed the flu to spread more quickly and to infect more agents, which resulted in higher epidemic totals and an accelerated epidemic. When the disease probability was set at 1.0, its maximum value, any contact between an infectious and susceptible agent would be expected to result in disease transfer. Surprisingly, however, even at this value, only 51% of the agents in the NHOHGL model are infected. This indicates the important role that contact rates play in the spread of the simulated epidemic. Even if all contacts result in disease, without any contacts, the flu can not spread.

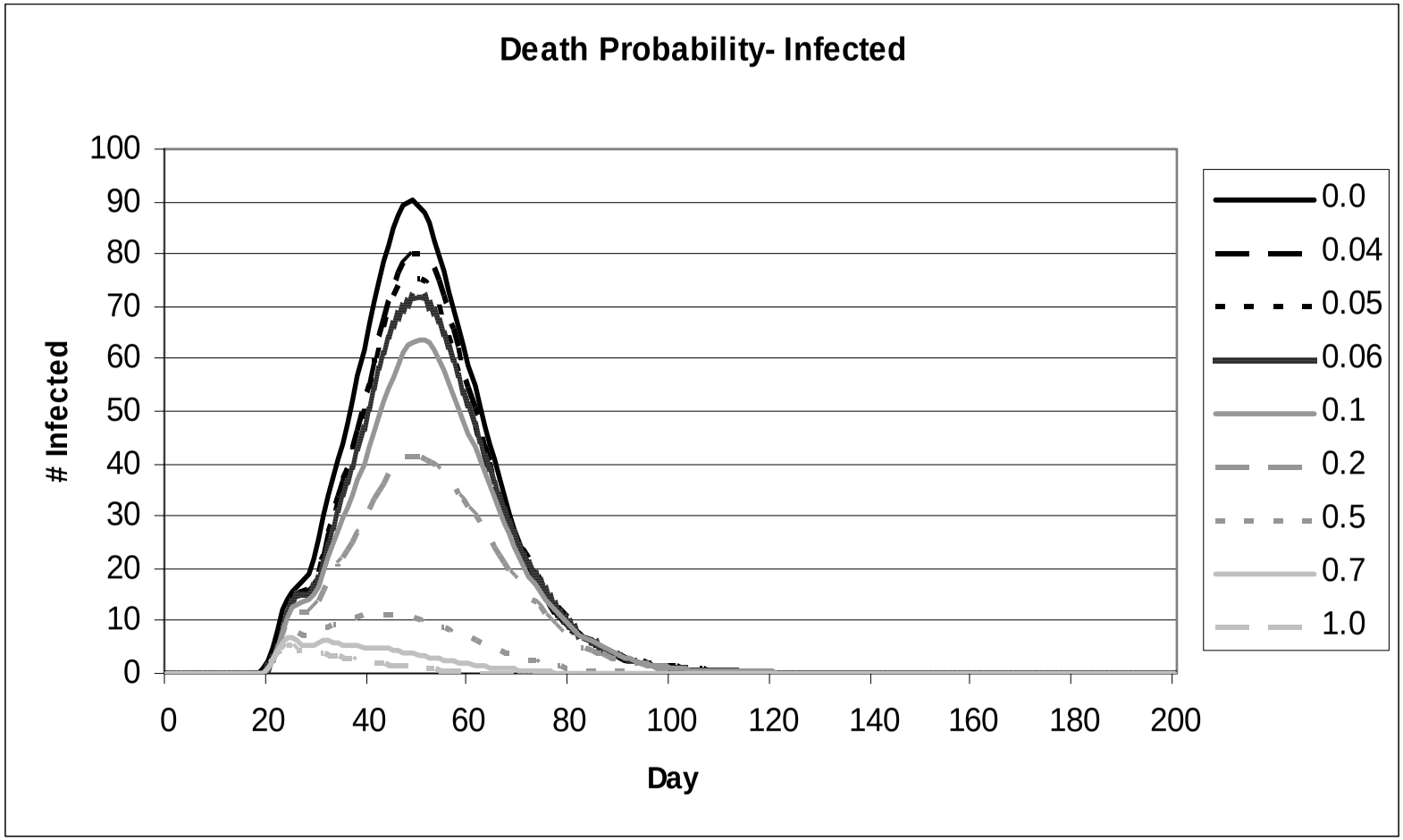
- **Death Probability**

The probability of death parameter specifies the probability that an infection will result in death rather than recovery. Death may occur on any day following the end of the death time period and before the end of the infectious period and essentially removes the dead agent from the active agent population. Death was a newer addition to the model and was first added in the Norway House Mortality model following the completion of Carpenter's thesis (Carpenter, 2004). Because the testing of this later version of the Norway House model has yet to be published, the probability of death parameter was thoroughly tested in this project. Probability of death was tested over a wide range of values in order to verify that it was functioning properly and to identify a value that would allow the model to reproduce historical mortality rates for the 1918-1919 flu epidemic at Norway House. The results of the sensitivity test may be found in Figure 6.16 and Table 6.15.

As would be expected, when the death probability was reduced from 0.2 to 0.1, the death toll dropped well below the control range minima because fewer infections were resulting in death. Conversely, the duration of infection and length of the epidemic increased, as did the peak number infected and the total number recovered and infected. This occurred because fewer of the infected agents were dying. The surviving agents contributed directly to the higher epidemic peak and number of infections, but because these agents were now remaining infected for the full length of the recovery period, their presence increased the probability of disease transmission, thereby adding further to these totals. When the death probability was reduced, there was also an increased incidence of

Figure 6.16: Results of the death probability sensitivity test

a. The number of infected agents per day of the simulation at different death probabilities



b. The number of dead agents per day of the simulation at different death probabilities

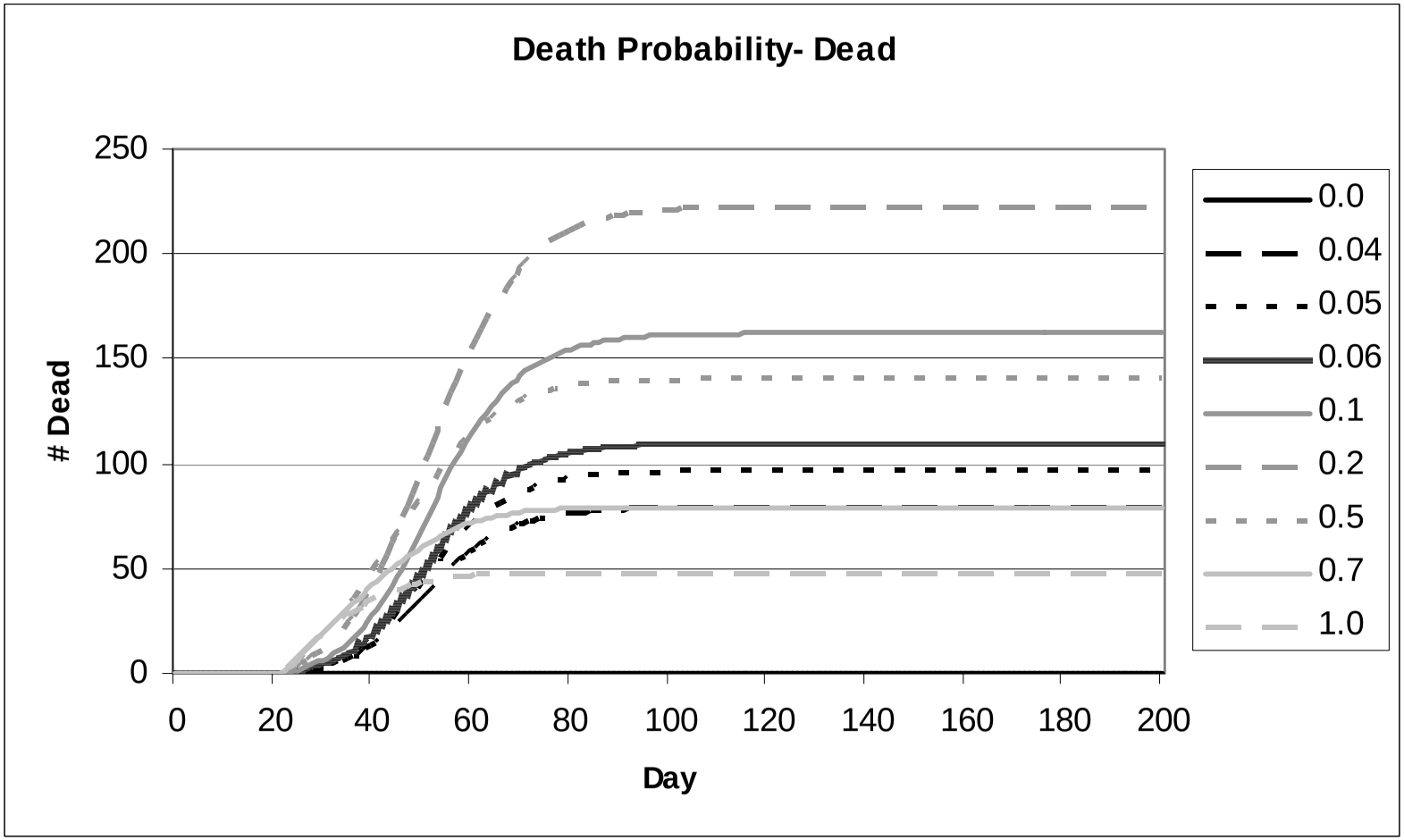


Table 6.15: Epidemic characteristics for death probability

Death Probability	0	0.04	0.05	0.06	0.1	0.2 *	0.5	0.7	1.0
Duration of Infection	107	130	125	115	148 ↑	123	107	98 ↓	62 ↓
Length of the Epidemic	0	127	120	112	138 ↑	122	106	97 ↓	61 ↓
Peak Time of Infection	30	30	30	30	31	30	20 ↓	5 ↓	5 ↓
Peak Number Infected	90.17 ↑	80.33 ↑	75.30 ↑	71.77 ↑	63.76 ↑	41.62	11.21 ↓	6.92 ↓	5.34 ↓
Total Recovered	565.21 ↑	447.22 ↑	422.13 ↑	392.48 ↑	310.25 ↑	153.41	9.27 ↓	0.66 ↓	0.00 ↓
Total Dead	0.00 ↓	79.31 ↓	96.85 ↓	109.64 ↓	162.37 ↓	221.19	140.49 ↓	78.91 ↓	47.69 ↓
Total Infected	565.21 ↑	526.52 ↑	518.97 ↑	502.12 ↑	472.62 ↑	374.60	149.75 ↓	79.57 ↓	47.69 ↓
% Infected	41.0% ↑	38.2% ↑	37.6% ↑	36.4% ↑	34.2% ↑	27.1%	10.9% ↓	5.8% ↓	3.5% ↓
% Dead	0.0% ↓	5.7% ↓	7.0% ↓	7.9% ↓	11.8% ↓	16.0%	10.2% ↓	5.7% ↓	3.5% ↓
Dead/Infected	0.00 ↓	0.15 ↓	0.19 ↓	0.22 ↓	0.34 ↓	0.59	0.94 ↑	0.99 ↑	1.00 ↑
Epidemics at OH	43 ↑	40 ↑	37 ↑	30 ↑	32 ↑	17	2 ↓	0 ↓	0 ↓
Epidemics at GL	19 ↑	13	8	8	8	5	0 ↓	0 ↓	0 ↓
Epidemics at OH & GL	0	0	0	0	0	0	0	0	0

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

the flu at Oxford House. Again, this was due to the larger number of infected agents, which gave the flu more opportunities to reach the other communities.

When death probability was decreased to 0.0, it would be expected that death would not occur, and this was true in the test simulation. Further, with no deaths, the duration of infection dropped to the lower end of the standard control range, the epidemic peak and the numbers recovered and infected exceeded the control range maximums, and there was a higher incidence of the flu at both Oxford House and God's Lake. This illustrates a continuation of the trends that were discussed in the previous paragraph and indicates that the death probability parameter is functioning properly and appropriately.

Increasing the probability of death produced a different pattern in the test data. As would be expected when death probability increased, a higher proportion of the infected agents died rather than recovered. However, other changes were also observed. When the death probability was increased from 0.2 to 0.5, the duration of infection and length of the epidemic both fell to the lower end of the standard control range, and dropped below the control range minima when the death probability was increased further. The time of peak infection also came increasingly earlier and the total number of infected dropped below the control range minima and continued to decrease as the probability of death increased. The total death toll also dropped and was reduced by nearly half each time the probability of death was increased. This occurred because the increasing death rate caused more and more infected agents to die before the end of the recovery period, thereby reducing the probability of disease transmission and resulting in ever lower epidemic totals. When the death probability was increased to its maximum, 1.0, this trend became even more pronounced. As would be expected, all of the infected

agents died and the epidemic statistics were all far below the control range minima. This occurred because with such a high death rate it was nearly impossible for the epidemic to sustain itself. Such a severe epidemic would quickly eliminate any of its carriers, providing it with very little chance to spread.

Historically, it is estimated that most of the population at Norway House was infected during the 1918-1919 flu epidemic and that approximately 15-20% died (Herring, 1994). Thus, it was important to find a death probability that could produce a similar mortality rate in the model simulations. The Norway House Mortality model had used a value of 0.2, the same value used for disease probability. Although this value would be expected to cause a 20% mortality rate, the NHOHGL standard parameter replication test established that a death probability of 0.2 actually causes a mortality rate of around 59%. This is because the probability of death does not directly specify the proportion of infections that will result in death, but rather determines the probability that an infectious agent will die on any day that he/she is infected. Because a death probability of 0.2 produced a mortality rate that was in excess of the historical estimates, other values were tested in order to find a more appropriate value. It was discovered that to replicate a death rate of between 15 and 20%, the death probability would have to be set at 0.04 (15% mortality), 0.05 (19% mortality), or 0.06 (22% mortality). Ultimately, a death probability of 0.04 was chosen for the winter and summer test simulations in order to provide a more accurate picture of the epidemic mortality. The results of these simulations will be discussed in chapter 7.

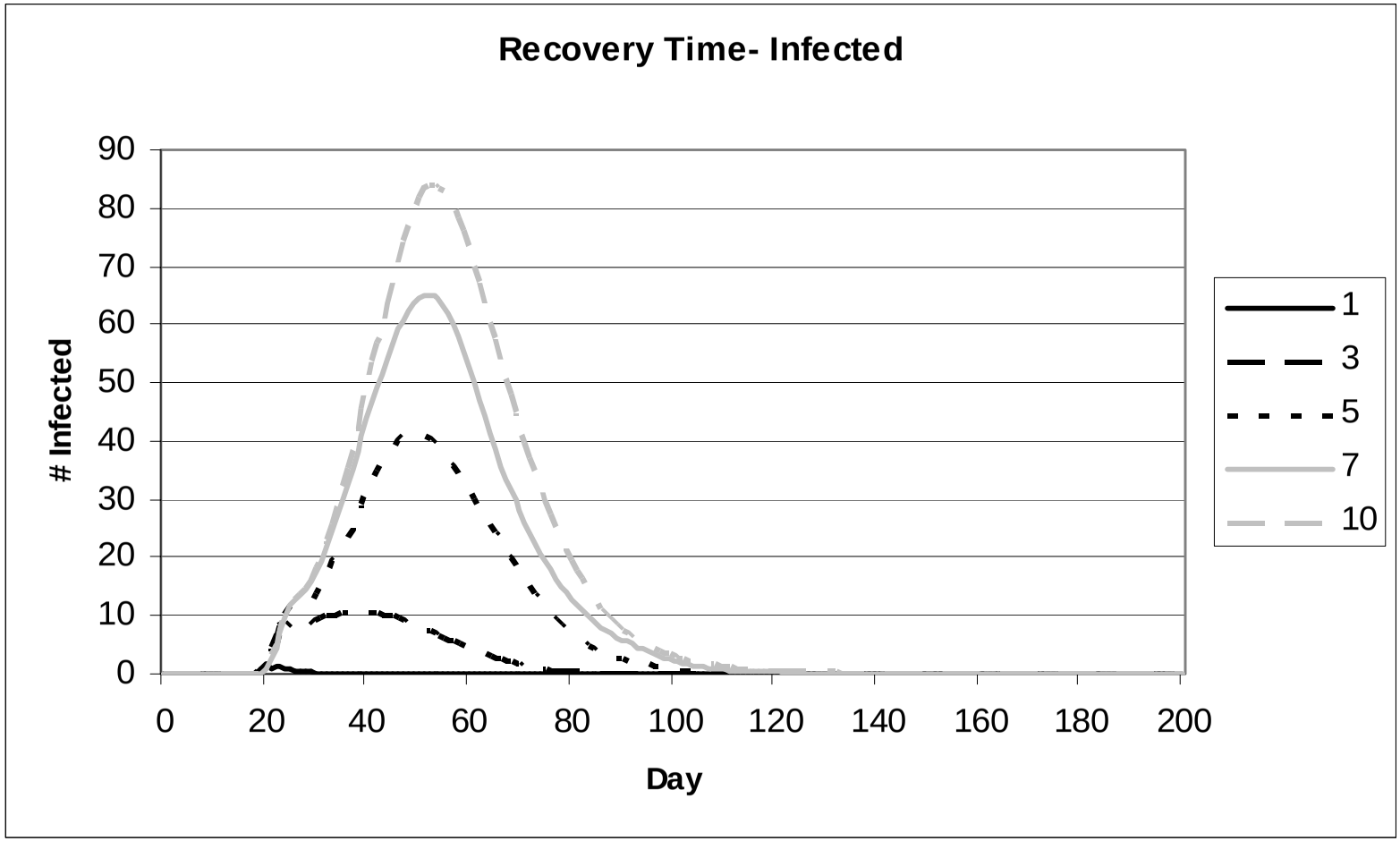
- **Recovery Time**

The recovery time parameter determines the length of the infectious period, or the period of time, in days, during which an infectious agent can transmit the flu to a susceptible agent. An infected agent remains infectious throughout the entire recovery period and, unless death occurs, becomes recovered following its completion. The length of the recovery period was set at 5 days for this project, which is the same value used in the original Norway House model and was based upon epidemiological data for influenza. For the replication test, several other values were also tested. The results of these simulations may be found in Figure 6.17 and Table 6.16.

When the length of the recovery period was reduced, the agents remained infectious for a shorter period of time. As would be expected, this led to shorter and milder epidemics. The duration of the epidemic and its associated mortality dropped to the lower end of the control range and then fell below its minimum values as the duration of infection was further decreased. The time of peak infection also came earlier as the length of the infectious period was reduced, and the height of the epidemic peak decreased dramatically. Additionally, the total number recovered, infected, and dead, and the number of epidemics at Oxford House and God's Lake all fell below the standard control range minima and decreased further along with the length of the recovery period. This occurred because, as the length of the infectious period was reduced, an infectious agent could transmit the flu to fewer and fewer susceptibles during this shrinking window of time. This made it more difficult for the flu to spread and also caused the mortality rates to plummet, since there were fewer days on which death could occur.

Figure 6.17: Results of the recovery time sensitivity test

a. The number of infected agents per day of the simulation at different values for the recovery time parameter



b. The number of dead agents per day of the simulation at different values for the recovery time parameter

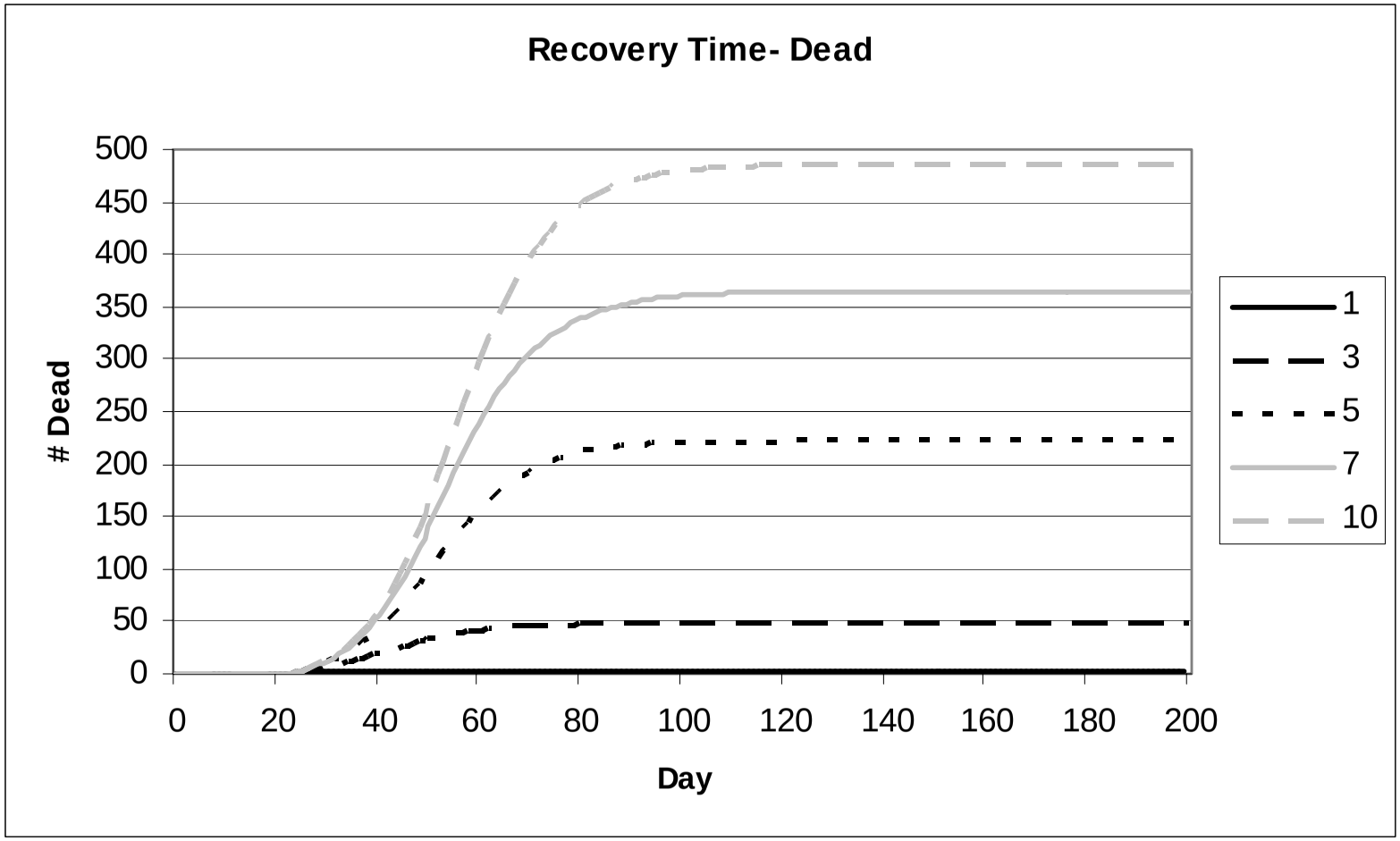


Table 6.16: Epidemic characteristics for recovery time

Recovery Time (days)	1	3	5 *	7	10
Duration of Infection	26 ↓	111	123	134 ↑	146 ↑
Length of the Epidemic	26 ↓	108	122	125	145 ↑
Peak Time of Infection	1 ↓	18 ↓	30	32	32
Peak Number Infected	1.56 ↓	10.76 ↓	41.62	65.15 ↑	83.94 ↑
Total Recovered	8.57 ↓	85.71 ↓	153.41	129.15 ↓	75.66 ↓
Total Dead	2.11 ↓	47.90 ↓	221.19	364.04 ↑	485.96 ↑
Total Infected	10.68 ↓	133.61 ↓	374.60	493.20 ↑	561.62 ↑
% Infected	0.8% ↓	9.7% ↓	27.1%	35.7% ↑	40.7% ↑
% Dead	0.2% ↓	3.5% ↓	16.0%	26.4% ↑	35.2% ↑
Dead/Infected	0.20 ↓	0.36 ↓	0.59	0.74 ↑	0.87 ↑
Epidemics at OH	0 ↓	4 ↓	17	36 ↑	57 ↑
Epidemics at GL	0 ↓	0 ↓	5	14 ↑	22 ↑
Epidemics at OH & GL	0	0	0	0	1

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

When the length of the recovery period was increased, the agents remained infectious for a longer period of time. This led to longer and more severe epidemics: the duration of infection and mortality increased; the time of peak infection occurred at the high end of the control range; the peak number infected rose above the standard control range values and grew higher as the length of the recovery period was increased; the epidemic totals all rose progressively higher; and the number of epidemics at Oxford House and God's Lake steadily increased. The longer infectious periods gave the infectious agents more opportunity to transmit the flu within their own communities and to carry it elsewhere. The number of infected individuals remained higher for a longer period of time as agents remained infectious for a longer period and as each infected more and more susceptibles. The death rate also increased because each day that an

agent remained infected, he/she had an additional chance to die, and with the extended infectious period, the number of these chances increased.

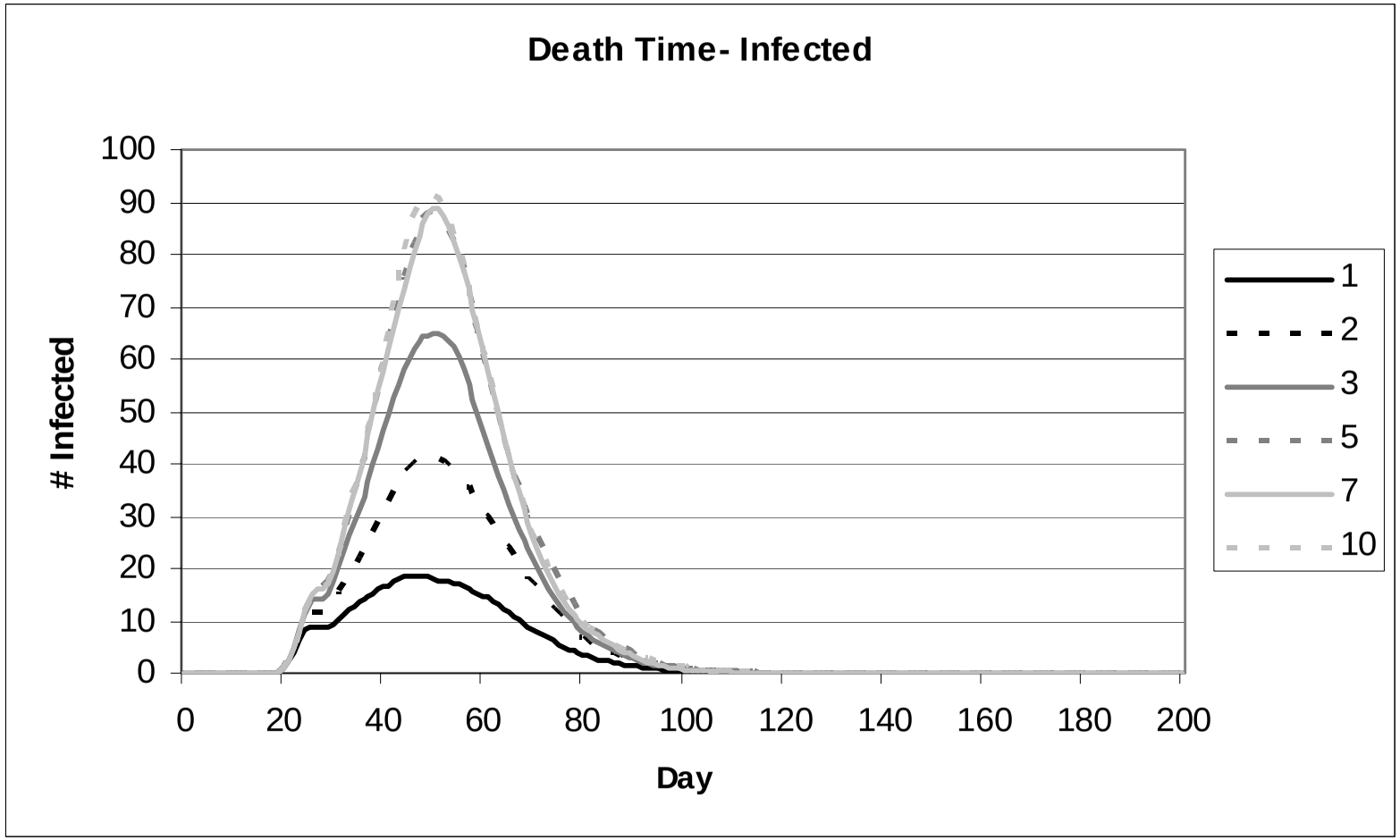
- **Death Time**

The death time parameter determines the average length of time, in days, that an agent is infected before death occurs. The standard value for this parameter in the NHOHGL model is 2 days, an estimate that was based upon data from the 1918-1919 flu pandemic (Crosby, 2003). This means that in the standard model scenario, death can occur between days 2 and 5 of the recovery period. If death does not occur within this allotted period, the agent automatically becomes recovered. Death time, like the probability of death, was added to the original Norway House following the completion of Carpenter's thesis (Carpenter, 2004) as part of the Norway House Mortality model revision. Because no tests of the death time parameter have yet been published, it was tested at a wide range of values in this project in order to determine whether it was performing appropriately and as designed. The results of this test may be found in Figure 6.18 and Table 6.17.

When death time was decreased from 2 days to 1 day, the duration of infection and length of the epidemic were relatively unaffected and remained well within the standard control range. The peak time of infection, however, dropped below the control range minima, as did the peak number infected, the total number recovered, dead, and infected, and the incidence of the flu at Oxford House and God's Lake. The shortened death time resulted in a milder epidemic because it allowed death to occur on an additional day of the infectious period, day 1. This increased the probability that death

Figure 6.18: Results of the death time sensitivity test

a. The number of infected agents per day of the simulation at different values for the death time parameter



b. The number of dead agents per day of the simulation at different values for the death time parameter

Table 6.17: Epidemic characteristics for death time

Death Time (days)	1	2 *	3	5	7	10
Duration of Infection	120	123	127	129	118	128
Length of the Epidemic	120	122	124	118	113	122
Peak Time of Infection	27 ↓	30	30	30	30	29
Peak Number Infected	18.62 ↓	41.62	64.88 ↑	88.13 ↑	88.85 ↑	91.62 ↑
Total Recovered	73.66 ↓	153.41	246.40 ↑	441.46 ↑	432.93 ↑	448.76 ↑
Total Dead	151.57 ↓	221.19	235.31 ↑	110.74 ↓	108.88 ↓	112.01 ↓
Total Infected	225.23 ↓	374.60	481.71 ↑	552.21 ↑	541.82 ↑	560.76 ↑
% Infected	16.3% ↓	27.1%	34.9% ↑	40.0% ↑	39.3% ↑	40.6% ↑
% Dead	11.0% ↓	16.0%	17.1% ↑	8.0% ↓	7.9% ↓	8.1% ↓
Dead/Infected	0.67 ↑	0.59	0.49 ↓	0.20 ↓	0.20 ↓	0.20 ↓
Epidemics at OH	7 ↓	17	31 ↑	37 ↑	42 ↑	29 ↑
Epidemics at GL	1 ↓	5	7	14 ↑	11	11
Epidemics at OH & GL	0	0	0	1	0	1

Epidemic data are combined totals for the three communities. The (*) indicates the value used in the standard parameter scenario and the statistics for this run are the mean values from the standard replication test. The (↑) indicates that the value is above the control range for that characteristic. The (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

would occur and, as would be expected, caused the mortality rate to increase to 67%. It also caused a number of infected agents to be removed from the simulation one day earlier than normal, thereby prematurely reducing the number of infected agents and forcing the simulated epidemic to have an earlier peak. Further, because dead agents are unable to spread the flu, increasing death time reduced the probability of disease transmission. This led to fewer overall infections and restricted the geographic spread of the simulated epidemic.

Increasing death time had a somewhat different effect on the simulated epidemic. Here too, the change in value appeared to have very little impact on the duration of infection and length of the epidemic; however, unlike the previous trial, when death time was increased, there also appeared to be no significant change in the timing of the

epidemic peak. Further, increasing death time caused the peak number infected and total number infected to increase rather than decrease and led to an increased incidence of the flu at Oxford House and God's Lake. This occurred because, when death time was increased, there was a smaller window during which death could occur. This left more infectious agents more time to transmit the flu to susceptibles and also decreased the probability that death would occur, as is evidenced by the lowered mortality rates. Together, this caused more agents to become infected and led to a larger epidemic peak.

Although these general trends were reproduced each time that death time was increased, some notable differences were observed between the data for the 3-day trial and the 5-, 7-, and 10-day trials. In the 3-day trial, the total number of deaths was higher than the control range maximums. In the 5-, 7-, and 10-day trials, however, the number of deaths were all below the standard control range. Further, the data for the 5-, 7-, and 10- day trials appears to be quite uniform, indicating that some threshold had been reached following the 3 day trial. The threshold was the length of the recovery period. When the death time was set at 3 days, this left a 2 day window during which infectious agents could die before the end of the 5-day recovery period. When the death time was set at 5, 7, or 10 days, however, there was only one day on which death could occur: day 5. This is because the model code specifies that death may occur only between the end of the death time and the end of the recovery time up to the final day of the recovery period. Thus, when the death time is greater or equal to the recovery time, death may only occur on this one day. Indeed, as would be expected in such a situation, 20% of those infected die, a value equal to the value of the daily death probability. This indicates that the death

time parameters, as well as the probability of death parameter, are functioning properly and as designed.

- **Turning Off the Displays**

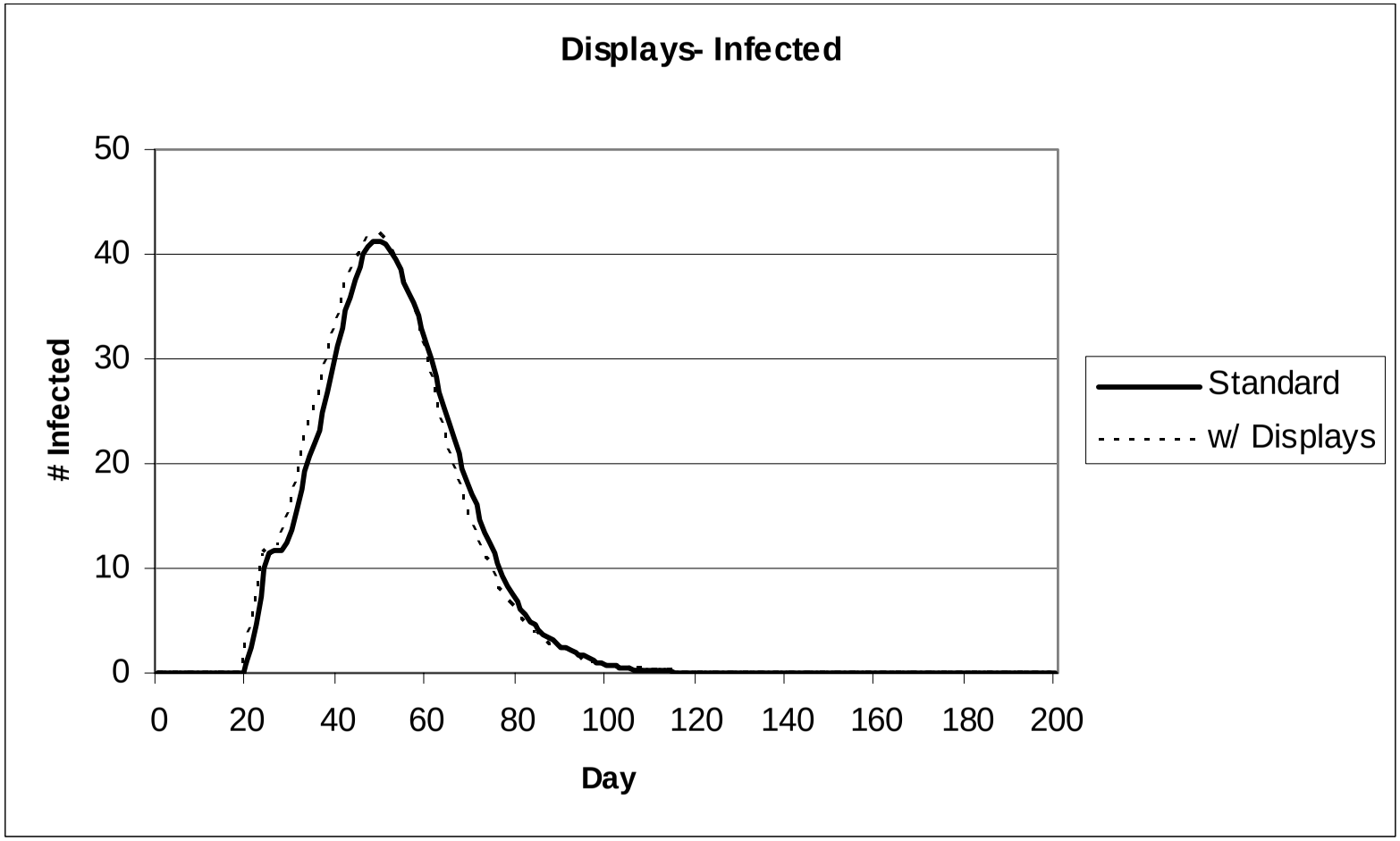
As noted in chapter 5, a new parameter was added to the NHOHGL model in order to allow the displays to be turned off while simulations were being run. Although the model displays provide for a more interesting and entertaining simulation experience, they require that the computer continuously update and display large amounts of data. This causes the simulation process to slow and lengthens run time. The ability to turn off the displays when they are not needed helps to increase the efficiency of data collection and makes run time more reasonable. The new displays on/off switch was tested in this project in order to ensure that turning off the model displays would not significantly affect the data that was generated by the NHOHGL model. The test compared the averaged data for two sets of 1000 runs of the model, one with the displays turned on and one with the displays turned off. The results of this test may be viewed in Figure 6.19 and Table 6.18.

Both the graphical and statistical comparison of the data for the two sets of runs reveals very little difference between them. The overall shape and height of the morbidity and mortality curves in Figure 6.19 is essentially identical, and the epidemic statistics for the two runs are also quite similar. When the displays were left on, all of the epidemic statistics fall within the limits of the standard control range with the exception of duration of infection, which was three days greater than the control range maximum. The reason for this is unclear, but it did not cause enough concern to merit the running of

Figure 6.19: Results of the displays on/off test

a. The number of infected agents per day of the simulation with the displays turned on/off

The data from the standard run were produced with the displays turned off.



b. The number of dead agents per day of the simulation with the displays turned on/off

The data from the standard run were produced with the displays turned off.

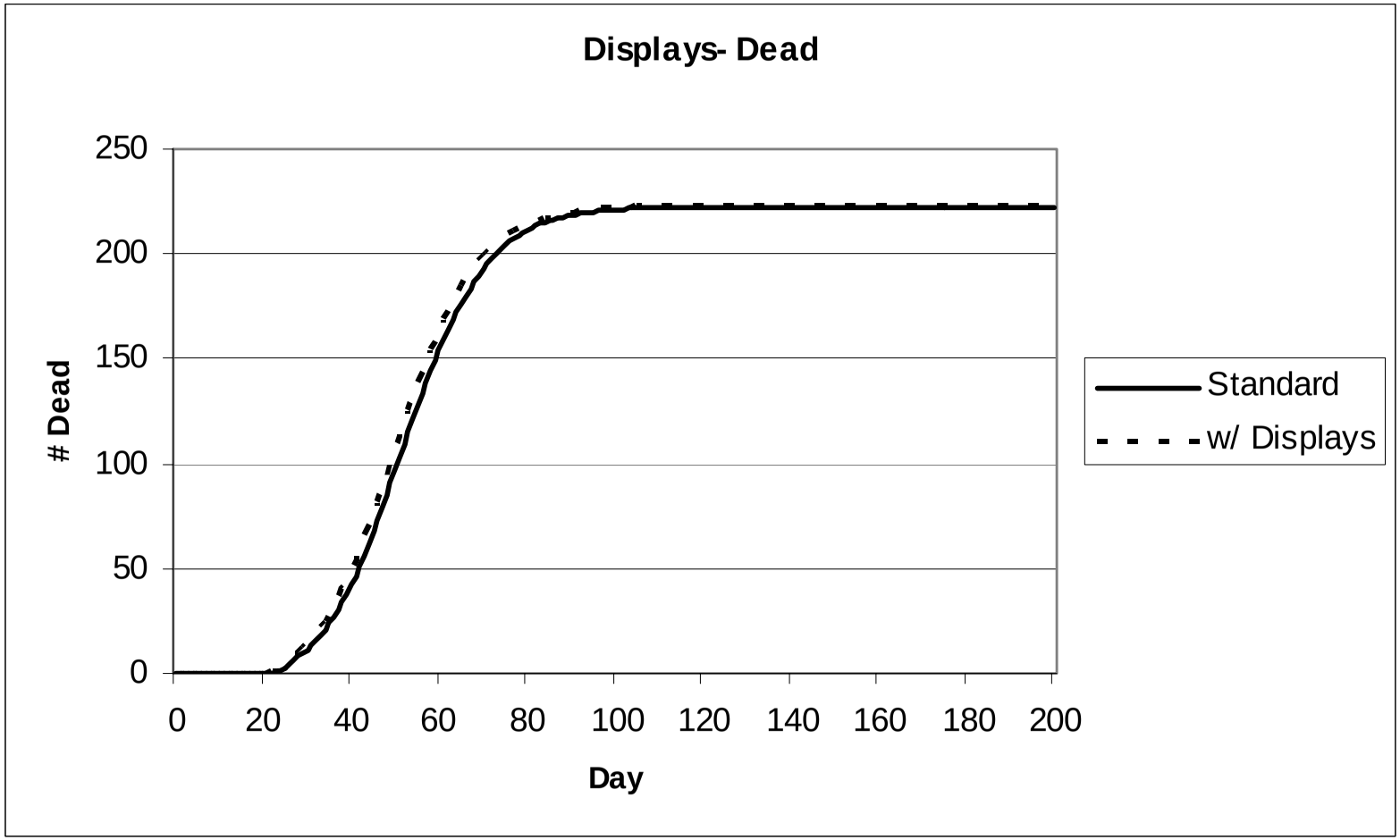


Table 6.18: Comparison of the data for the simulated epidemics with and without the onscreen displays turned on

Displays	Off	On
Time Needed for 1000 runs	5 hours	17 hours
Duration of Infection	123	136 ↑
Length of the Epidemic	122	131
Peak Time of Infection	30	29
Peak Number Infected	41.62	42.07
Total Number Recovered	153.41	154.14
Total Number Dead	221.19	223.56
Total Number Infected	374.60	377.70
% Infected	27.1%	27.4%
% Dead	16.0%	16.2%
Epidemics at OH	17	21
Epidemics at GL	5	6
Epidemics at OH and GL	0	0

In both tests, the standard parameter values were used. Epidemic data are combined totals for the three communities. Displays were turned off during the normal use of the NHOHGL model. The (↑) indicates that the value is above the control range for that characteristic and a (↓) indicates that the value is below the control range for that characteristic. The control range for the standard parameter scenario may be found in Table 6.1.

all model simulations with the displays left on, especially since it took 17 hours to run 1000 simulations with the displays turned on, as opposed to 5 hours with the displays turned off. This substantial time difference would have made running the model with the displays too inefficient and impractical. However, in order to eliminate any potential bias due to the possible influence of the displays on epidemic duration, all simulations in this project were run with the displays turned off.

Chapter 7: Investigation of the Research Questions and Assessment of the NHOHGL Model

Once the NHOHGL model parameters had been tested and were found to be performing both predictably and appropriately, the model could be used to investigate the research questions that were outlined for this project. These questions are: what impact might the traditional seasonal mobility and settlement patterns that were practiced by the Cree have had on the spread of the 1918-1919 flu epidemic in this region of Canada; why did the 1918-1919 flu epidemic never manage to reach Oxford House or God's Lake; and how do the results of the NHOHGL agent-based computer simulation compare to those of Sattenspiel's NHODE model, which used mathematical equations to model of the 1918-1919 flu epidemic in the same three communities (Sattenspiel and Dietz, 1995; Sattenspiel and Herring, 1998)? The NHOHGL model was used to investigate each of these questions. This chapter will present the simulation data that were collected in these investigations and will discuss what insights they may reveal into the research questions. The chapter will then conclude with an evaluation of whether the NHOHGL model is a reasonable representation of the 1918-1919 influenza epidemic in this region of Canada.

Question 1: What impact might the traditional Cree seasonal movement and settlement patterns have had on the spread of the 1918-1919 influenza epidemic in the Norway House region of Canada?

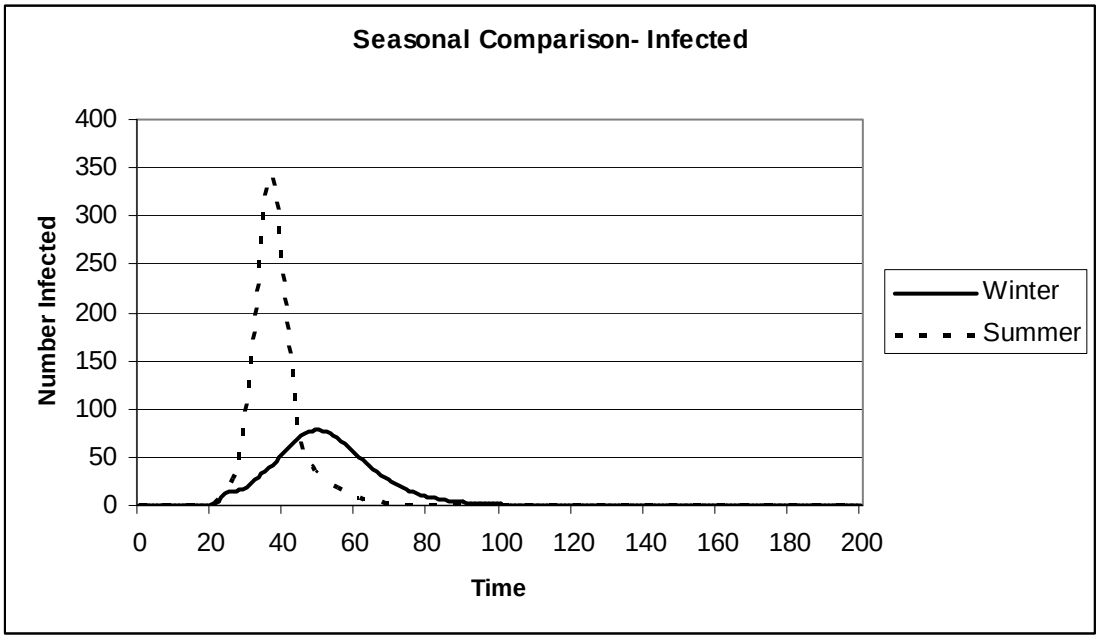
The 1918-1919 flu epidemic first reached Norway House in December of 1918, after most of the Cree family groups had already dispersed to their winter hunting camps

and traplines in the hinterland (Herring and Sattenspiel, 2003). Because this would have reduced the population density at the post, it has been hypothesized that the historical epidemic may have led to lower overall infection and mortality rates than would have resulted from a summer epidemic. A test of this hypothesis using the NHOHGL model required the use of the winter and summer parameter scenarios as outlined in Chapter 5. The control ranges for these model scenarios may be found in Table 6.2 (winter) and Table 6.3 (summer) and a graphical comparison of the averaged data for the winter and summer replication tests may be viewed in Figure 7.1.

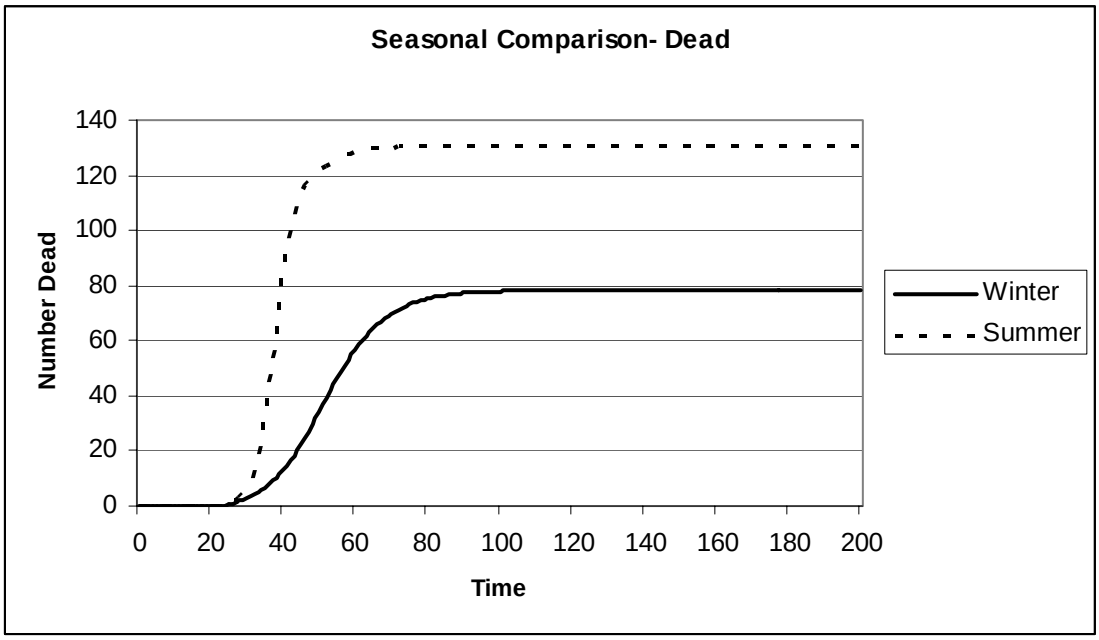
A comparison of the data for the winter and summer replication tests lends support to the above hypothesis. It is clear both from the graphs in Figure 7.1 and the control range tables in the previous chapter that there are significant differences between the data for the summer and winter model scenarios. First, the duration of infection and the length of the epidemic (the period between the first and last associated death) are consistently shorter for a summer epidemic than for a winter epidemic. In fact, the maximum length of a summer epidemic was still one day shorter than the minimum length of a winter epidemic. Second, a summer epidemic peaks nearly 20 days earlier than a winter epidemic and the size of the peak is more than four times greater for a summer epidemic. Third, a summer epidemic produces numbers of recovered, dead, and infected agents that are significantly and consistently higher than the totals for a winter epidemic. Finally, the flu manages to reach Oxford House much more frequently in a summer epidemic, an average of 374 times per 1000 runs versus an average of 13 times

Figure 7.1: Comparison of the data for the winter and summer replication tests. Graphs show the average number of infected agents per day of the simulation (a) and the average number of dead agents per day of the simulation (b) for each seasonal test.

a.



b.



per 1000 runs under the winter scenario. In both scenarios, however, the flu never manages to reach God's Lake.

These differences indicate that, as hypothesized, the simulated summer epidemics are significantly and consistently shorter and more severe than the simulated winter epidemics. This can be explained by the differences in both population density and travel rates between the two model scenarios. In the summer scenario, all of the family groups are located at the three community posts. This allows agents from different families to more frequently come into contact with one another, thus facilitating the spread of the flu. Additionally, travel is faster and more frequent in the summer scenario, allowing the flu a greater opportunity to reach Oxford House from Norway House. In the winter scenario, in contrast, the family groups at each community are divided evenly among the post and the four outlying camps. This greater dispersal of the population across the model landscape reduces contact among agents from different family groups, making it more difficult for the flu to spread within a community. Further, the reduced rate and speed of travel in the winter scenario decreases the probability that the flu will successfully reach Oxford House.

These seasonal factors, which have a clear influence on the NHOHGL model simulations, were based upon data regarding the conditions at Norway House, Oxford House, and God's Lake during the time of the 1918-1919 flu pandemic. It can thus be predicted that if the 1918-1919 flu had reached this region of Canada during the summer rather than the winter, the outcome may well have been much more devastating.

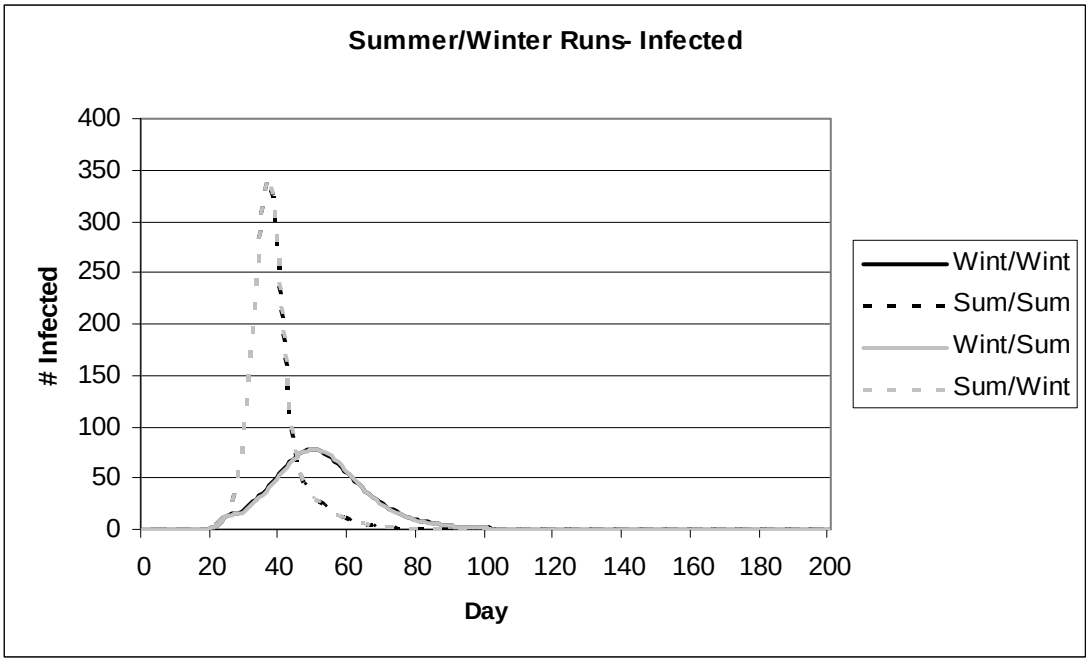
- **Corollary to Question 1: Do mobility rates or settlement patterns play a larger role in the observed differences between the winter and summer simulations?**

Although the data from the winter and summer replication tests indicate that seasonal differences in settlement patterns and in the speed and frequency of travel led to consistently shorter, more severe, and more geographically widespread epidemics in the summer than in the winter, these tests provide no information on the role that each of these seasonal factors plays in producing the observed seasonal epidemic patterns. For this reason, additional simulations were run in order to determine whether the seasonal differences in mobility rates or the seasonal differences in settlement patterns may have more strongly influenced the observed differences between the winter and summer epidemic patterns. These simulations were run using hybridized winter and summer parameter sets which combined the settlement structure from one season with the historical mobility data from the other season. The results of these simulations may be viewed in Figure 7.2 and Table 7.1 and are listed alongside the data from the winter and summer replication tests in order to facilitate comparison.

The data in Figure 7.2 and Table 7.1 clearly indicate that the epidemic data are almost entirely determined by the settlement structure rather than by the mobility structure. The data for the standard winter runs are identical to those for the run using winter settlement structure and summer mobility and the data for the standard summer runs are the same as those for the run using summer settlement structure and winter mobility. All of the epidemic statistics fall within the control ranges for the standard seasonal run sharing their population structure except for the peak number infected in the

Figure 7.2: Results of the winter/summer hybrid simulations shown alongside the averaged data for the winter and summer replication tests. Graphs show the number of infected agents per day of the simulation (a) and the number of dead agents per day of the simulation (b). In the keys for the graphs, "winter" and "summer" have been abbreviated. The first designation refers to the settlement structure and the second refers to the mobility scenario that was used in the simulation.

a.



b.

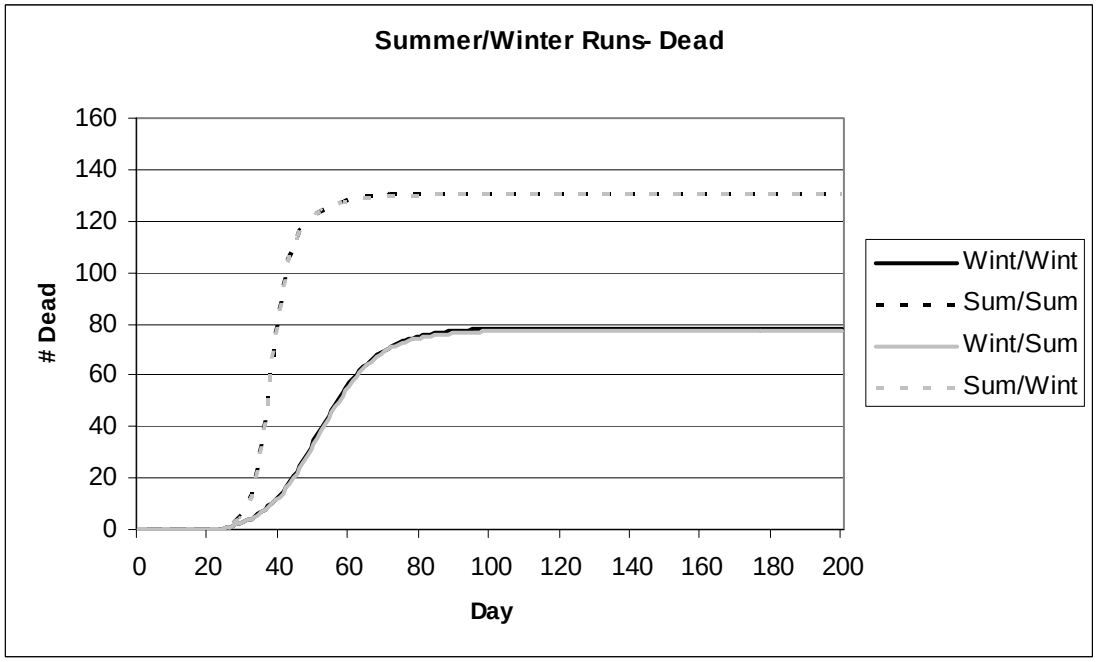


Table 7.1: Epidemic data for the winter/summer hybrid simulations
The averages for the winter and summer control range tests are also included to facilitate comparison.

Settlement Structure	Winter	Summer	Winter	Summer
Between Fort Move Probability	Historical Winter	Historical Summer	Historical Summer	Historical Winter
Duration of Infection	127	86	143	79
Length of the Epidemic	120	82	136	73
Peak Time of Infection	49	37	50	37
Peak Number Infected	78.06	342.62	77.64	344.39 ↑
Total Number Recovered	442.08	734.68	437.30	736.93
Total Number Dead	78.33	130.55	77.32	130.32
Total Number Infected	520.40	865.23	514.63	867.24
% Infected	37.7%	62.7%	37.3%	62.8%
% Dead	5.7%	9.5%	5.6%	9.4%
Epidemics at OH	13	374	11	371
Epidemics at GL	0	0	0	0
Epidemics at both OH and GL	0	0	0	0

Epidemic data from the replication study standard winter and summer scenarios is averaged data from 10 sets of 1000 runs of the NHOHGL (probability of disease 0.2; probability of death 0.04). Data for the hybrid runs is averaged data for 1000 runs of the NHOHGL model (probability of disease 0.2; probability of death 0.04). The values from the hybrid runs is compared to the control ranges for the listed seasonal settlement structure (not the move probability), (↑) indicates a value is above the standard control range for the appropriate control range (↓) indicates values that are below the appropriate control range.

summer structure/winter mobility run. Because this value is just slightly above the summer control range maximum, and because winter mobility pattern would be expected to lower the size of the epidemic peak rather than increase it, this is more likely an effect of stochasticity rather than an effect of the parameter change in this run.

Most of the epidemic data would be expected to be the same for all the runs with winter settlement structure and all the runs with summer settlement structure. The duration of infection, length of the epidemic, peak time of infection, peak number infected, and totals of the number recovered, dead, and infected would be expected to be more strongly influenced by contact within a community rather than contact between

communities. The exception, however, is the number of instances of the epidemic at Oxford House (God's lake was flu-free in all of the seasonal simulations). Surprisingly, this also appears to be more a product of the settlement structure than the inter-community mobility patterns. This is probably because the size and density of the agent population at the post has a strong influence on the number of inter-community travelers, as well as the likelihood that one of these travelers would be infected. Having more agents at the post increased the pool of available travelers, making intercommunity travel more likely, even as the between-post-move probabilities were decreased to their winter values. It is thus clear that the settlement structure is dominant factor controlling the spread, as well as the severity, of the seasonal simulated epidemics.

- **Corollary to Question 1: How do the findings of the NHOHGL model regarding the impact of seasonality compare to the data from the original Norway House model?**

The NHOHGL model was not the first agent-based computer simulation to investigate the impact of the traditional Cree seasonal movement and settlement patterns on the spread of the 1918-1919 flu epidemic. This question was also of central importance in the original Norway House model project (Carpenter, 2004). Because the Norway House model served as a foundation for the new NHOHGL model, it was important that the data from the two models regarding seasonality be compared in order to verify that the integrity of the original one-post model was not compromised when it was extended into a three-post model or when death was added to the simulations. This

required that simulations of the NHOHGL model be run using the same seasonal parameters that were used in the original Norway House model project and with death and inter-community mobility turned off. The resulting data for Norway House could then be compared to the winter and summer control ranges from the original Norway House model in order to determine whether the results were consistent.

The winter and summer control ranges for the Norway House model winter and summer epidemics may be found in Table 7.2 while the data for the NHOHGL model Norway House simulations may be viewed in Table 7.3. All of the data produced by the NHOHGL model in this test simulation clearly fall within the appropriate Norway House model seasonal control range with one exception: duration of infection. The NHOHGL model data indicate an average winter duration that is a full 17 days longer than the Norway House model control range maximum and an average summer duration that is 5 days longer. The reason for this discrepancy is unclear but it may indicate that the changes that were made to the original Norway House model, most likely the addition of inter-community mobility in this project or the addition of death following the completion of Carpenter's thesis, may have had an unintended effect upon the epidemic simulations. Whatever the cause, it is important to note this difference between the data from the NHOHGL model and the data from the original Norway House model, especially if it may indicate an overrepresentation of the duration of infection in the NHOHGL model simulations. This issue will be discussed further in the analysis of question number 4.

Table 7.2: Winter (a) and summer (b) control range data for the original Norway House model. Source: (Carpenter, 2004). The original Norway House Model included only the Norway House community (population 750) and did not model death or inter-community mobility.

a.

Winter	Maximum	Minimum	Std. Dev.
Duration of Infection	96	80	3.83
Peak Time of Infection	31	28	0.96
Peak Number Infected	98	85	3.07
Total Number Infected	574	530	11.44
Percent Infected	76%	71%	

b.

Summer	Maximum	Minimum	Std. Dev.
Duration of Infection	41	32	2.01
Peak Time of Infection	17	16	0.22
Peak Number Infected	369	356	3.17
Total Number Infected	749	746	1.16
Percent Infected	100%	99%	

Table 7.3: Epidemic data for the Norway House tests of the NHOHGL model
In these tests, the parameters were set equal to the values used in the original Norway House project, death was turned off, movement between the posts was halted, and the epidemic statistics were calculated only for the population at Norway House. The listed values are averages from 10 sets of 1000 runs of the NHOHGL model and would be expected to fall within the control ranges listed in Table 7.2. Values falling above the appropriate Norway House Model Control range are indicated by a (↑). No values fell below these control ranges.

Scenario	Winter	Summer
Duration of Infection (Days)	113 ↑	46 ↑
Peak Time of Infection (Day)	30	16
Peak Number Infected	88.36	359.13
NH Infected	545.16	746.76
NH % Infected	72.7%	99.6%

Question 2: Why did the 1918-1919 influenza epidemic never reach Oxford House or God's Lake?

This is one of the most interesting mysteries surrounding the 1918-1919 flu epidemic in this region of Canada. While the HBC post journals and the Anglican Church of Canada burial records indicate that the epidemic ravaged Norway House, killing approximately 15-20% of the population in just six weeks time, there is no record of the epidemic having visited either Oxford House or God's Lake (Herring, 1994; Herring and Sattenspiel, 2003; Sattenspiel and Herring, 2002). Given the proximity of the three communities, the absence of the flu at Oxford House and God's Lake is somewhat surprising, especially because contact between the three posts continued throughout the epidemic (Herring, 1994; Herring and Sattenspiel, 2003).

The reason for the absence of the flu at Oxford House and God's lake is unclear, although several theories have been suggested. One possible explanation is that it was simply luck that kept the 1918-1919 flu localized at Norway House. It is also possible that travel between the posts was too infrequent during the winter to have allowed the flu to easily spread between them. It could also be that the travel rates were sufficient but that the length of the journey between Norway House and the other communities was longer than the duration of the infectious period, making it impossible for an infectious traveler to still be contagious upon arrival at Oxford House or God's Lake. Another possibility is that the infectious trippers who might have potentially carried the flu from Norway House to Oxford House or God's Lake either died somewhere along the journey

or were forced to stop long enough that they would no longer have been infectious when they did eventually arrive (Sattenspiel and Herring, 2002).

While these theories appear to be plausible, a previous investigation of this mystery using the population-based NHODE model could determine no natural reason for the absence of the flu at Oxford House and God's Lake (Sattenspiel and Herring, 1998). In contrast to the historical data, the model predicted that, under normal winter conditions, there should have been cases of the flu at both of these communities. This discordance led Herring and Sattenspiel to suggest that perhaps the actions of local leaders, including quarantine efforts, may have reduced travel enough to have kept the historical flu from spreading to Oxford House and God's Lake (Herring and Sattenspiel, 2003; Sattenspiel and Herring, 1998). However, because their later studies have indicated that these quarantine efforts were probably ineffective (Sattenspiel and Herring, 2003) and because the historical data indicate that other nearby posts with no quarantine also did not experience the epidemic (Herring and Sattenspiel, 2003), this is probably not the most likely explanation.

The data from the NHOHGL model appear to support the theory that it was more than luck or the actions of local leaders that kept the 1918-1919 flu epidemic from reaching Oxford House and God's Lake. Under the winter parameters, which have been designed to mimic as closely as possible the conditions during the time of the 1918-1919 flu epidemic at Norway House, the model shows that the flu only rarely makes it to Oxford House (in an average of 1.3% of the runs) and never reaches God's Lake. This would appear to indicate that it would have been extremely unlikely for the flu to have

reached these communities even in the absence of quarantine measures. The reason is probably some combination of the distance between the posts, which is admittedly a rough estimate, the short infectious period of influenza, and the infrequency of travel. Together, these factors would have reduced the likelihood of an infectious traveler carrying the flu from Norway House to Oxford House or God's Lake to nearly zero.

In the summer simulations, the epidemic fared a bit better, managing to travel between communities more frequently than in the winter simulations. In the 10,000 runs of the model that were performed during the summer replication test, the flu reached Oxford House in an average of 37.4% of the runs. Here too, however, the flu never once reached God's Lake. Thus, although there appears to be a greater probability that a summer influenza epidemic would have spread to Oxford House than a winter influenza epidemic (like the 1918-1919 flu), it would still not be surprising if a summer epidemic at Norway House failed to spread to the rest of the region.

- **Corollary to Question 2: What would it take for the simulated epidemic to routinely reach Oxford House and God's Lake?**

As noted above, the data from the seasonal replication tests indicate that the flu was unlikely to have reached Oxford House or God's Lake in the wintertime and that it would not have been unusual for a summer epidemic to also fail to spread to these additional communities. This observation provokes an interesting question: what would it take to make the simulated epidemic spread to Oxford House and God's Lake on a regular basis? This question has important implications for the study of modern flu

epidemics because today, travel is faster and more frequent and contact rates are often higher than they were historically. This means that a modern influenza epidemic could have the potential to spread more widely and more rapidly than the 1918-1919 flu epidemic. Being able to determine which factors would have encouraged the geographic spread of the 1918-1919 flu could thus be useful in predicting and preventing the spread of the next major flu pandemic.

To investigate this question, several tests of the NHOHGL model were performed in which one or more of the standard seasonal parameter values were increased. These parameters were selected because the sensitivity testing had indicated that increasing their respective values led to more widespread epidemics. For each test, 1000 runs of the NHOHGL model were performed using the revised input parameters. The changes that were made in each test, along with the resulting epidemic data, may be viewed in Table 7.4 (winter) and Table 7.5 (summer).

In the winter severity tests, merely increasing steps-per-day to 12 or increasing the probability of contact between members of different family groups to 0.1 did manage to slightly increase the incidence of the flu at Oxford House, but failed to produce any epidemics at God's Lake. To consistently produce an epidemic that reached both additional communities, multiple parameter increases were necessary. Increasing the disease probability to 0.5, the probability of contact outside a family to 0.1, and steps-per-day to 6, while multiplying the between post move probabilities by 100, led to an increased incidence of the flu at both locations, with the flu reaching Oxford House in just under half the runs and reaching God's Lake in nearly 5%. Retaining these same

Table 7.4: Epidemic data for the winter severity tests

Increases made:	Run A	Run B	Run C	Run D	Run E
Disease Probability			0.5	0.5	0.5
Outside Contact		0.1	0.1	0.1	0.1
Steps-Per-Day	12		6	12	12
Between Post Move Prob.			100x	100x	100x
Recovery Time					10
Death Time					4
Duration of Infection	150	50	45	43	60
Length of the Epidemic	141	47	44	42	55
Peak time of Infection	30	12	10	10	11
Peak Number Infected	76.66	350.16	435.81	455.10	710.68
Total Number Recovered	439.11	622.92	733.62	770.72	794.50
Total Number Dead	78.22	110.46	130.59	137.20	262.07
Total Number Infected	517.32	733.38	864.20	907.91	1056.58
% Infected	37.49%	53.14%	62.62%	65.79%	76.56%
% Dead	5.67%	8.00%	9.46%	9.94%	18.99%
Epidemics at OH	57	28	454	665	883
Epidemics at GL	0	0	45	58	173
Epidemics at OH & GL	0	0	45	58	173

Epidemic data are the combined totals for the three communities.

Table 7.5: Epidemic data for the summer severity tests

Increases made:	Run A	Run B	Run C	Run D	Run E
Disease Probability			0.5	0.5	0.5
Outside Contact		0.1	0.1	0.1	0.1
Steps-Per-Day	12			12	12
Between Post Move Prob.			100x	100x	100x
Recovery Time					10
Death Time					4
Duration of Infection	86	25	51	23	31
Length of the Epidemic	75	23	42	18	22
Peak time of Infection	17	4	5	5	6
Peak Number Infected	352.87	736.77	948.39	1267.69	1339.09
Total Number Recovered	788.10	752.30	1166.36	1170.96	1037.66
Total Number Dead	139.89	133.60	207.24	207.59	342.34
Total Number Infected	927.99	885.90	1373.60	1378.55	1380.00
% Infected	67.25%	64.20%	99.54%	99.89%	100.00%
% Dead	10.14%	9.68%	15.02%	15.04%	24.81%
Epidemics at OH	563	412	999	999	1000
Epidemics at GL	0	0	992	999	1000
Epidemics at OH & GL	0	0	992	999	1000

Epidemic data are the combined totals for the three communities.

changes but increasing steps-per-day a bit further to 12 caused these totals to increase even further, with the flu reaching Oxford House in 66.5% of the runs and the flu reaching God's Lake in just under 6%. Doubling both the recovery time and the death time allowed the flu to finally reach Oxford House in 88.3% of the runs, but still left the incidence of the flu at Gods' lake at only 17.3%.

The alterations to the summer scenario were more successful in furthering the geographic spread of the simulated epidemic. As in the winter severity tests, merely increasing steps-per-day to 12 or increasing the probability of contact outside of a family group to 0.1 caused the flu to reach Oxford House more often, but did not increase the incidence of the flu at God's Lake. However, increasing disease probability to 0.5, upping the probability of contact outside a family to 0.1, and multiplying the between post move probabilities by 100 caused the flu to reach both additional communities in nearly all of the 1000 runs. This also remained the result when additional modifications, identical to those made for the winter severity tests, were performed.

The results of these severity tests indicate that it takes multiple modifications of the standard seasonal parameter values in order to produce an epidemic that routinely spreads to all three communities. Yet, although the modifications made to the parameters in these severity tests may appear to be extreme, they are not unrealistic by modern standards. The increased travel rates used in these simulations allow the agents to travel between Norway House and Oxford House in 2 days; between Oxford House and God's Lake in 1 day; and between God's Lake and Norway House in 3 days. Such travel rates may still underestimate the speed of modern travel by airplane, car, railroad, or

snowmobile. Further, modern travel has reduced the isolation of these communities and increased their contact with one another and with other communities. This, along with the increased chances for social interaction that are provided by modern schools, workplaces, shopping and leisure activities could well account for the increased contact rates that were used in these simulations. As for the epidemic parameters, these were only estimates in the first place, and even if the values used in the severity tests overestimate the virulence and infectious period of the 1918-1919 flu epidemic, it is impossible to tell what a future epidemic holds in store.

The severity tests also indicate that it takes less modification for the simulated flu epidemic to consistently spread to Oxford House and God's Lake under the summer scenario, with its aggregated population, than under the winter scenario, with its dispersed population. Such a result would be expected based upon the results of the seasonal replication tests, but is of particular importance because the data from the summer tests may well be a more accurate representation of the potential impact of a modern-day flu epidemic in this region. This is because most of the residents of former fur-trading communities like Norway House, Oxford House, and God's Lake now live near the central "post" area year-round (Flannery, 1995; Hallowell, 1992; Hanks, 1982). The results of these tests thus provide a message of caution, indicating that even though Oxford House and God's Lake may have been safe from the 1918-1919 flu epidemic, they might not be so lucky today.

- **Corollary to Question 2: Would the flu epidemic have been more or less likely to spread if it was introduced at Oxford House or God's Lake instead of at Norway House?**

Because Norway House occupied a more prominent position along the trade routes than did Oxford House or God's Lake, it is not surprising that the 1918-1919 flu epidemic first appeared at this community. As shown above, the results of the NHOHGL model simulation indicate that the failure of the 1918-1919 flu epidemic to spread to Oxford House or God's Lake should also not be surprising. However, what would have happened if the epidemic had first appeared at Oxford House or God's Lake rather than at Norway House? Would the epidemic have been more or less likely to spread? This question was previously considered by Sattenspiel and Herring (Sattenspiel and Herring, 2002) who, using the population-based NHODE model, concluded that, because the travel rates into Norway House were higher than the travel rates from Norway House, the 1918-1919 flu epidemic would have been more likely to have spread if it had been introduced at Oxford House or God's Lake. This is assuming, however, that the epidemic could have reached one of these communities without first going through Norway House, which would probably have been unlikely.

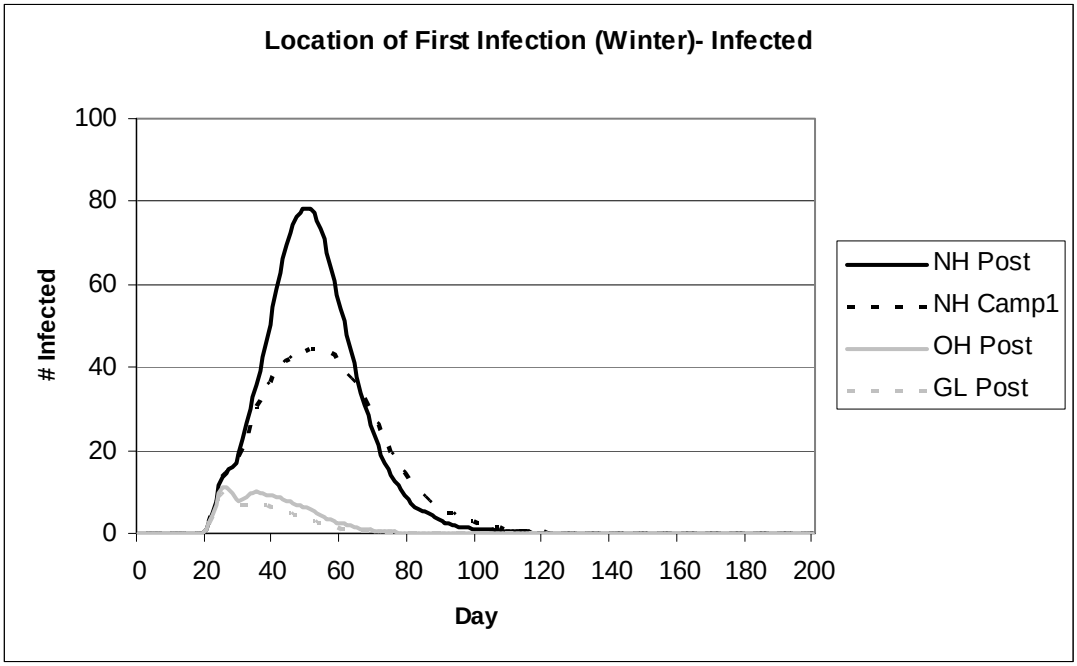
At first glance, the data from the NHOHGL model would appear to contradict the prediction made by Sattenspiel and Herring. The data for the "Location of First Infection" parameter test, as described in chapter 6, indicate that if the flu were introduced at Oxford House or God's Lake rather than at Norway House, the regional epidemic would have been shorter and milder and the flu would inevitably have failed to

spread to the other communities. However, it is important to note that in the NHOHGL parameter tests, the between-post move probabilities were all set at 0.001 whereas Sattenspiel's model used historical winter travel rates that were more similar to the values used in the NHOHGL winter simulations (see Table 5.3). For this reason, additional NHOHGL model simulations had to be run in which the location of first infection was altered in the winter scenario. The results of these tests are provided in Figure 7.3 and Table 7.6.

The data from these tests appear to support the findings of the initial NHOHGL model parameter test as opposed to the findings of Sattenspiel's NHODE model. Even despite the inclusion of the historical between-post move probabilities, the simulated epidemics continue to be shorter, milder, and isolated at the site of introduction when the flu is introduced at Oxford House or God's Lake instead of at Norway House. This result is really not surprising given the fact that the historical between-post move probabilities are all considerably lower than the value used in the original parameter test. Nonetheless, the discordance between these results and the results of Sattenspiel's NHODE model deserve further exploration. It is possible that the reason for the difference is simply due to differences in the values of the other input parameters. However, it is also possible that the difference is related to the more fundamental difference between the two modeling techniques. These possibilities will be explored further in the following section.

Figure 7.3: Results of the winter location of introduction tests showing the number of infected agents per day of the simulation (a) and the number of dead agents per day of the simulation (b). Historical between-post move probabilities were used in these tests.

a.



b.

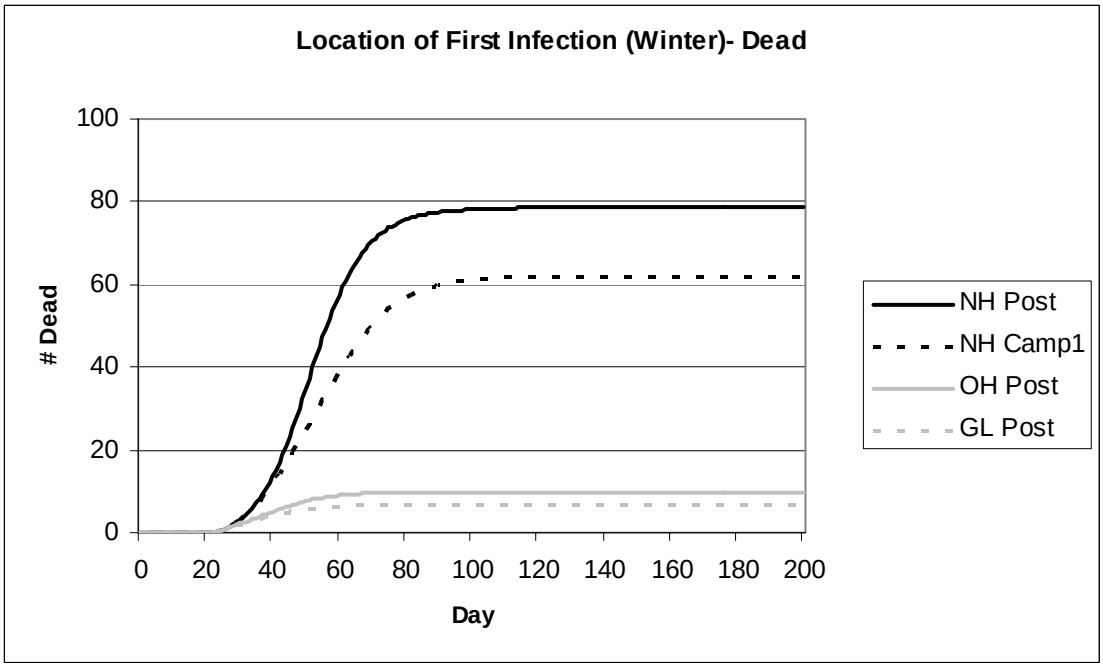


Table 7.6: Epidemic data for the winter location of first infection tests

Location 1st Infection	NH	OH	GL
Duration of Infection	127	87 ↓	72 ↓
Length of the Epidemic	120	82 ↓	65 ↓
Peak Time of Infection	29	6 ↓	6 ↓
Peak Number Infected	78.06	11.19 ↓	9.56 ↓
Total Recovered	442.08	54.86 ↓	37.60 ↓
Total Dead	78.33	9.73 ↓	6.65 ↓
Total Infected	520.40	64.58 ↓	44.26 ↓
% Infected	37.7%	4.7% ↓	3.2% ↓
% Dead	5.68%	0.7% ↓	0.5% ↓
Epidemics at NH	1000	0 ↓	0 ↓
Epidemics at OH	13	1000 ↑	0 ↓
Epidemics at GL	0	0 ↓	1000 ↑
Epidemics at OH & GL	0	0	0
Epidemics at NH & GL	0	0 ↓	0 ↓
Epidemics at NH & OH	13	0 ↓	0 ↓

Epidemic data are the combined totals for the three communities. In each test, the first infectious individual is introduced at the community post. Data for Norway House are the averages from the winter replication test. Values higher than the winter control range maximums are indicated by a (↑); values lower than the winter control range are indicated by a (↓).

Question 3: How do the results of the agent-based simulations compare to those of Sattenspiel's mathematical model of the 1918-1919 flu epidemic?

As noted in chapter 5, the primary objective of the NHOHGL model project was to expand upon the basic design of the original Norway House model in order to produce a stochastic, agent-based computer simulation of the 1918-1919 influenza epidemic that included not only Norway House, but also Oxford House and God's Lake. One major advantage of this new model over its predecessor was that the results of the model simulations could be more directly compared to the data produced by Sattenspiel's Norway House Ordinary Differential Equation (NHODE) model, a deterministic population-based mathematical model which had modeled the 1918-1919 flu epidemic at

all three communities rather than just at Norway House. It was hoped that such a comparison might provide insight into the relative merits of the two different modeling techniques for the study of epidemic disease in small populations.

While the differences between the techniques used in the NHOHGL computer simulation and the NHODE differential-equations system have already been detailed in chapter 4, it is important to reiterate here the two key features that set the two models apart from one another. First, the NHOHGL model uses an individual- or agent-based approach whereas the NHODE model uses a population-based approach. Because an individual-based model is better able to replicate individual diversity than a population-based model, which considers all members of a population to have equal characteristics, the NHOHGL model is better able to replicate a heterogeneous population than the NHODE model. Second, the NHOHGL model is stochastic whereas the NHODE model is deterministic. Because stochastic models account for the effects of randomness upon the data, while deterministic models produce the same results each time the model is run at the same set of parameter values, the NHOHGL model is more appropriate for the modeling of small populations, where randomness can have a greater influence upon the outcome of events.

As might be expected, a comparison of the data from the NHOHGL model with the data from the NHODE model reveals several differences that may be attributed to these fundamental differences in modeling technique. Two of these- the differing predictions regarding the intercommunity-spread of the 1918-1919 flu epidemic and the differing predictions regarding what would have happened if the epidemic had begun at

Oxford House or God's Lake instead of at Norway House- have already been discussed in the previous section. Additional differences, however, were also observed. To begin with, the NHODE model indicated that the timing of the epidemic and the extent of its spread were influenced by both the mobility variables and the population variables, including contact rates, whereas the severity of the epidemic was primarily determined by the contact rates within a community (Herring and Sattenspiel, 2003; Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002; Sattenspiel et al., 2000). The data from the NHOHGL model did not show as clear cut a pattern. Instead, the mobility parameters, contact parameters, and other population parameters in the NHOHGL model were all found to have multiple and complex influences upon the timing, spread, and severity of the simulated epidemic (see chapter 6). Additionally, while the NHODE model indicated that the location of introduction affected the timing of the epidemic, but not its severity (Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002), the data from the NHOHGL model indicated that the location of introduction affected almost every measure of the epidemic, including its severity. A further difference was observed between the simulated epidemic totals from the two models: the NHODE model predicted that the total number of cases would be 190 at Norway House and about 7 at both Oxford House and God's Lake (Sattenspiel and Herring, 2002) whereas, in the standard parameter replication test, the NHOHGL model predicted, on average, 374 at Norway House, and less than one at Oxford House and at God's Lake.

Although the differences between the data from the NHOHGL model and the data from the NHODE model are notable, the comparison also reveals many important

similarities. First of all, the overall shape of the epidemic curve is similar in both models, with an initial case building up to a rather short and defined epidemic peak. Second, in the NHOHGL model, just as in the NHODE model, the epidemic peak was larger and occurred earlier at Norway House than at the other communities (see Figure 6.2) (Sattenspiel and Herring, 1998; Sattenspiel et al., 2000). Finally, in both models, the traditional Cree seasonal patterns of residence and mobility were found to have a similar effect upon the simulated epidemic, with a summer epidemic having a significantly earlier and more severe epidemic peak (Herring and Sattenspiel, 2003). These similarities indicate that, at their most basic level, the models do agree, even though their more specific predictions regarding the epidemic details differ.

- **Corollary to Question 3: Is the difference in modeling technique directly responsible for the observed differences between the data from the NHOHGL model and the NHODE model?**

Although it would be easy to simply attribute the differences between the data produced by the NHOHGL model and the data produced by Sattenspiel's NHODE model to the different modeling techniques that they employ, this may be an oversimplification. The differing demands of a stochastic, individual-based, computer simulation versus a deterministic, population-based, mathematical model have led parameter design in the NHOHGL and NHODE models to differ considerably. These differences complicate any comparison of the two models and make it difficult to determine whether the observed differences truly result from a difference in modeling technique or are instead the result

of a difference in input parameter values. Because even a small difference in initial parameter values can have a large impact upon the epidemic output, it was important that this possibility be considered when making a comparison of the two models.

One important difference between the parameter design of the two models is the number of input parameters. While the NHOHGL model is controlled by a series of 41 input parameters, the NHODE model has only 4 adjustable parameters, which control the size of the population at each community, the fraction of contacts which result in disease transmission, the average number of contacts per unit of time, and the rate of recovery (Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002). Consequently, many of the NHOHGL model parameters do not have a direct parallel in the NHODE model.

To further complicate the issue, some of the parameters that are present in both the NHOHGL and NHODE models differ greatly in their design and are thus difficult to compare. For example, in the NHODE model, inter-community mobility is controlled by two variables, the probability of leaving one post to visit another and the distribution of travel by post, but these have been combined into a single series of parameters in the NHOHGL model, the between-post move probabilities. However, the NHOHGL model has also combined the standard values for the two NHODE mobility parameters in order to derive values for the between-post move probabilities and thus, the overall effect is the same in both models. For example, while the NHODE model specifies that the probability of leaving Oxford House is 0.000169 and that 60% of the travel from Oxford House is to Norway House (Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002), the NHOHGL model has multiplied these values in order to determine the

between-post move probability for trips from Oxford House to Norway House: 0.0001014.

The parameters controlling the length of the recovery/infectious period are also formulated quite differently in the two models, but here too, the values work out to be equivalent. The NHOHGL model uses a recovery time whereas the NHODE model uses a recovery rate. While recovery time in the NHOHGL model was set to last for 5 days, a period consistent with the length of the infectious period for influenza, the recovery rate in the NHODE model was set at 0.2 per day, a value that also led, on average, to recovery in 5 days (Sattenspiel and Herring, 1998).

Although the values for the inter-community mobility and recovery period parameters in the NHOHGL and NHODE models are relatively easy to compare, despite the differences in their design, the contact parameters present more of a challenge. The demands of an individual-based *versus* a population-based model have led to fundamental differences in the formulation of contact in the two models. While the population-based NHODE model includes a separate contact rate for each community, the individual-based NHOHGL model uses a separate contact rate for individuals, based upon whether the contact is or is not between members of the same family group. This difference makes comparison of these contact parameter values impossible.

For the few parameters that are shared and retain a similar form in both models, a more direct value comparison is possible. This comparison reveals that even though the parameter values in both models have been based upon the same set of historical, ethnographic, and epidemiological data, the values used in the NHOHGL model do not

always correspond with the values used in Sattenspiel's NHODE model. The notable differences include: the sizes of the populations at each of the three posts, which have been rounded in the NHOHGL model but which were left at their census estimates in the NHODE model, and the probability/rate of disease transmission parameters, which, being difficult to estimate, were set at 0.2 in the NHOHGL model but at 0.5 in the NHODE model.

Given these differences, it is quite possible that some of the observed variation between the data from the NHOHGL model and the data from Sattenspiel's NHODE model may result from differences in initial parameter values rather than from the differences in modeling technique. In order to test this, an additional set of 1000 simulations of the NHOHGL model was performed in which the initial parameter values were adjusted in order to replicate as closely as possible the parameter values that were used in the NHODE model. During these simulations, the value of disease probability was increased from 0.2 to 0.5 and death was turned off, because it had not been included in the NHODE model. The population sizes at the three communities, however, were left at their standard NHOHGL values, because using the rough census figures would have made dividing the agents among the post and camps impossible. The other NHOHGL parameters were also left at their standard winter values because, as shown above, these either had no correlate in the NHODE model or already had values that were similar to the values used in the NHODE model.

The data from this test may be viewed in Table 7.7, where they are listed alongside the averaged data from the NHOHGL winter replication test. While the data

Table 7.7: Epidemic data for the tests of the NHOHGL model using parameter values that approximate those used in Sattenspiel's NHODE model
These data are shown alongside the averaged data for the standard parameter replication test of the NHOHGL model in order to facilitate comparison.

	Standard Winter	NHODE Parameters
Duration of Infection	126	64
Length of the Epidemic	114	0
Peak Time of Infection	29	20
Peak Number Infected	78.20	220.17
Total Recovered	441.85	716.74
Total Dead	78.59	0
Total Infected	520.44	716.74
% Infected	37.7%	51.94%
% Dead	5.7%	0.0%
Epidemics at OH	18	35
Epidemics at GL	0	0
Epidemics at OH & GL	0	0
NH Dead	78.38	0
NH Infected	519.08	709.41
OH Dead	0.21	0
OH Infected	1.36	7.33
GL Dead	0.00	0
GL Infected	0.00	0

The standard NHOHGL winter parameters use a probability of infection of 0.2. The value of this parameter in the runs replicating the NHODE parameters is 0.5. Death has also been turned off in these test simulations.

from the test simulations using the NHODE parameter values clearly differ from the normal winter NHOHGL model data, they also differ from the data produced by Sattenspiel's NHODE model. Using the NHODE input values, the NHOHGL model predicts 709 cases of the flu at Norway House, 7 at Oxford House, and 0 at God's Lake.

These totals are even more divergent from the NHODE model predictions than were the original NHOHGL model estimates. Further, even when using the NHODE model input values, the NHOHGL model continues to indicate that the 1918-1919 flu

epidemic is unlikely to reach Oxford House, managing to do so just 3.5% of the time, and unable to reach God's Lake. The results of this test thus indicate that the observed differences between the predictions of the NHOHGL and NHODE models, at least regarding the impact and geographic spread of the epidemic, are related more to a difference in modeling technique than to a difference in input parameter values. The other differences between the predictions of the NHOHGL and NHODE models, including the impact of the location of first infection, mobility rates, and population density, could not be tested without running a larger series of tests.

Assessment of the NHOHGL Model: Is the model a good representation of the 1918-1919 flu epidemic in this region of Canada?

The process of determining whether a computer simulation is a reasonable representation of the system or phenomenon that is being modeled is known as validation testing (Gilbert and Troitzsch, 1999). In the case of the NHOHGL model, the validation testing required that the data from the winter simulations be compared to the historical data for the 1918-1919 influenza epidemic at Norway House. As noted in chapter 3, these data, largely derived from the HBC post journals and the Anglican Church of Canada burial records (Herring, 1994), provide an estimate of the duration of the epidemic at Norway House, an estimate of the death toll, and confirmation that the epidemic failed to reach either Oxford House or God's Lake. These statistics are limited in scope and could well have been influenced by any number of factors, so they are not ideal for model validation, yet, because they provide the only source of historical data on

the 1918-1919 flu epidemic at Norway House, they are the only means for testing the accuracy of the NHOHGL model.

A comparison of the data for the simulated NHOHGL epidemics with the historical epidemic data reveals several important differences. To start, while the HBC post journals indicate that the influenza epidemic at Norway House lasted only 6 weeks (Herring, 1994; Herring, 2000), or, about 42 days, the simulated winter epidemics have an average duration of 126.5 days, or about 4 months. The NHOHGL model thus appears to overestimate epidemic duration. Much of this difference, however, may be attributed to two key differences in how the duration of the epidemic is calculated in the model versus historically. First, tabulation of the length of the simulated epidemics began with the introduction of the first infected individual and ended when there were absolutely no more traces of the epidemic, whereas historically, the epidemic may have gone largely unnoticed until multiple deaths occurred and was considered to have ended when the last infection occurred. Second, the duration of the simulated epidemic considers the presence of the flu at all three communities whereas historically, the epidemic failed to reach either Oxford House or God's Lake. A recalculation of the duration of the simulated winter epidemics (from the winter parameter replication test) that considers only the epidemic at Norway House and stops counting upon the recovery or death of the final infected individual indicates an average duration of infection of 85.8 days. Although this estimate is much closer to the historical estimate, it is still a bit high and may indicate that the NHOHGL parameter values, especially the probability of infection, need adjustment.

A comparison of the mortality data from the NHOHGL winter simulations with the historical mortality estimates for the 1918-1919 flu epidemic at Norway House indicates that the value chosen for the probability of death parameter, and possibly also the death time parameter, could also use some adjustment. The Anglican Church of Canada burial records indicate that about 18-20% of the Norway House population died as a result of the 1918-1919 flu epidemic (Herring, 1994; Herring, 2000). The data from the NHOHGL model winter replication test, however, predict an average death rate at Norway House of 10.43%. This underestimation of epidemic mortality may be attributed to the fact that probability of death is extremely difficult to estimate, especially for a historical epidemic like the 1918-1919 flu, and to the unusual design of this parameter in the NHOHGL model. Thus, the value chosen for this parameter during the winter simulations (0.04) may simply not be a realistic estimate. Closer estimates of the historical death rates were achieved during the replication testing when the probability of death was increased to 0.1 (21.51%) or to 0.5 (18.73%) and when death time was decreased to one day (20.18%). Because these values lead to a closer approximation of the historical mortality rate, they may be more similar to the historical death rates and should be considered for any future modeling of the 1918-1919 flu at Norway House, Oxford House, and God's Lake.

One point upon which the historical data and the NHOHGL simulation data appear to readily agree is the lack of flu at Oxford House and God's Lake. As noted in the discussion of question 2, data from the HBC post journals indicate that the 1918-1919 flu epidemic failed to spread form Norway House to either Oxford House or God's Lake

(Herring, 1994; Herring and Sattenspiel, 2003; Sattenspiel and Herring, 2002). This observation is fully supported by the NHOHGL model simulations which indicate that, for any given set of 1000 runs of the model at the winter parameter values, the flu makes it to Oxford House only an average of 13 times and never makes it to God's Lake. Thus, the model generally predicts what was historically documented: an epidemic that remains confined at Norway House.

Overall, even though the NHOHGL model data fail to accurately predict the historical epidemic duration and death toll, it can still be considered to be a reasonable representation of the 1918-1919 influenza epidemic at Norway House, Oxford House and God's Lake. No model is an exact representation of reality. Thus, the NHOHGL model could not be expected to exactly replicate the historical epidemic 100% accurately 100% of the time. Further, given the possibility that the observed differences between the simulated data and the historical data could potentially be corrected by simply adjusting some of the more difficult to estimate parameter values, as well as the limited data regarding the 1918-1919 flu epidemic in this region, the NHOHGL model can indeed be considered to be a success. Future projects that apply the model to a more complete data set, for which firm estimates are available for more of the parameter values, may well confirm its utility.

Chapter 8: Discussion and Conclusions

In this project, agent-based computer simulation has provided a unique opportunity to clarify and expand the picture of the 1918-1919 flu epidemic in the Norway House region of Manitoba. While the utility of agent-based computer simulation for the modeling of historical epidemics in a small population had already been demonstrated by the original Norway House model (Carpenter, 2004), this project has greatly expanded the capabilities of this model through the addition of two additional communities, Oxford House and God's Lake, and inter-community mobility. The resulting Norway House, Oxford House, God's Lake (NHOHGL) model has allowed for the study of the 1918-1919 flu epidemic on a wider geographic scale, has enabled new questions to be addressed, and has facilitated a comparison between individual-based and population-based modeling techniques.

Following its construction, the new NHOHGL model was thoroughly tested in order to ensure that it was working properly and that it was an appropriate representation of the 1918-1919 flu epidemic at Norway House, Oxford House, and God's Lake. The verification study established that each of the model parameters was behaving as designed and that the data produced by the model simulations was roughly consistent. These tests also served to demonstrate how modification of the initial social and epidemiological framework of the model affected the spread and severity of the ensuing influenza epidemic. The validation testing then established that the new model is indeed a reasonable representation of the 1918-1919 flu epidemic at Norway House, Oxford

House and God's Lake, but indicated that it could possibly be made more accurate with the adjustment of the values of some of the epidemic parameters. Once the NHOHGL model had been thoroughly tested, additional simulations were run in order to investigate the three primary research questions.

The first question concerned the impact of the traditional Cree seasonal population movement and settlement patterns and seasonal differences in mobility rates upon the spread and severity of 1918-1919 flu epidemic in this region. In keeping with the findings of the original Norway House model (Carpenter, 2004), the NHOHGL model indicated that a summer epidemic would likely have been shorter and more severe than the historical epidemic, which had struck during the winter. This is because the higher population densities during the summer months, when the Cree extended family bands would gather at the posts, would have encouraged the rapid spread of the epidemic to more members of the population. The NHOHGL model also suggests that the epidemic would have been more likely to reach Oxford House if it had occurred during the summer rather than during the winter, due to increased summer travel rates and the higher proportion of infectious travelers. Additional investigations using the NHOHGL model have shown that it is the seasonal differences in settlement patterns rather than the seasonal differences in mobility rates, that account for most of the observed differences between the winter and summer simulations.

The second question to be investigated by the NHOHGL model concerned one of the most interesting mysteries surrounding the 1918-1919 flu epidemic, its highly variable pattern of spread. The HBC post journals have indicated that even as the flu

raged at Norway House, it was notably absent at both Oxford House and God's Lake, even though contact between the three communities continued. Such a scattered pattern of epidemic distribution was not unique to the Norway House region, or to Canada. The NHOHGL model suggests that, at least in this case, the reason that the epidemic failed to spread was simply that it couldn't. This was because the combination of low wintertime travel rates, the short infectious period of influenza, and the length of the journey between Norway House and its neighbors ensured that the few travelers who did journey from Norway House to Oxford House or God's Lake would generally either recover or die prior to their arrival, and would thus be unable to transmit the flu. While the NHOHGL model indicates that the 1918-1919 flu epidemic was highly unlikely to have reached Oxford House and almost completely incapable of reaching God's Lake under normal conditions, it has also shown that this natural epidemic protection had limits. Increasing the speed and frequency of travel, increasing contact rates, and/or increasing the virulence of the epidemic all led to an increased incidence of the flu at both Oxford House and God's Lake and, in the summer simulations, enabled the simulated epidemic to routinely reach all three posts. Because these parameter changes are not out of line with modern estimates, this finding offers a strong message of caution, especially given current concerns that the H5N1 avian flu virus could soon spark another major influenza pandemic.

The third question pertained more to the model itself than to its subject matter and involved a comparison of two very different epidemic modeling techniques. The addition of Oxford House, God's Lake, and inter-community mobility in the NHOHGL model has

enabled a more direct comparison to be made between this stochastic, individual-based computer simulation and Sattenspiel's NHODE model, a deterministic, population-based mathematical model, which used a series of differential equations to model the 1918-1919 flu epidemic at Norway House, Oxford House, and God's Lake (Sattenspiel and Dietz, 1995; Sattenspiel, 1990; Sattenspiel and Herring, 1998; Sattenspiel and Herring, 2002; Sattenspiel and Herring, 2003; Sattenspiel et al., 2000). It is important to note, however, that fundamental differences in model and parameter design mean that an exact comparison of the NHODE and NHOHGL models may never be entirely possible.

A comparison of the data produced by the two models has revealed several notable differences. First, the NHODE model always predicted at least a small number of cases of the flu at Oxford House and God's Lake but the NHOHGL model has provided support for the view that the flu would have been unlikely to spread to Oxford House and unable to spread to God's Lake, a conclusion that is in agreement with the historical data. Second, while the NHODE model predicted that an epidemic that began at Oxford House or God's Lake would have spread more quickly than an epidemic initiated at Norway House, the NHOHGL model predicted that an epidemic initiated at Oxford House or God's Lake would have failed to spread. Third, while the NHODE model indicated that the location of introduction affected the timing but not the severity of the simulated epidemic, the NHOHGL model indicated that this parameter influenced almost every measure of the epidemic, including severity. Fourth, the data from the NHODE model indicated that, while mobility rates, travel patterns, and contact rates influenced the timing and spread of the epidemic, contact rates were primarily responsible for the

severity of the epidemic. The NHOHGL model, in contrast, indicated that each of these parameters had multiple and complex influences. Finally, the NHODE model predicted a total of 190 cases of the flu at Norway House, 7 at Oxford House, and 7 at God's Lake, but the NHOHGL model, when set at the same input parameter values, predicted 709 cases at Norway House, 7 at Oxford House, and 0 at God's Lake. Further tests have shown that these differences most likely extend from the different modeling techniques that the two models employ.

The NHOHGL model has clearly provided a more in-depth view of the 1918-1919 flu epidemic in the Norway House region of Canada, and has also offered insight into the practice of epidemic modeling, but it was also found to have several important limitations. The first was time. Even with the displays turned off, it still takes five hours to run a single set of 1000 runs of the NHOHGL model, and this can be a serious constraint for data collection. Additional modifications to the model or perhaps a shift to a different running platform may help to alleviate this issue, but the time element is an important factor to consider when using this model in the future. Another important limitation was the data set. Several of the model parameters, including disease probability, death probability, and the time that it would have taken to complete a journey between each of the posts had to be estimated due to a lack of sufficient archival data. It may thus be useful to test the NHOHGL model using a more complete data set from another community or region, in order to better establish its level of accuracy. A third limitation concerns usability. In its current state, the NHOHGL model is difficult to use or modify without an extensive introduction to the model program code. The creation of

a more straightforward user interface would be helpful in this respect. Other limitations of the model stem from its overall design which, at this time, limits the length of the incubation period to one day, requires the length of a return journey to be equal to the length of the original journey, and allows the size of the agent population to be increased but not decreased. These limitations could potentially be alleviated with minor modifications to the NHOHGL model program.

Despite these limitations, much of the potential of the NHOHGL model still remains untapped. In its current form, the NHOHGL model could easily be used to investigate the impact and effectiveness of quarantine measures by simply reducing or halting travel between the three communities. With a modification to the model output, the NHOHGL model could also be used to investigate the pattern of epidemic spread within the individual family groups at each community or to track specific agents throughout the simulation. Such an application would be of particular interest given evidence that the 1918-1919 flu epidemic at Norway House appears to have affected certain families much more severely than others (Herring and Sattenspiel, 2003) and that this characteristic may also vary seasonally. The NHOHGL Model also has the potential to be used to investigate the role that differing infection or death rates among adult males, adult females, and the children and elderly may have played in the spread and severity of this or other epidemics.

The NHOHGL model also has almost limitless potential to be expanded and adapted to model different flu epidemics, different diseases, different communities, or different cultural groups. Data from other flu epidemics could be applied to the

NHOHGL model and then compared to the results of the present simulations in order to determine trends in infection and spread. The NHOHGL model could also be easily adapted to investigate other infectious diseases, such as measles or whooping cough, and could be used to study more complex diseases if additional disease states, such as a formal latency period, were added to the epidemic sequence. The model landscape could also be expanded by the addition of more communities and inter-connecting pathways, which would enable the modeling of larger geographic areas. A geographic information system (GIS) map could be integrated into the model landscape in order to create a more realistic environment for the agents (Kohler et al., 1999). Other factors, such as additional human behavioral patterns, like decision-making, family support systems, or underlying medical conditions, like endemic tuberculosis or nutritional factors, could be added to the model in order to enhance its explanatory potential.

The data presented in this thesis have shown that agent-based computer simulation, when used in combination with historical, archival, and epidemiological data, can be a useful tool for the study of epidemic disease, especially in small populations or where the available data are incomplete. By expanding the framework of the original Norway House model, the NHOHGL model has enabled the 1918-1919 flu epidemic to be modeled on a more regional scale and has also allowed for a valuable comparison to be made between population-based and individual-based modeling techniques. The data produced by the NHOHGL model simulations has helped to expand the picture of the 1918-1919 flu epidemic at Norway House, Oxford House, and God's Lake and, more generally, has provided important insights into how various human social factors and

epidemiological characteristics may influence the spread and severity of an influenza epidemic. These insights can help to shed light upon other influenza or virgin-soil epidemics for which less archival data is available. They can also be used to help predict, prevent, or minimize the spread and severity of future influenza epidemics. Today, as the world faces the potential of another major influenza pandemic in the form of the H5N1 avian flu virus, such insights may prove to be especially valuable.

Appendix A: The NHOHGL Model Program

Appendix A will present the computer code for the NHOHGL model, written in the Java™ computer language. The program consists of three interdependent segments: the parameters; the agents, the model, and the paths. The parameter section allows the researcher to easily adjust the parameter values by listing them all on one page and then inserting the correct values in the appropriate sections of the model. In the following code, the parameters are set at their standard values. The agents section keeps track of all the individual agent information. The model section is the most important section, and holds the bulk of the program code. The paths section is the smallest section and merely sets up the pathways between the three communities.

The code for each of these sections will be presented exactly as it appears when put into JBuilder. Any lines that are preceded by a "//" indicate notes or programmer's comments and are not active parts of the computer program. These comments may help to assist those not familiar with Java™ to understand the purpose of each section of the program.

NHOHGL Parameters

```
//NHOHGL SEIRD Study started 07/13/05 (Standard Parameters)

runs: 1000

numOfTimeSteps {
 set: 200
}
ScreenOutput {
 set: 0
}
graphs {
 set: 0
}
shouldDisplayWorld {
 set: 0
}
NumStepsPerDay {
 set: 6
}
MDiseaseProb  {
 set: 0.2
}
FDiseaseProb  {
  set: 0.2
}
CDiseaseProb {
  set:  0.2
}
MDeathProb  {
 set: 0.2
}
FDeathProb    {
  set:  0.2
}
CDeathProb    {
  set:  0.2
}
RecoveryTime{
  set: 5
}
DeathTime     {
  set:  2
```

```
}
NHFortPop     {
 set: 150
}
NHCampPop     {
  set: 150
}
NHMaleProp  {
  set:  0.25
}
NHFemaleProp          {
  set:  0.25
}
NHCampStay  {
  set:  0.99
}
NHNoCliques {
  set:  10
}
OHFortPop     {
 set: 66
}
OHCampPop     {
  set: 66
}
OHMaleProp  {
  set:  0.25
}
OHFemaleProp          {
  set:  0.25
}
OHCampStay  {
  set:  0.99
}
OHNoCliques {
  set:  5
}
GLFortPop     {
 set: 60
}
GLCampPop     {
  set: 60
}
GLMaleProp   {
```

```
  set:  0.25
}
GLFemaleProp          {
  set:  0.25
}
GLCampStay  {
  set:  0.99
}
GLNoCliques  {
  set:  5
}
PMeetingWI   {
  set:  0.5
}
PMeetingWO  {
  set:  0.001
}
TimeOfFirstInfection  {
  set: 20
}
NHtoOHMoveProb {
  set: 0.001
}
NHtoGLMoveProb {
  set: 0.001
}
OHtoNHMoveProb {
  set: 0.001
}
OHtoGLMoveProb {
  set: 0.001
}
GLtoNHMoveProb {
  set: 0.001
}
GLtoOHMoveProb {
  set: 0.001
}
```

Agents

```
//NHOHGL SEIRD Study started 07/13/05

package nhohgl;
import java.util.*;

public class Agents {
  //Intrinsic person stuff:
  String Sex;
  String DiseaseState;
  int InfectionLength;
  int Clique;

  //Location stuff:
  String Home;
  int HomeX;
  int HomeY;
  int XCoor;
  int YCoor;
  String CurrentLocation;

  //Movement stuff:
  int DirToHome;
  String Direction;
  int LastAction;
  static Random Select = new Random();
//for multiple forts, need to add probability to move from fort to fort

  public Agents(int X, int Y, String loc, double M, double F, double C,
              int Number, String Disease, int CliqueNo, int CampDir, int last){
    double T = (double) Select.nextInt(1000)/1000;
    Direction="none";
    if(T<M)
      Sex ="M";
    else if(T<M+F)
      Sex = "F";
    else
      Sex = "C";

    DiseaseState=Disease;
    XCoor=X;
    YCoor=Y;
    CurrentLocation = loc;
```

```
    HomeX=X;
    HomeY=Y;
    Home = loc;
    DirToHome=CampDir;
    InfectionLength=0;
    Clique =  CliqueNo;
    LastAction = last;


   /*  OLD:
      if(Number!=1)
        Direction= "fort";*/

    if(Number==1) //number 1 means they are a NH fort Agent
      Direction = "toOH";
    else if(Number==6)   //number 6 means they are a OH fort Agent
      Direction = "toNH";
    else
      Direction= "fort";

   }
}
```

Model

```
//NHOHGL SEIRD Study started 07/13/05
//modification of original norwayhouse model- Connie Carpenter
//New expanded Version- Carrie Ahillen

package nhohgl;

import uchicago.src.sim.space.*;
import java.util.*;
import uchicago.src.sim.util.SimUtilities;
import uchicago.src.sim.gui.*;
import uchicago.src.sim.engine.*;
import uchicago.src.sim.analysis.*;
import java.awt.Color;
import uchicago.src.sim.space.*;
import java.io.*;


public class Model extends SimpleModel {    //the world for the display colors

  Object2DGrid displayWorld;    //the world to deal with the paths and the camps
  Object2DGrid pathsWorld;     //some stuff for the display
  DisplaySurface dsurf;
  DataRecorder recorder;
  public OpenSequenceGraph graph;    //dynamic graph chart
  public OpenSequenceGraph NHgraph;    //dynamic graph chart
  public OpenSequenceGraph OHgraph;    //dynamic graph chart
  public OpenSequenceGraph GLgraph;    //dynamic graph chart

  //other variables

    static BufferedWriter datafile;

    //This determines whether the data is output to the screen
    //false means you only get a output to the file
    //true means you get both
    boolean ScreenOutput;

    //if graphs is true, then you get pretty graphs when you run it
    //false, no graphs, hopefully faster
    boolean graphs;
```

```
    //Describes how many squares on the world an agent moves per time step in the model
between the forts
    //1 means they move 1 square per day, 2 is 2 per day, etc
    int NumStepsPerDay;

    //this is the the number of steps (days) we run the model
    int numOfTimeSteps = 200;

    //this lets you turn the world display on and off
    //true means it displays
    //false means it will not be displayed
    boolean shouldDisplayWorld;



  int size;
  double p;

  String initDiseaseState;
  double MDiseaseProb;
  double FDiseaseProb;
  double CDiseaseProb;
  double MDeathProb;
  double FDeathProb;
  double CDeathProb;

//  int fortX;
  //int fortY;
  //first the Norway House fort
//NH fort coordinate variables
  int NHFortX;
  int NHFortY;
  double NHCampStay;
  int NHNoCliques;
  int NHPop;
  int NHFortPop;
  int NHCampPop;
  double NHMaleProp;
  double NHChildProp;
  double NHFemaleProp;


  int OHFortX;
  int OHFortY;
```

```
int OHPop;
int OHFortPop;
int OHCampPop;
double OHCampStay;
int OHNoCliques;
double OHMaleProp;
double OHChildProp;
double OHFemaleProp;

int GLFortX;
int GLFortY;
int GLPop;
int GLFortPop;
int GLCampPop;
double GLCampStay;
int GLNoCliques;
double GLMaleProp;
double GLChildProp;
double GLFemaleProp;


//Move Probabilities between the forts
double NHtoOHMoveProb;
double NHtoGLMoveProb;
double OHtoNHMoveProb;
double OHtoGLMoveProb;
double GLtoNHMoveProb;
double GLtoOHMoveProb;



int RecoveryTime;
int DeathTime;

int Time;
int TimeOfFirstInfection;
int Pop;
int Infected;
int Exposed;
int Susceptible;
int Recovered;
int Dead;
double PMeetingWI,PMeetingWO;
```

```
//NHFortInfected,NHFortDead,NHCamp1Infected,NHCamp1Dead,NHCamp2Infected,N
HCamp2Dead,NHCamp3Infected,NHCamp3Dead,NHCamp4Infected,NHCamp4Dead,O
HPop,OHSusc,OHInfected,OHRecvd,OHExp,OHDead,OHFortInfected,OHFortDead,OH
Camp1Infected,OHCamp1Dead,OHCamp2Infected,OHCamp2Dead,OHCamp3Infected,
OHCamp3Dead,OHCamp4Infected,OHCamp4Dead,GLPop,GLSusc,GLInfected,GLRec
vd,GLExp,GLDead,GLFortInfected,GLFortDead,GLCamp1Infected,GLCamp1Dead,GL
Camp2Infected,GLCamp2Dead,GLCamp3Infected,GLCamp3Dead,GLCamp4Infected,G
LCamp4Dead,
//Ok.  For the following post-wide (fort + camps) arrays, the indices are as follows:
   //0 is unused, but could/should be the totals for the whole model
   //1 is NH, 2 is OH, 3 is GL
   int postsPop[] = {0,0,0,0};
   int postsSusc[] = {0,0,0,0};
   int postsInf[] = {0,0,0,0};
   int postsRec[] = {0,0,0,0};
   int postsExp[] = {0,0,0,0};
   int postsDead[] = {0,0,0,0};
   //For these camps arrays,
   //0 is NHFort, 1 is NHCamp1, 2 is camp2,etc
   //5 is OHFort, 6 is OHCamp1. You get the idea
   //10 is GLFort. 11 is GLCamp1.  14 is GLCamp4. The end.
   int campsPop[] = {0,0,0,0,0,0,0,0,0,0,0,0,0,0,0};
   int campsSusc[] = {0,0,0,0,0,0,0,0,0,0,0,0,0,0,0};
   int campsInf[] = {0,0,0,0,0,0,0,0,0,0,0,0,0,0,0};
   int campsRec[] = {0,0,0,0,0,0,0,0,0,0,0,0,0,0,0};
   int campsExp[] = {0,0,0,0,0,0,0,0,0,0,0,0,0,0,0};
   int campsDead[] = {0,0,0,0,0,0,0,0,0,0,0,0,0,0,0};


 static Random select = new Random();    //allows random variables easily

 public Model() {                    //some stuff for the model
   Controller.ALPHA_ORDER = false;
 }

   class SeqSus implements Sequence {     // The sequence of the number
                                 //of susceptable
     public double getSValue() {        // This is the method to be defined
                                 //for Sequence
     return (double) Susceptible;      // Returns the appropriate value
     }
```

```
}

class SeqInf implements Sequence {    // The sequence of the number
                                //of infected
  public double getSValue() {       // This is the method to be defined
                               //for Sequence
  return (double) Infected;         // Returns the appropriate value
  }
}

class SeqRec implements Sequence {       // The sequence of the number
                                 // of recovered
  public double getSValue() {           // This is the method to be
                                //defined for Sequence
  return (double) Recovered;            // Returns the appropriate value
  }
}

 class SeqDea implements Sequence {        // The sequence of the number
                                  // of dead
  public double getSValue() {             // This is the method to be
                                //  defined for Sequence
  return (double) Dead;                 // Returns the appropriate value
  }
}


class SeqNHSus implements Sequence {     // The sequence of the number
                                //of susceptable
   public double getSValue() {        // This is the method to be defined
                               //for Sequence
   return (double) postsSusc[1];      // Returns the appropriate value
  }
}

class SeqNHInf implements Sequence {    // The sequence of the number
                                //of infected
  public double getSValue() {       // This is the method to be defined
                               //for Sequence
  return (double) postsInf[1];         // Returns the appropriate value
  }
}

class SeqNHRec implements Sequence {       // The sequence of the number
```

```
                                // of recovered
  public double getSValue() {          // This is the method to be
                               //defined for Sequence
  return (double) postsRec[1];          // Returns the appropriate value
  }
}

 class SeqNHDea implements Sequence {        // The sequence of the number
                                // of dead
  public double getSValue() {           // This is the method to be
                                //  defined for Sequence
  return (double) postsDead[1];                // Returns the appropriate value
  }
}

class SeqOHSus implements Sequence {     // The sequence of the number
                               //of susceptable
   public double getSValue() {        // This is the method to be defined
                              //for Sequence
   return (double) postsSusc[2];      // Returns the appropriate value
  }
}

class SeqOHInf implements Sequence {    // The sequence of the number
                              //of infected
  public double getSValue() {        // This is the method to be defined
                             //for Sequence
  return (double) postsInf[2];         // Returns the appropriate value
  }
}

class SeqOHRec implements Sequence {       // The sequence of the number
                                // of recovered
  public double getSValue() {          // This is the method to be
                               //defined for Sequence
  return (double) postsRec[2];          // Returns the appropriate value
  }
}

 class SeqOHDea implements Sequence {        // The sequence of the number
                                // of dead
  public double getSValue() {           // This is the method to be
                                //  defined for Sequence
  return (double) postsDead[2];                // Returns the appropriate value
```

```
    }
  }

  class SeqGLSus implements Sequence {     // The sequence of the number
                                //of susceptable
    public double getSValue() {        // This is the method to be defined
                               //for Sequence
    return (double) postsSusc[3];      // Returns the appropriate value
    }
  }

  class SeqGLInf implements Sequence {    // The sequence of the number
                               //of infected
    public double getSValue() {        // This is the method to be defined
                              //for Sequence
    return (double) postsInf[3];         // Returns the appropriate value
    }
  }

  class SeqGLRec implements Sequence {       // The sequence of the number
                                 // of recovered
    public double getSValue() {          // This is the method to be
                               //defined for Sequence
    return (double) postsRec[3];          // Returns the appropriate value
    }
  }

   class SeqGLDea implements Sequence {        // The sequence of the number
                                  // of dead
    public double getSValue() {            // This is the method to be
                                //  defined for Sequence
    return (double) postsDead[3];                // Returns the appropriate value
    }
  }

 public void setup() {
   super.setup();

   // Specify the parameters to be displayed for setting by the user
   params = new String[]
{"MDiseaseProb","FDiseaseProb","CDiseaseProb","MDeathProb","FDeathProb","CDeat
hProb","recoveryTime","deathTime","fortPop","campPop","maleProp","femaleProp","ca
mpStay","noCliques","pMeetingWI","pMeetingWO","timeOfFirstInfection"};
```

```
  setStoppingTime(numOfTimeSteps);
  size=100;                    //size of the world
  shouldDisplayWorld = true;

  //first the Norway House fort
  NHFortX = 25;   //NH fort coordinate variables
  NHFortY = 75;
  NHMaleProp=0.25;              //proportions of people types
  NHFemaleProp=0.25;
  NHChildProp=1-NHMaleProp-NHFemaleProp;
  NHCampStay=0.99;        //probability of staying in camp
  NHNoCliques = 10;        // Number of groups within fort
  NHCampPop=0;          // Population of each of the camps
  NHFortPop=750;       // Population of the fort
  NHPop = (NHFortPop + NHCampPop * 4);

  // the Oxford House fort
  OHFortX = 49;   //OH fort coordinate variables
  OHFortY = 51;
  OHMaleProp=0.25;              //proportions of people types
  OHFemaleProp=0.25;
  OHChildProp=1-OHMaleProp-OHFemaleProp;
  OHCampStay=0.99;        //probability of staying in camp
  OHNoCliques = 5;         // Number of groups within fort
  OHCampPop=0;          // Population of each of the camps
  OHFortPop=330;       // Population of the fort
  OHPop = (OHFortPop + OHCampPop * 4);


 // the God's Lake fort
  GLFortX = 61;   //GL fort coordinate variables
  GLFortY = 63;
  GLMaleProp=0.25;              //proportions of people types
  GLFemaleProp=0.25;
  GLChildProp=1-GLMaleProp-GLFemaleProp;
  GLCampStay=0.99;        //probability of staying in camp
  GLNoCliques = 5;         // Number of groups within fort
  GLCampPop=0;          // Population of each of the camps
  GLFortPop=300;       // Population of the fort
  GLPop = (GLFortPop + GLCampPop * 4);



//Move Probabilities between the forts
```

```
    NHtoOHMoveProb = 0.001;
    NHtoGLMoveProb = 0.001;
    OHtoNHMoveProb = 0.001;
    OHtoGLMoveProb = 0.001;
    GLtoNHMoveProb = 0.001;
    GLtoOHMoveProb = 0.001;

    p=1;                        //probability of moving along a path

    initDiseaseState="S";
    MDiseaseProb=0.2;    //probability of becoming infected upon contact
                    // with an infected person
    FDiseaseProb=0.2;
    CDiseaseProb=0.2;
    MDeathProb=0.5;    //probability of an infected person dying
    FDeathProb=0.5;
    CDeathProb=0.5;
    RecoveryTime=5;    //The amount of time that agent is infectious
    DeathTime=2;       //The person can die 2 days after infection up to including
recoveryTime
    PMeetingWI = 0.5;      // probability of meeting if members are
                    //of same clique
    PMeetingWO = 0.001;   // probability of meeting if members are
                    // of different clique
    Time=0;
    TimeOfFirstInfection=20;      // time when first infection introduced to fort

    displayWorld = new Object2DGrid(size,size);     //initializing the worlds
    pathsWorld = new Object2DGrid(size,size);

    // If there is already a graphics object from a previous run, we need
    // to delete it to clean up the screen.

    if (graph != null)
      graph.dispose();
    if (NHgraph != null)
      NHgraph.dispose();
    if (OHgraph != null)
      OHgraph.dispose();
    if (GLgraph != null)
      GLgraph.dispose();

    //stuff for the display
```

```
  if(dsurf != null)
  dsurf.dispose();
  dsurf= new DisplaySurface(this, "The World");
  registerDisplaySurface("The World", dsurf);

 }

 public void buildModel() {

  //recorder = new DataRecorder("./writeMe.txt", this);
  //recorder.createNumericDataSource("MDiseaseProb", this, "getMDiseaseProb");
  //initialize the world spaces

  for(int i=0;i<size;i++){
   for(int j=0;j<size;j++){
    displayWorld.putValueAt(i,j,0);
    Paths P = new Paths(i,j);
    pathsWorld.putObjectAt(i,j,P);
   }
  }

  //create fort, camps, & agents for each


//NH fort
  Paths fort = (Paths) pathsWorld.getObjectAt(NHFortX,NHFortY);
  fort.pop=NHFortPop;
  fort.moveProbs[8]=NHtoOHMoveProb;    //probability of moving towards OH from
NH
  fort.moveProbs[2]=NHtoGLMoveProb;    //probability of moving towards GL from
NH
  fort.moveProbs[4]=1-(NHtoOHMoveProb + NHtoGLMoveProb);

  fort.susceptible=fort.pop;
  fort.number=1;
  fort.color=3;
  fort.dirToNH=4;
  fort.dirToOH=8;
  fort.dirToGL=2;
  fort.locationType = "fort";
  //fortX=fort.xcoor;
  //fortY=fort.ycoor;
  displayWorld.putValueAt(fort.xcoor,fort.ycoor, fort.color);
```

```
//Making a camp ARRAY for all the camps. Weeee.
   //can change these numbers to change the distances of the paths from the forts
    Paths Camp[] = {null, (Paths) pathsWorld.getObjectAt(NHFortX+2,NHFortY),
(Paths) pathsWorld.getObjectAt(NHFortX,NHFortY+3), (Paths)
pathsWorld.getObjectAt(NHFortX-4,NHFortY), (Paths)
pathsWorld.getObjectAt(NHFortX,NHFortY-5)} ;

//Initializes the camps as much as we can in a for loop
    for(int i=1; i<5; i++){
     Camp[i].pop=NHCampPop;
     Camp[i].susceptible=Camp[i].pop;
     Camp[i].moveProbs[4]=NHCampStay;
     Camp[i].number=i+1;
     Camp[i].color=2;
     Camp[i].locationType = "camp";
    }

//taking care of stuff that can't be loop-ized in camp1 setup
    //moveProbs is an array of movement probabilities
    //[4] is probability of staying
  //can change these numbers to change the distances of the camps from the forts
    Camp[1].moveProbs[3]=1-Camp[1].moveProbs[4];  //move prob towards the fort
    displayWorld.putValueAt(NHFortX+2, NHFortY,Camp[1].color);

//taking care of stuff that can't be loop-ized in camp2 setup
    Camp[2].moveProbs[7]=1-Camp[2].moveProbs[4];
    displayWorld.putValueAt(NHFortX,NHFortY+3,Camp[2].color);

//taking care of stuff that can't be loop-ized in camp3 setup
    Camp[3].moveProbs[5]=1-Camp[3].moveProbs[4];
    displayWorld.putValueAt(NHFortX-4, NHFortY,Camp[3].color);

//taking care of stuff that can't be loop-ized in camp4 setup
    Camp[4].moveProbs[1]=1-Camp[4].moveProbs[4];
    displayWorld.putValueAt(NHFortX,NHFortY-5,Camp[4].color);

   //create the paths for camp1

    for(int i=NHFortX+1; i<NHFortX+2; i++){
     Paths river = (Paths) pathsWorld.getObjectAt(i,NHFortY);
     river.moveProbs[3]=p;    //probability of moving towards the fort
     river.moveProbs[5]=1-p;  //probability of moving towards the camp
     river.color=1;
```

```
  river.locationType="path";
  displayWorld.putValueAt(i,NHFortY,river.color);
}

//create the paths for camp2

for(int i=NHFortY+1;i<NHFortY+3;i++){
  Paths river = (Paths) pathsWorld.getObjectAt(NHFortX,i);
  river.moveProbs[7]=p;
  river.moveProbs[1]=1-p;
  river.color=3;
  river.locationType="path";
  displayWorld.putValueAt(NHFortX,i,river.color);
}


//create the paths for camp3

for(int i=NHFortX-1;i>NHFortX-4;i--){
  Paths river = (Paths) pathsWorld.getObjectAt(i,NHFortY);
  river.moveProbs[5]=p;
  river.moveProbs[3]=1-p;
  river.color=5;
  river.locationType="path";
  displayWorld.putValueAt(i,NHFortY,river.color);
   }

//create the paths for camp4

for(int i=NHFortY-1;i>NHFortY-5;i--){
  Paths river = (Paths) pathsWorld.getObjectAt(NHFortX,i);
  river.moveProbs[1]=p;
  river.moveProbs[7]=1-p;
  river.color=6;
  river.locationType="path";
  displayWorld.putValueAt(NHFortX,i,river.color);
}

//create the agents in the fort

for(int i=0;i<fort.pop;i++){
  int clique = select.nextInt(NHNoCliques)+ 1;
  Agents A = new Agents(fort.xcoor,fort.ycoor,"fort",NHMaleProp,NHFemaleProp,
                  NHChildProp,fort.number,initDiseaseState,clique,4,10);
```

```
        agentList.add(A);
    }

//creates the agents for camp1

    for(int i=0;i<Camp[1].pop;i++){
      int clique = select.nextInt(NHNoCliques)+ 1;
      Agents A = new
Agents(Camp[1].xcoor,Camp[1].ycoor,"camp",NHMaleProp,NHFemaleProp,
                        NHChildProp,Camp[1].number,initDiseaseState,clique,5,10);
      agentList.add(A);
    }

 //creates the agents for camp2

    for(int i=0;i<Camp[2].pop;i++){
      int clique = select.nextInt(NHNoCliques)+ 1;
      Agents A = new
Agents(Camp[2].xcoor,Camp[2].ycoor,"camp",NHMaleProp,NHFemaleProp,NHChildPr
op,Camp[2].number,initDiseaseState,clique,1,10);
      agentList.add(A);
    }

//creates the agents for camp3

    for(int i=0;i<Camp[3].pop;i++){
      int clique = select.nextInt(NHNoCliques)+ 1;
      Agents A = new
Agents(Camp[3].xcoor,Camp[3].ycoor,"camp",NHMaleProp,NHFemaleProp,NHChildPr
op,Camp[3].number,initDiseaseState,clique,3,10);
      agentList.add(A);
    }

 //creates the agents for camp4

    for(int i=0;i<Camp[4].pop;i++){
      int clique = select.nextInt(NHNoCliques)+ 1;
      Agents A = new
Agents(Camp[4].xcoor,Camp[4].ycoor,"camp",NHMaleProp,NHFemaleProp,NHChildPr
op,Camp[4].number,initDiseaseState,clique,7,10);
      agentList.add(A);
      }
```

```
// now initialize Oxford House fort, camps & agents

    //create fort, OHcamps, & agents for each

    Paths OHFort = (Paths) pathsWorld.getObjectAt(OHFortX,OHFortY);
    OHFort.moveProbs[0]=OHtoNHMoveProb;  //the probablility that an agent will
move towards NH
    OHFort.moveProbs[2]=OHtoGLMoveProb;  //the probablility that an agent will move
GL
    OHFort.moveProbs[4]=1-(OHtoNHMoveProb + OHtoGLMoveProb);
    OHFort.pop=OHFortPop;
    OHFort.susceptible=OHFort.pop;
    OHFort.number=6;
    OHFort.color=3;
    OHFortX=OHFort.xcoor;
    OHFortY=OHFort.ycoor;
    OHFort.dirToNH=0;
    OHFort.dirToOH=4;
    OHFort.dirToGL=2;
    OHFort.locationType = "fort";
    displayWorld.putValueAt(OHFort.xcoor,OHFort.ycoor, OHFort.color);


//Making a OH Camp ARRAY for all the OHCamps. Weeee.
    //can change these numbers to change the distances of the camps from the forts
     Paths OHCamp[] = {null, (Paths) pathsWorld.getObjectAt(OHFortX+2,OHFortY),
(Paths) pathsWorld.getObjectAt(OHFortX,OHFortY+3), (Paths)
pathsWorld.getObjectAt(OHFortX-4,OHFortY), (Paths)
pathsWorld.getObjectAt(OHFortX,OHFortY-5)};


//Initializes the OHcamps as much as we can in a for loop
     for(int i=1; i<5; i++){
       OHCamp[i].pop=OHCampPop;
       OHCamp[i].susceptible=OHCamp[i].pop;
       OHCamp[i].moveProbs[4]=OHCampStay;
       OHCamp[i].number=i+6;
       OHCamp[i].color=2;
       OHCamp[i].locationType = "camp";
     }

//taking care of stuff that can't be loop-ized in OHcamp1 setup
     //moveProbs is an array of movement probabilities
     //[4] is probability of staying
```

```
   //can change these numbers to change the distances of the camps from the forts
      OHCamp[1].moveProbs[3]=1-OHCamp[1].moveProbs[4];  //move prob towards the
fort
      displayWorld.putValueAt(OHFortX+2, OHFortY,OHCamp[1].color);

//taking care of stuff that can't be loop-ized in OHcamp2 setup
      OHCamp[2].moveProbs[7]=1-OHCamp[2].moveProbs[4];
      displayWorld.putValueAt(OHFortX,OHFortY+3,OHCamp[2].color);

//taking care of stuff that can't be loop-ized in OHcamp3 setup
      OHCamp[3].moveProbs[5]=1-OHCamp[3].moveProbs[4];
      displayWorld.putValueAt(OHFortX-4, OHFortY,OHCamp[3].color);

//taking care of stuff that can't be loop-ized in OHcamp4 setup
      OHCamp[4].moveProbs[1]=1-OHCamp[4].moveProbs[4];
      displayWorld.putValueAt(OHFortX,OHFortY-5,OHCamp[4].color);

    //create the paths for OHcamp1

     for(int i=OHFortX+1; i<OHFortX+2; i++){
       Paths river = (Paths) pathsWorld.getObjectAt(i,OHFortY);
       river.moveProbs[3]=p;    //probability of moving towards the fort
       river.moveProbs[5]=1-p;  //probability of moving towards the OHcamp
       river.color=1;
       river.locationType="path";
       displayWorld.putValueAt(i,OHFortY,river.color);
     }

     //create the paths for OHcamp2

     for(int i=OHFortY+1;i<OHFortY+3;i++){
       Paths river = (Paths) pathsWorld.getObjectAt(OHFortX,i);
       river.moveProbs[7]=p;
       river.moveProbs[1]=1-p;
       river.color=3;
       river.locationType="path";
       displayWorld.putValueAt(OHFortX,i,river.color);
     }


     //create the paths for OHcamp3

     for(int i=OHFortX-1;i>OHFortX-4;i--){
       Paths river = (Paths) pathsWorld.getObjectAt(i,OHFortY);
```

```
      river.moveProbs[5]=p;
      river.moveProbs[3]=1-p;
      river.color=5;
      river.locationType="path";
      displayWorld.putValueAt(i,OHFortY,river.color);
       }

    //create the paths for OHcamp4

    for(int i=OHFortY-1;i>OHFortY-5;i--){
      Paths river = (Paths) pathsWorld.getObjectAt(OHFortX,i);
      river.moveProbs[1]=p;  //probability of moving towards the fort once on the path
      river.moveProbs[7]=1-p; //probability of moving towards the camp once on the path
      river.color=6;
      river.locationType="path";
      displayWorld.putValueAt(OHFortX,i,river.color);
    }


    //create the agents in the OH fort

    for(int i=0;i<OHFort.pop;i++){
      int clique = select.nextInt(OHNoCliques)+ (NHNoCliques + 1);
      Agents A = new
Agents(OHFort.xcoor,OHFort.ycoor,"fort",OHMaleProp,OHFemaleProp,OHChildProp,
OHFort.number,initDiseaseState,clique,4,10);
        agentList.add(A);
    }

//creates the agents for OHcamp1

    for(int i=0;i<OHCamp[1].pop;i++){
      int clique = select.nextInt(OHNoCliques)+ (NHNoCliques + 1);
      Agents A = new
Agents(OHCamp[1].xcoor,OHCamp[1].ycoor,"camp",OHMaleProp,OHFemaleProp,OH
ChildProp,OHCamp[1].number,initDiseaseState,clique,5,10);
      agentList.add(A);
    }

  //creates the agents for OHcamp2

    for(int i=0;i<OHCamp[2].pop;i++){
      int clique = select.nextInt(OHNoCliques)+ (NHNoCliques + 1);
```

```
      Agents A = new
Agents(OHCamp[2].xcoor,OHCamp[2].ycoor,"camp",OHMaleProp,OHFemaleProp,OH
ChildProp,OHCamp[2].number,initDiseaseState,clique,1,10);
      agentList.add(A);
    }

//creates the agents for OHcamp3

    for(int i=0;i<OHCamp[3].pop;i++){
      int clique = select.nextInt(OHNoCliques)+ (NHNoCliques + 1);
      Agents A = new
Agents(OHCamp[3].xcoor,OHCamp[3].ycoor,"camp",OHMaleProp,OHFemaleProp,OH
ChildProp,OHCamp[3].number,initDiseaseState,clique,3,10);
      agentList.add(A);
    }

  //creates the agents for OHcamp4

    for(int i=0;i<OHCamp[4].pop;i++){
      int clique = select.nextInt(OHNoCliques)+ (NHNoCliques + 1);
      Agents A = new
Agents(OHCamp[4].xcoor,OHCamp[4].ycoor,"camp",OHMaleProp,OHFemaleProp,OH
ChildProp,Camp[4].number,initDiseaseState,clique,7,10);
      agentList.add(A);
      }



// now initialize God's Lake fort, camps & agents

          //create fort, GLcamps, & agents for each

          Paths GLFort = (Paths) pathsWorld.getObjectAt(GLFortX,GLFortY);
          GLFort.moveProbs[0]=GLtoNHMoveProb;  //probablility that an agent will
move to NH
          GLFort.moveProbs[6]=GLtoOHMoveProb;  //probablility that an agent will
move to OH
          GLFort.moveProbs[4]=1-(GLtoNHMoveProb + GLtoOHMoveProb);
          GLFort.pop=GLFortPop;
          GLFort.susceptible=GLFort.pop;
          GLFort.number=6;
          GLFort.color=3;
          GLFortX=GLFort.xcoor;
          GLFortY=GLFort.ycoor;
```

```
            GLFort.dirToNH=0;
            GLFort.dirToGL=4;
            GLFort.dirToOH=6;
            GLFort.locationType = "fort";
            displayWorld.putValueAt(GLFort.xcoor,GLFort.ycoor, GLFort.color);


//Making a GL Camp ARRAY for all the GLCamps. Weeee.
           //can change these numbers to change the distances of the camps from the forts
            Paths GLCamp[] = {null, (Paths)
pathsWorld.getObjectAt(GLFortX+2,GLFortY), (Paths)
pathsWorld.getObjectAt(GLFortX,GLFortY+3), (Paths)
pathsWorld.getObjectAt(GLFortX-4,GLFortY), (Paths)
pathsWorld.getObjectAt(GLFortX,GLFortY-5)};


//Initializes the GLcamps as much as we can in a for loop
            for(int i=1; i<5; i++){
              GLCamp[i].pop=GLCampPop;
              GLCamp[i].susceptible=GLCamp[i].pop;
              GLCamp[i].moveProbs[4]=GLCampStay;
              GLCamp[i].number=i+6;
              GLCamp[i].color=2;
              GLCamp[i].locationType = "camp";
            }

//taking care of stuff that can't be loop-ized in GLcamp1 setup
            //moveProbs is an array of movement probabilities
            //[4] is probability of staying
          //can change these numbers to change the distances of the camps from the forts
            GLCamp[1].moveProbs[3]=1-GLCamp[1].moveProbs[4];  //move prob towards
the fort
            displayWorld.putValueAt(GLFortX+2, GLFortY,GLCamp[1].color);

//taking care of stuff that can't be loop-ized in GLcamp2 setup
            GLCamp[2].moveProbs[7]=1-GLCamp[2].moveProbs[4];
            displayWorld.putValueAt(GLFortX,GLFortY+3,GLCamp[2].color);

//taking care of stuff that can't be loop-ized in GLcamp3 setup
            GLCamp[3].moveProbs[5]=1-GLCamp[3].moveProbs[4];
            displayWorld.putValueAt(GLFortX-4, GLFortY,GLCamp[3].color);

//taking care of stuff that can't be loop-ized in GLcamp4 setup
            GLCamp[4].moveProbs[1]=1-GLCamp[4].moveProbs[4];
```

```
 displayWorld.putValueAt(GLFortX,GLFortY-5,GLCamp[4].color);

//create the paths for GLcamp1

for(int i=GLFortX+1; i<GLFortX+2; i++){
  Paths river = (Paths) pathsWorld.getObjectAt(i,GLFortY);
  river.moveProbs[3]=p;    //probability of moving towards the fort
  river.moveProbs[5]=1-p;  //probability of moving towards the GLcamp
  river.color=1;
  river.locationType="path";
  displayWorld.putValueAt(i,GLFortY,river.color);
}

//create the paths for GLcamp2

for(int i=GLFortY+1;i<GLFortY+3;i++){
  Paths river = (Paths) pathsWorld.getObjectAt(GLFortX,i);
  river.moveProbs[7]=p;
  river.moveProbs[1]=1-p;
  river.color=3;
  river.locationType="path";
  displayWorld.putValueAt(GLFortX,i,river.color);
}


//create the paths for GLcamp3

for(int i=GLFortX-1;i>GLFortX-4;i--){
  Paths river = (Paths) pathsWorld.getObjectAt(i,GLFortY);
  river.moveProbs[5]=p;
  river.moveProbs[3]=1-p;
  river.color=5;
  river.locationType="path";
  displayWorld.putValueAt(i,GLFortY,river.color);
   }

//create the paths for GLcamp4

for(int i=GLFortY-1;i>GLFortY-5;i--){
  Paths river = (Paths) pathsWorld.getObjectAt(GLFortX,i);
  river.moveProbs[1]=p;  //probability of moving towards the fort once on the
path
  river.moveProbs[7]=1-p; //probability of moving towards the camp once on the
path
```

```
            river.color=6;
            river.locationType="path";
            displayWorld.putValueAt(GLFortX,i,river.color);
          }


          //create the agents in the GL fort

          for(int i=0;i<GLFort.pop;i++){
            int clique = select.nextInt(GLNoCliques)+ (NHNoCliques + OHNoCliques+
1);
            Agents A = new
Agents(GLFort.xcoor,GLFort.ycoor,"fort",GLMaleProp,GLFemaleProp,GLChildProp,G
LFort.number,initDiseaseState,clique,4,10);
              agentList.add(A);
          }

//creates the agents for GLcamp1

          for(int i=0;i<GLCamp[1].pop;i++){
            int clique = select.nextInt(GLNoCliques)+ (NHNoCliques + OHNoCliques+
1);
            Agents A = new
Agents(GLCamp[1].xcoor,GLCamp[1].ycoor,"camp",GLMaleProp,GLFemaleProp,GLC
hildProp,GLCamp[1].number,initDiseaseState,clique,5,10);
            agentList.add(A);
          }

        //creates the agents for GLcamp2

          for(int i=0;i<GLCamp[2].pop;i++){
            int clique = select.nextInt(GLNoCliques)+ (NHNoCliques + OHNoCliques+
1);
            Agents A = new
Agents(GLCamp[2].xcoor,GLCamp[2].ycoor,"camp",GLMaleProp,GLFemaleProp,GLC
hildProp,GLCamp[2].number,initDiseaseState,clique,1,10);
            agentList.add(A);
          }

//creates the agents for GLcamp3

          for(int i=0;i<GLCamp[3].pop;i++){
            int clique = select.nextInt(GLNoCliques)+ (NHNoCliques + OHNoCliques+
1);
```

```
            Agents A = new
Agents(GLCamp[3].xcoor,GLCamp[3].ycoor,"camp",GLMaleProp,GLFemaleProp,GLC
hildProp,GLCamp[3].number,initDiseaseState,clique,3,10);
            agentList.add(A);
           }

         //creates the agents for GLcamp4

           for(int i=0;i<GLCamp[4].pop;i++){
            int clique = select.nextInt(GLNoCliques)+ (NHNoCliques + OHNoCliques+
1);
            Agents A = new
Agents(GLCamp[4].xcoor,GLCamp[4].ycoor,"camp",GLMaleProp,GLFemaleProp,GLC
hildProp,Camp[4].number,initDiseaseState,clique,7,10);
            agentList.add(A);
            }




            //Create paths between the NH and OH forts

            for (int i = 1; i < NHFortY - OHFortY; i++) {
              Paths river = (Paths) pathsWorld.getObjectAt(NHFortX + i, NHFortY - i);
              river.moveProbs[0] = p; //probability of moving towards the NH fort
              river.moveProbs[8] = 1 - p; //probability of moving towards the OH fort
              river.color = 3;
              river.locationType = "fortpath";
              displayWorld.putValueAt(NHFortX + i, NHFortY - i, river.color);
            }

            //Create paths between the OH and GL forts

            for (int i = 1; i < GLFortY - OHFortY; i++) {
              Paths river = (Paths) pathsWorld.getObjectAt(OHFortX + i, OHFortY + i);
              river.moveProbs[0] = p; //probability of moving towards the NH fort
              river.moveProbs[8] = 1 - p; //probability of moving towards the OH fort
              river.color = 3;
              river.locationType = "fortpath";
              displayWorld.putValueAt(OHFortX + i, OHFortY + i, river.color);
            }

            //Create paths between the NH and GL forts, with a bend
```

```
  int newPathX = NHFortX;
  int newPathY = NHFortY + 1;
  boolean bend = false;
 for (int i = 1; i < GLFortX - NHFortX; i++) {
  if (GLFortX - newPathX == newPathY - GLFortY && bend == false)
  {
    bend = true;
  }
  if (bend)
     {
     newPathX++;
     newPathY--;
  }
  else {
     newPathX++;
  }
  Paths river = (Paths) pathsWorld.getObjectAt(newPathX, newPathY);
  river.moveProbs[0] = p; //probability of moving towards the NH fort
  river.moveProbs[8] = 1 - p; //probability of moving towards the OH fort
  river.color = 3;
  river.locationType = "fortpath";
  displayWorld.putValueAt(newPathX, newPathY, river.color);
 }
```

```
//variables for Norway House males, females, children
   int NHMales=0;
   int NHFemales=0;
   int NHChildren=0;
//variables for Oxford House males, females, children
   int OHMales=0;
   int OHFemales=0;
   int OHChildren=0;
//variables for God's Lake males, females, children
   int GLMales=0;
   int GLFemales=0;
   int GLChildren=0;


//this loop counts the number of males, females, and children in Norway House
//because these are randomized parameters
   for(int i=0;i<NHPop;i++){
```

```
    Agents A = (Agents) agentList.get(i);
    if(A.Sex=="M")
      NHMales++;
    if(A.Sex=="F")
      NHFemales++;
    else if(A.Sex=="C")
      NHChildren++;
   }


//this loop counts the number of males, females, and children in Oxford House
//because these are randomized parameters
      for(int i=NHPop;i< NHPop + OHPop;i++){
        Agents A = (Agents) agentList.get(i);
        if(A.Sex=="M")
          OHMales++;
        if(A.Sex=="F")
          OHFemales++;
        else if(A.Sex=="C")
          OHChildren++;
       }


//this loop counts the number of males, females, and children in God's Lake
//because these are randomized parameters
            for(int i=NHPop + OHPop;i< NHPop + OHPop + GLPop;i++){
              Agents A = (Agents) agentList.get(i);
              if(A.Sex=="M")
                GLMales++;
              if(A.Sex=="F")
                GLFemales++;
              else if(A.Sex=="C")
                GLChildren++;
            }

//IF ScreenOutput is set to true, then we print to the screen as the program runs
   if (ScreenOutput) {
     System.out.println("Total Pop: " + (NHPop + OHPop + GLPop) + ", NHPop: " +
NHPop + ", NHMales: " + NHMales +
                    ", NHFemales: " + NHFemales + ", NHChildren: " +
                    NHChildren + ", OHPop: " + OHPop + ", OHMales: " +
                    OHMales + ", OHFemales: " + OHFemales +
                    ", OHChildren: " + OHChildren + ", GLPop: " + GLPop + ", GLMales: "
+
```

```
                    GLMales + ", GLFemales: " + GLFemales +
                    ", GLChildren: " + GLChildren);
    }

    try {
      datafile.write("Total Pop: " + (NHPop + OHPop + GLPop) + ", NHPop: " + NHPop +
", NHMales: " + NHMales +
                    ", NHFemales: " + NHFemales + ", NHChildren: " +
                    NHChildren + ", OHPop: " + OHPop + ", OHMales: " +
                    OHMales + ", OHFemales: " + OHFemales +
                    ", OHChildren: " + OHChildren + ", GLPop: " + GLPop + ", GLMales: "
+
                    GLMales + ", GLFemales: " + GLFemales +
                    ", GLChildren: " + GLChildren + "\r\n");
      datafile.flush();
    }
    catch(IOException e) {
      System.out.println("could not write to data file");
      System.exit(0);
    }


  // Create a sequence chart graph and set it up
  // Labels and ranges for graph

  //if graphs is true, print the graphs
  //if graphs is false, no graphs
  if (graphs) {
    graph = new OpenSequenceGraph("SEIRD", this);
    graph.setXRange(0, numOfTimeSteps);
    graph.setYRange(0, 100);
    graph.setAxisTitles("Days", "Number of Cases");

    // Add the sequences we have defined

    graph.addSequence("INF", new SeqInf());
    graph.addSequence("SUS", new SeqSus());
    graph.addSequence("REC", new SeqRec());
    graph.addSequence("DEA", new SeqDea());

    graph.display(); // Display the graph right away
    graph.step(); // Do the first update to it

    NHgraph = new OpenSequenceGraph("NH-SEIRD", this);
```

```
NHgraph.setXRange(0, numOfTimeSteps);
NHgraph.setYRange(0, 100);
NHgraph.setAxisTitles("Days", "Number of Cases");

// Add the sequences we have defined

NHgraph.addSequence("INF", new SeqNHInf());
NHgraph.addSequence("SUS", new SeqNHSus());
NHgraph.addSequence("REC", new SeqNHRec());
NHgraph.addSequence("DEA", new SeqNHDea());

NHgraph.display(); // Display the graph right away
NHgraph.step(); // Do the first update to it

OHgraph = new OpenSequenceGraph("OH-SEIRD", this);
OHgraph.setXRange(0, numOfTimeSteps);
OHgraph.setYRange(0, 100);
OHgraph.setAxisTitles("Days", "Number of Cases");

// Add the sequences we have defined

OHgraph.addSequence("INF", new SeqOHInf());
OHgraph.addSequence("SUS", new SeqOHSus());
OHgraph.addSequence("REC", new SeqOHRec());
OHgraph.addSequence("DEA", new SeqOHDea());

OHgraph.display(); // Display the graph right away
OHgraph.step(); // Do the first update to it



   GLgraph = new OpenSequenceGraph("GL-SEIRD", this);
   GLgraph.setXRange(0, numOfTimeSteps);
   GLgraph.setYRange(0, 100);
   GLgraph.setAxisTitles("Days", "Number of Cases");

   // Add the sequences we have defined

   GLgraph.addSequence("INF", new SeqGLInf());
   GLgraph.addSequence("SUS", new SeqGLSus());
   GLgraph.addSequence("REC", new SeqGLRec());
   GLgraph.addSequence("DEA", new SeqGLDea());

   GLgraph.display(); // Display the graph right away
```

```
      GLgraph.step(); // Do the first update to it

  }


    DisplayConstants.CELL_WIDTH = 6;      //and then the dimensions of the grid
    DisplayConstants.CELL_HEIGHT = 6;

     //create the colors
     //a different color for each type of agent

     ColorMap Color_world = new ColorMap();
     Color_world.mapColor(0,Color.darkGray);
     Color_world.mapColor(1,Color.red);
     Color_world.mapColor(2,Color.green);
     Color_world.mapColor(3,Color.yellow);
     Color_world.mapColor(4,Color.blue);
     Color_world.mapColor(5,Color.white);
     Color_world.mapColor(6,Color.orange);


//this tells the program that the displays are there

     Value2DDisplay display_world = new Value2DDisplay(displayWorld,Color_world);
     if (shouldDisplayWorld) {
       dsurf.addDisplayable(display_world, "display1");
       addSimEventListener(dsurf);
     }

//this displays them right away

    for(int i=0;i<size;i++){
    for(int j=0;j<size;j++){
    Paths check = (Paths) pathsWorld.getObjectAt(i,j);
    if(check.pop>0)
    displayWorld.putValueAt(i,j,4);
     }
   }

   if (shouldDisplayWorld) {
     dsurf.display();
   }

     int pop=0;
```

```
    int infected=0;
    int susceptible=0;
    int recovered=0;
    int dead=0;

     //updating the display

    for(int i=0;i<size;i++){
    for(int j=0;j<size;j++){
    Paths check = (Paths) pathsWorld.getObjectAt(i,j);
    if(check.pop>0)
    displayWorld.putValueAt(i,j,4);
    else
    displayWorld.putValueAt(i,j,check.color);

    pop+=(int) check.pop;
    infected+=(int) check.infected;
    susceptible+=(int) check.susceptible;
    recovered+=(int) check.recovered;
    dead+=(int) check.dead;
      }
    }

    if (ScreenOutput) {
     String fortLetters = "NH";
     System.out.print("Time,Pop,Susceptible,Exposed,Infected,Recovered,Dead,");
     for (int numforts = 0; numforts < 3; numforts++) {
      if (numforts == 1) {fortLetters = "OH";}
      if (numforts == 2) {fortLetters = "GL";}
      System.out.print(fortLetters + "Pop," + fortLetters + "Susc," + fortLetters + "Exp,"
+
                  fortLetters + "Infected," + fortLetters + "Recvd," +
                  fortLetters + "Dead," +
                  fortLetters + "FortInfected," + fortLetters +
                  "FortDead,");
      for (int campnum = 1; campnum < 5; campnum++) {
       System.out.print(fortLetters + "Camp" + campnum + "Infected," +
                   fortLetters + "Camp" + campnum + "Dead,");
      }
     }
     System.out.println(""); //end the line
    }
```

```
  //This sends data to a text file

  try {
   String fortLetters = "NH";
   datafile.write("Time,Pop,Susceptible,Exposed,Infected,Recovered,Dead,");
   for (int numforts = 0; numforts < 3; numforts++) {
    if (numforts == 1) {fortLetters = "OH";}
    if (numforts == 2) {fortLetters = "GL";}
    datafile.write(fortLetters + "Pop," + fortLetters + "Susc,"  +
                   fortLetters + "Exp," + fortLetters + "Infected," +
                   fortLetters + "Recvd,"+ fortLetters + "Dead," +
                   fortLetters + "FortInfected," + fortLetters +
                   "FortDead,");
    for (int campnum = 1; campnum < 5; campnum++) {
     datafile.write(fortLetters + "Camp" + campnum + "Infected," +
                    fortLetters + "Camp" + campnum + "Dead,");
    }
   }
   datafile.write("\r\n"); //end the line
       datafile.flush();

  }

  catch(IOException e) {
   System.out.println("could not write to data file");
  }
  Time++;

  }

public void step() {

   //adds an infected agent

    if(Time==TimeOfFirstInfection){

    //initialize with one infected, change this so that the initial infected
   // agent happens outside the transient state of the model

   //this introduces the infection to NH fort (to intro. to OH fort,
   Agents X = (Agents) agentList.get(0);
   Paths fort = (Paths) pathsWorld.getObjectAt(X.XCoor,X.YCoor);
   X.DiseaseState="I";
   fort.susceptible--;
```

```
      fort.infected++;
      }
//checking and updating the status of the NH agents
      for(int i=0;i<NHPop;i++){
        Agents person = (Agents) agentList.get(i);

           if(person.DiseaseState == "E") {
           person.DiseaseState = "I";
           Paths update = (Paths) pathsWorld.getObjectAt(person.XCoor, person.YCoor);
           update.infected++;
           update.exposed--;
         }

         if (person.DiseaseState != "S") {
           Disease(person);
           continue;
         }

         encounters(person, i);
         Disease(person);
        }

       for(int i=0;i<NHPop;i++){
         Agents person = (Agents) agentList.get(i);
         if(person.Sex=="M" && person.DiseaseState != "D"){
           for (int numMoves = 0; numMoves < NumStepsPerDay; numMoves++) {
             //everyone needs to move once
             //people on fortpaths need to move numStepsPerDay
             if (person.CurrentLocation == "fortpath" || numMoves == 0) {
               Move(person);
             }
           }
         }
       }
//end of check/update NH agents

//checking and updating the status of the OH agents
        for(int i=NHPop;i<(NHPop + OHPop);i++){
         //this for loop loops through the OH agents in the agents list
         //to do this, we loop from the first agent after the last NH agent
         //to the last OH agent
          Agents person = (Agents) agentList.get(i);

             if(person.DiseaseState == "E") {
```

```
            person.DiseaseState = "I";
            Paths update = (Paths) pathsWorld.getObjectAt(person.XCoor, person.YCoor);
            update.infected++;
            update.exposed--;
          }

          if (person.DiseaseState != "S") {
            Disease(person);
            continue;
          }

          encounters(person, i);
          Disease(person);
         }

       for(int i=NHPop;i<(NHPop + OHPop);i++){
          Agents person = (Agents) agentList.get(i);
          if(person.Sex=="M" && person.DiseaseState != "D"){
           if (person.CurrentLocation == "fortpath") {
            for (int numMoves = 0; numMoves < NumStepsPerDay; numMoves++) {
             Move(person);
            }
           }
           else {
            Move(person);
           }
          }
        }
//end of check/update OH agents


    //checking and updating the status of the GL agents
          for(int i=(NHPop + OHPop);i<(NHPop + OHPop + GLPop);i++){
           //this for loop loops through the GL agents in the agents list
           //to do this, we loop from the first agent after the last OH agent
           //to the last GL agent
            Agents person = (Agents) agentList.get(i);

              if(person.DiseaseState == "E") {
              person.DiseaseState = "I";
              Paths update = (Paths) pathsWorld.getObjectAt(person.XCoor,
person.YCoor);
              update.infected++;
              update.exposed--;
```

```
            }

            if (person.DiseaseState != "S") {
               Disease(person);
               continue;
            }

            encounters(person, i);
            Disease(person);
           }

          for(int i=(NHPop + OHPop);i<(NHPop + OHPop + GLPop);i++){
            Agents person = (Agents) agentList.get(i);
            if(person.Sex=="M" && person.DiseaseState != "D") {
              if (person.CurrentLocation == "fortpath") {
                for (int numMoves = 0; numMoves < NumStepsPerDay; numMoves++) {
                  Move(person);
                }
              }
              else {
                Move(person);
              }
            }
           }
//end of check/update GL agents


        Pop=0;
        Infected=0;
        Susceptible=0;
        Recovered=0;
        Exposed=0;
        Dead=0;

        //updating the display

        for(int i=0;i<size;i++){
         for(int j=0;j<size;j++){
           Paths check = (Paths) pathsWorld.getObjectAt(i,j);
           if(check.pop>0)
             displayWorld.putValueAt(i,j,4);
           else
             displayWorld.putValueAt(i,j,check.color);
```

```
        //pop+=(int) check.pop;

        Infected+=(int) check.infected;
        Susceptible+=(int) check.susceptible;
        Recovered+=(int) check.recovered;
        Exposed+=(int)check.exposed;
        Dead+=(int)check.dead;
      }
     }

    Pop=Infected + Susceptible + Recovered + Exposed + Dead;

if (ScreenOutput) {
       System.out.print(Time + "," + Pop + "," + Susceptible + "," + Exposed +
                    "," + Infected + "," + Recovered + "," + Dead + ",");
    }

  try {
    datafile.write(Time + "," + Pop + "," + Susceptible + "," + Exposed + "," + Infected +
                "," + Recovered + "," + Dead + ",");
  } catch (IOException e) {
  System.out.println("could not open data file");
  System.exit(0);
}


for (int i = 0; i < 4; i++) {
  postsPop[i] = 0;
  postsSusc[i] = 0;
  postsInf[i] = 0;
  postsRec[i] = 0;
  postsExp[i] = 0;
  postsDead[i] = 0;
}
for (int i = 0; i < 15; i++) {
  campsPop[i] = 0;
  campsSusc[i] = 0;
  campsInf[i] = 0;
  campsRec[i] = 0;
  campsExp[i] = 0;
  campsDead[i] = 0;
}
```

```
for (int numagent = 0; numagent < agentList.size(); numagent++) {
 int campArrayIndex = -1; //init to out of bounds to help with error checking
 Agents person = (Agents) agentList.get(numagent);
 if (person.HomeX == NHFortX && person.HomeY == NHFortY)
  campArrayIndex = 0;
 else if (person.HomeX == NHFortX+2 && person.HomeY == NHFortY)
  campArrayIndex = 1;
 else if (person.HomeX == NHFortX && person.HomeY == NHFortY+3)
  campArrayIndex = 2;
 else if (person.HomeX == NHFortX-4 && person.HomeY == NHFortY)
  campArrayIndex = 3;
 else if (person.HomeX == NHFortX && person.HomeY == NHFortY-5)
  campArrayIndex = 4;
 else if (person.HomeX == OHFortX && person.HomeY == OHFortY)
  campArrayIndex = 5;
 else if (person.HomeX == OHFortX+2 && person.HomeY == OHFortY)
  campArrayIndex = 6;
 else if (person.HomeX == OHFortX && person.HomeY == OHFortY+3)
  campArrayIndex = 7;
 else if (person.HomeX == OHFortX-4 && person.HomeY == OHFortY)
  campArrayIndex = 8;
 else if (person.HomeX == OHFortX && person.HomeY == OHFortY-5)
  campArrayIndex = 9;
 else if (person.HomeX == GLFortX && person.HomeY == GLFortY)
  campArrayIndex = 10;
 else if (person.HomeX == GLFortX+2 && person.HomeY == GLFortY)
  campArrayIndex = 11;
 else if (person.HomeX == GLFortX && person.HomeY == GLFortY+3)
  campArrayIndex = 12;
 else if (person.HomeX == GLFortX-4 && person.HomeY == GLFortY)
  campArrayIndex = 13;
 else if (person.HomeX == GLFortX && person.HomeY == GLFortY-5)
  campArrayIndex = 14;

 (campsPop[campArrayIndex])++;

 if (person.DiseaseState == "S") {
  (campsSusc[campArrayIndex])++;
 }
 else if (person.DiseaseState == "E") {
  (campsExp[campArrayIndex])++;
 }
 else if (person.DiseaseState == "I") {
```

```
      (campsInf[campArrayIndex])++;
    }
    else if (person.DiseaseState == "R") {
      (campsRec[campArrayIndex])++;
    }
    else if (person.DiseaseState == "D") {
      (campsDead[campArrayIndex])++;
    }
  }

  for (int fortnum = 1; fortnum <= 3; fortnum++) {
    for (int campnum = 0; campnum < 5; campnum++) {
      postsPop[fortnum] += campsPop[campnum + 5*(fortnum - 1)];
      postsSusc[fortnum] += campsSusc[campnum + 5*(fortnum - 1)];
      postsInf[fortnum] += campsInf[campnum + 5*(fortnum - 1)];
      postsRec[fortnum] += campsRec[campnum + 5*(fortnum - 1)];
      postsExp[fortnum] += campsExp[campnum + 5*(fortnum - 1)];
      postsDead[fortnum] += campsDead[campnum + 5*(fortnum - 1)];
    }
  }

/* Debugging code that prints all the camps*[] arrays
  System.out.println("Category,1,2,3,4,5,6,7,8,9,10,11,12,13,14");
  System.out.print("CampsPop,");
  for (int index = 0; index < 15; index++) {
        System.out.print(campsPop[index] + ",");
  }
  System.out.println("");
  System.out.print("CampsSusc,");
  for (int index = 0; index < 15; index++) {
    System.out.print(campsSusc[index] + ",");
  }
  System.out.println("");
  System.out.print("CampsInf,");
  for (int index = 0; index < 15; index++) {
    System.out.print(campsInf[index] + ",");
  }
  System.out.println("");
  System.out.print("CampsRec,");
  for (int index = 0; index < 15; index++) {
    System.out.print(campsRec[index] + ",");
  }
  System.out.println("");
  System.out.print("CampsExp,");
```

```
   for (int index = 0; index < 15; index++) {
     System.out.print(campsExp[index] + ",");
   }
   System.out.println("");
   System.out.print("CampsDead,");
   for (int index = 0; index < 15; index++) {
     System.out.print(campsDead[index] + ",");
   }
   System.out.println("");
*/



   if (ScreenOutput) {
     for (int postNum = 1; postNum < 4; postNum++) {
       System.out.print(postsPop[postNum] + "," + postsSusc[postNum] + "," +
postsExp[postNum] +
                        "," + postsInf[postNum] + "," + postsRec[postNum] + "," +
                        postsDead[postNum] + ",");
       for (int index = 0; index < 5; index++) {
         System.out.print(campsInf[index + 5*(postNum - 1)] + "," + campsDead[index +
5*(postNum - 1)] + ",");
       }
     }
}

//This sends data to a text file

try {
     for (int postNum = 1; postNum < 4; postNum++) {
       datafile.write(postsPop[postNum] + "," + postsSusc[postNum] + "," +
postsExp[postNum] +
                        "," + postsInf[postNum] + "," + postsRec[postNum] + "," +
                        postsDead[postNum] + ",");
       for (int index = 0; index < 5; index++) {
         datafile.write(campsInf[index + 5*(postNum - 1)] + "," + campsDead[index +
5*(postNum - 1)] + ",");
       }
     }

  datafile.write("\r\n");
  datafile.flush();
}
catch (IOException e) {
```

```
  System.out.println("could not open data file");
  System.exit(0);
}

/*
for (int numforts = 0; numforts < 3; numforts ++) {

        Paths camp1 = (Paths) pathsWorld.getObjectAt(NHFortX+2,NHFortY);
        Paths camp2 = (Paths) pathsWorld.getObjectAt(NHFortX,NHFortY+3);
        Paths camp3 = (Paths) pathsWorld.getObjectAt(NHFortX-4,NHFortY);
        Paths camp4 = (Paths) pathsWorld.getObjectAt(NHFortX,NHFortY-5);
        Paths fort = (Paths) pathsWorld.getObjectAt(NHFortX, NHFortY);


      if (ScreenOutput) {
        System.out.print( fort.infected + "," + fort.dead + "," + camp1.infected + "," +
camp1.dead + "," +
                    camp2.infected + "," + camp2.dead + "," + camp3.infected +
                    "," + camp3.dead + "," + camp4.infected + "," +
                    camp4.dead + ",");
      }

      //This sends data to a text file

      try {
        datafile.write(Time + "," + Pop + "," + Susceptible + "," + Infected +
                  "," + Recovered + "," + Exposed + "," + Dead + "," +
                  fort.dead + "," + camp1.dead + "," + camp2.dead + "," +
                  camp3.dead + "," + camp4.dead + "," + fort.infected +
                  "," + camp1.infected + "," + camp2.infected + "," +
                  camp3.infected + "," + camp4.infected + "\r\n");
        datafile.flush();
      }
      catch (IOException e) {
        System.out.println("could not open data file");
        System.exit(0);
      }
    }
*/
     //recorder.record();
     // ... and write it into the file right away.
    // recorder.writeToFile();

    if (graphs) {
```

```
      graph.step();
      NHgraph.step();
      OHgraph.step();
      GLgraph.step();
    }
    if (shouldDisplayWorld) {
      dsurf.updateDisplay();
    }
     Time++;
    }

//Disease and Death probability implemented
  public void Disease(Agents person){
        Paths P = (Paths) pathsWorld.getObjectAt(person.XCoor,person.YCoor);
        double disProb = (double) (select.nextInt(1000)+1)/1000;  // selects a number
between 0 & 1  but not = to 0
        double deaProb = (double) (select.nextInt(1000)+1)/1000;
        double disease=0;
        int infected = 0;
        int dead = 0;

        if(person.InfectionLength >= DeathTime && person.InfectionLength <
RecoveryTime && MDeathProb >= deaProb){
                person.DiseaseState = "D";
                P.infected--;
                P.dead++;
                person.InfectionLength=0;
        }

        if(person.InfectionLength==RecoveryTime && MDeathProb >= deaProb){
          person.DiseaseState = "D";
          P.infected--;
          P.dead++;
          person.InfectionLength=0;

        } else if (person.InfectionLength == RecoveryTime && person.DiseaseState !=
"D") {
          person.DiseaseState = "R";
          P.infected--;
          P.recovered++;
          person.InfectionLength=0;
        }

        if(person.DiseaseState=="I") {
```

```
          person.InfectionLength++;
        }
      }

private int RollDice(Agents person) {
      Paths spot = (Paths) pathsWorld.getObjectAt(person.XCoor,person.YCoor);
      double pM = (double) (select.nextInt(1000)+1)/1000;
      double sumprobs=0;
      int nextAction=10;
      int i=0;

      while (i < 9 && pM > sumprobs) {
        if(spot.moveProbs[i] != -1)
          sumprobs += spot.moveProbs[i];
        if (pM <= sumprobs) {
          nextAction = i;
        }
        i++;
      }
      return nextAction;
    }


 public void Move(Agents person){          //chooses only the males
    int nextAction=10;
    Paths spot = (Paths) pathsWorld.getObjectAt(person.XCoor,person.YCoor);


    if (person.CurrentLocation == "camp") {
      nextAction = RollDice(person);
    }
    else if (person.CurrentLocation == "fort") {
      //if you're from a camp, go back

      if (person.Home == "camp") {
        nextAction = person.DirToHome;
      }
      /*So now we know person.Home == "fort"*/
      //if you're from this fort, decide whether you're going out
      else {
        if (person.HomeX == person.XCoor && person.HomeY == person.YCoor) {
          nextAction = RollDice(person);
        }
        //if you're not from this fort, go back
```

```
    else {
      if (person.HomeX == NHFortX && person.HomeY == NHFortY) {
        nextAction = spot.dirToNH;
      }
      else if (person.HomeX == OHFortX && person.HomeY == OHFortY) {
        nextAction = spot.dirToOH;
      }
      else if (person.HomeX == GLFortX && person.HomeY == GLFortY) {
        nextAction = spot.dirToGL;
      }
    }
  }
}
else if (person.CurrentLocation == "fortpath")  {
  //this block takes care of the west bend on the path between NH & GL
  if (person.XCoor == NHFortX + 1 && person.YCoor == NHFortY + 1) {
    if (person.LastAction == 2)
      nextAction = 5;
    else if (person.LastAction == 3)
      nextAction = 6;
  }
  //this block takes care of the east bend on the path between NH & GL
  else if (GLFortX - person.XCoor == person.YCoor - GLFortY &&
          person.YCoor == NHFortY + 1) {
    if (person.LastAction == 5)
      nextAction = 8;
    else if (person.LastAction == 0)
      nextAction = 3;
  }
  else
    nextAction = person.LastAction;
}  //implies (person.Location == "path")
else nextAction = person.LastAction;

  //actually execute the movements of the agents
  //start by subtracting them from the current square
  //updates the population of the spot

  if (person.DiseaseState=="S") {
    spot.susceptible--;
  }
  if (person.DiseaseState=="I") {
    spot.infected--;
  }
```

```
if (person.DiseaseState=="R") {
  spot.recovered--;
}
if (person.DiseaseState=="D") {
  spot.dead--;
}

spot.pop=spot.infected+spot.recovered+spot.susceptible+spot.dead;


//Then move them
if(nextAction==0){
  person.XCoor--;
  person.YCoor++;
}
if(nextAction==1){
  person.YCoor++;
}
if(nextAction==2){
  person.XCoor++;
  person.YCoor++;
}
if(nextAction==3){
  person.XCoor--;
}
if(nextAction==5){
  person.XCoor++;
}
if(nextAction==6){
  person.XCoor--;
  person.YCoor--;
}
if(nextAction==7){
  person.YCoor--;
}
if(nextAction==8){
  person.XCoor++;
  person.YCoor--;
}

person.LastAction = nextAction;  //Ok, we just did "nextAction", whatever it was
                                 //Now we need to record what action we took this time
                                 //So the agents have a memory of their actions
```

```
        //newSpot is the square the person just moved onto
         Paths newSpot = (Paths) pathsWorld.getObjectAt(person.XCoor,person.YCoor);

         person.CurrentLocation = newSpot.locationType;

        //Now add them to the population of the spot

         if(person.DiseaseState=="S")
           newSpot.susceptible++;

         if(person.DiseaseState=="I")
           newSpot.infected++;

         if(person.DiseaseState=="R")
           newSpot.recovered++;

         if(person.DiseaseState=="D")
           newSpot.dead++;


newSpot.pop=newSpot.infected+newSpot.recovered+newSpot.susceptible+newSpot.dead
;


    }

  public void  encounters(Agents person, int i) {
     double probMeeting = 0;

     for(int j=0; j < agentList.size(); j++) {
           if (i==j){
           continue;
         }

         Agents personEncountered = (Agents)agentList.get(j);
         if (personEncountered.DiseaseState != "I") {
           continue;
         }

          if (person.XCoor == personEncountered.XCoor && person.YCoor ==
personEncountered.YCoor)   {
              if (person.Clique == personEncountered.Clique  ) {
                 probMeeting = PMeetingWI;
               } else  {
```

```
                probMeeting = PMeetingWO;
                 }
        }

            double prob = (double) (select.nextInt(1000)+1)/1000;

            if (probMeeting >= prob) {
               prob = (double) (select.nextInt(1000)+1)/1000;

               if (person.Sex=="M" && MDiseaseProb >= prob) {
                  person.DiseaseState = "E";
                  Paths P = (Paths) pathsWorld.getObjectAt(person.XCoor,person.YCoor);
                  P.exposed++;
                  P.susceptible--;
                  break;
               }

               if (person.Sex=="F" && FDiseaseProb >= prob) {
                  person.DiseaseState = "E";
                  Paths P = (Paths) pathsWorld.getObjectAt(person.XCoor,person.YCoor);
                  P.exposed++;
                  P.susceptible--;
                  break;
               }

               if (person.Sex=="C" && CDiseaseProb >= prob) {
                  person.DiseaseState = "E";
                  Paths P = (Paths) pathsWorld.getObjectAt(person.XCoor,person.YCoor);
                  P.exposed++;
                  P.susceptible--;
                  break;
               }

            } else {}
        }
}

//Setting up Mortality Parameter page

public double getMDiseaseProb() {
   return MDiseaseProb;
 }

public void setMDiseaseProb (double n) {
```

```
    MDiseaseProb = n;
  }

public double getFDiseaseProb() {
    return FDiseaseProb;
  }

public void setFDiseaseProb (double n) {
    FDiseaseProb = n;
  }

public double getCDiseaseProb() {
    return CDiseaseProb;
   }

public void setCDiseaseProb (double n) {
    CDiseaseProb = n;
  }

public double getMDeathProb() {
    return MDeathProb;
  }

public void setMDeathProb (double n) {
    MDeathProb = n;
  }

public double getFDeathProb() {
    return FDeathProb;
  }

public void setFDeathProb (double n) {
    FDeathProb = n;
  }

public double getCDeathProb() {
    return CDeathProb;
   }

public void setCDeathProb (double n) {
    CDeathProb = n;
  }

public int getRecoveryTime ()  {
```

```
    return RecoveryTime;
  }

public void setRecoveryTime (int n)  {
    RecoveryTime = n;
  }
  public int getdeathTime ()  {
     return DeathTime;
   }

 public void setdeathTime (int n)  {
     DeathTime = n;
   }
public int getNHFortPop ()  {
    return NHFortPop;
  }
public void setNHFortPop (int n) {
    NHFortPop = n;
 }
public int getNHCampPop ()  {
    return NHCampPop;
  }
public void setNHCampPop (int n)  {
    NHCampPop = n;
  }
public double getNHMaleProp ()  {
    return NHMaleProp;
  }
public void setNHMaleProp (double n)  {
    NHMaleProp = n;
  }
public double getNHFemaleProp ()  {
    return NHFemaleProp;
  }
public void setNHFemaleProp (double n)  {
    NHFemaleProp = n;
  }
public double getNHCampStay ()  {
    return NHCampStay;
  }
public void setNHCampStay (double n)  {
    NHCampStay = n;
  }
public int getNHNoCliques ()  {
```

```
    return NHNoCliques;
  }
public void setNHNoCliques (int n)  {
    NHNoCliques = n;
  }
  public int getOHFortPop ()  {
    return OHFortPop;
  }
public void setOHFortPop (int n) {
    OHFortPop = n;
 }
public int getOHCampPop ()  {
    return OHCampPop;
  }
public void setOHCampPop (int n)  {
    OHCampPop = n;
  }
public double getOHMaleProp ()  {
    return OHMaleProp;
  }
public void setOHMaleProp (double n)  {
    OHMaleProp = n;
  }
public double getOHFemaleProp ()  {
    return OHFemaleProp;
  }
public void setOHFemaleProp (double n)  {
    OHFemaleProp = n;
  }
public double getOHCampStay ()  {
    return OHCampStay;
  }
public void setOHCampStay (double n)  {
    OHCampStay = n;
  }
public int getOHNoCliques ()  {
    return OHNoCliques;
  }
public void setOHNoCliques (int n)  {
    OHNoCliques = n;
  }
  public int getGLFortPop ()  {
      return GLFortPop;
    }
```

```
  public void setGLFortPop (int n) {
     GLFortPop = n;
   }
  public int getGLCampPop ()  {
     return GLCampPop;
    }
  public void setGLCampPop (int n)  {
     GLCampPop = n;
    }
  public double getGLMaleProp ()  {
     return GLMaleProp;
    }
  public void setGLMaleProp (double n)  {
     GLMaleProp = n;
    }
  public double getGLFemaleProp ()  {
     return GLFemaleProp;
    }
  public void setGLFemaleProp (double n)  {
     GLFemaleProp = n;
    }
  public double getGLCampStay ()  {
     return GLCampStay;
    }
  public void setGLCampStay (double n)  {
     GLCampStay = n;
    }
  public int getGLNoCliques ()  {
     return GLNoCliques;
    }
  public void setGLNoCliques (int n)  {
     GLNoCliques = n;
    }
public double getPMeetingWI ()  {
    return PMeetingWI;
  }
public void setPMeetingWI (double n)  {
    PMeetingWI = n;
  }
public double getPMeetingWO ()  {
    return PMeetingWO;
  }
public void setPMeetingWO (double n)  {
    PMeetingWO = n;
```

```
    }
  public int getTimeOfFirstInfection ()  {
     return TimeOfFirstInfection;
    }
  public void setTimeOfFirstInfection (int n)  {
     TimeOfFirstInfection = n;
    }

 //Accessors for the Move Probabilities between the forts
 //allows them to be parameters
  public double getNHtoOHMoveProb ()  {
    return NHtoOHMoveProb;
  }
public void setNHtoOHMoveProb (double n)  {
    NHtoOHMoveProb = n;
  }
  public double getNHtoGLMoveProb ()  {
  return NHtoGLMoveProb;
}
public void setNHtoGLMoveProb (double n)  {
  NHtoGLMoveProb = n;
}
  public double getOHtoNHMoveProb ()  {
  return OHtoNHMoveProb;
}
public void setOHtoNHMoveProb (double n)  {
  OHtoNHMoveProb = n;
}
  public double getOHtoGLMoveProb ()  {
  return OHtoGLMoveProb;
}
public void setOHtoGLMoveProb (double n)  {
  OHtoGLMoveProb = n;
}
  public double getGLtoNHMoveProb ()  {
  return GLtoNHMoveProb;
}
public void setGLtoNHMoveProb (double n)  {
  GLtoNHMoveProb = n;
}
  public double getGLtoOHMoveProb ()  {
  return GLtoOHMoveProb;
}
public void setGLtoOHMoveProb (double n)  {
```

```
  GLtoOHMoveProb = n;
}
public int getScreenOutput ()  {
  if (ScreenOutput)
    return 1;
  else return 0;
}
public void setScreenOutput (int n)  {
  if (n == 0)
    ScreenOutput = false;
  else ScreenOutput = true;
}
public int getgraphs ()  {
  if (graphs)
    return 1;
  else return 0;
}
public void setgraphs (int n)  {
  if (n == 0)
    graphs = false;
  else graphs = true;
}
public int getNumStepsPerDay ()  {
  return NumStepsPerDay;
}
public void setNumStepsPerDay (int n)  {
  NumStepsPerDay = n;
}
public int getnumOfTimeSteps ()  {
  return numOfTimeSteps;
}
public void setnumOfTimeSteps (int n)  {
  numOfTimeSteps = n;
}
public int getshouldDisplayWorld ()  {
  if (shouldDisplayWorld)
    return 1;
  else return 0;
}
public void setshouldDisplayWorld (int n)  {
  if (n == 0)
    shouldDisplayWorld = false;
  else shouldDisplayWorld = true;
}
```

```
  public String[] getInitParam() {
      String[] params =
{"MDiseaseProb","FDiseaseProb","CDiseaseProb","MDeathProb","FDeathProb","CDeat
hProb","RecoveryTime","DeathTime","NHFortPop","NHCampPop","NHMaleProp","N
HFemaleProp","NHCampStay","NHNoCliques","OHFortPop","OHCampPop","OHMale
Prop","OHFemaleProp","OHCampStay","GLFortPop","GLCampPop","GLMaleProp","
GLFemaleProp","GLCampStay","GLNoCliques","OHNoCliques","PMeetingWI","PMee
tingWO","TimeOfFirstInfection", "NHtoOHMoveProb", "NHtoGLMoveProb",
"OHtoNHMoveProb", "OHtoGLMoveProb", "GLtoNHMoveProb",
"GLtoOHMoveProb",
"ScreenOutput","graphs","NumStepsPerDay","numOfTimeSteps","shouldDisplayWorld"
};
      return params;
   }

// Make sure the path is appropriate for the computer you are working on (Where is your
parameter sheet stored)

  public static void main(String[] args) {

   try {
    datafile = new BufferedWriter (new FileWriter ("testdata.txt",true));

    SimInit init = new SimInit();
    Model m = new Model();
    init.loadModel(m, "C:/Documents and Settings/Tom
Ahillen/jbproject/NHOHGL/NHOHGL Parameters.txt",true);
    datafile.close();
   }

   catch(IOException e) {
    System.out.println("could not open data file");
   }

  //init.loadModel(m, null, false);


 }

}
```

Paths

```
//NHOHGL SEIRD Study started 07/13/05

package nhohgl;

public class  Paths {
  double moveProbs[];
  int exposed;
  int infected;
  int recovered;
  int dead;
  int susceptible;
  int pop;
  int xcoor;
  int ycoor;
  int number;
  int color;
  int dirToNH;
  int dirToOH;
  int dirToGL;
  String locationType;


  public Paths(int x, int y){
     moveProbs = new double[9];
     exposed=0;
     infected=0;
     recovered=0;
     dead=0;
     susceptible=0;
     number=0;
     color=0;
     pop=infected+recovered+susceptible+dead;
     xcoor=x;
     ycoor=y;
     for(int i=0;i<9;i++){
       moveProbs[i]=-1;
     }

  }
}
```

Appendix B: The Excel Macro

The program code for the Microsoft Excel Macro program that was used in the analysis of the NHOHGL simulation data is presented in this section. The code is written in Visual Basic and is designed to take the rough output from a set of NHOHGL simulations and average the data from each run in order to produce summary epidemic data. The Macro then takes these averaged data and places them into a Microsoft Excel table for easy access.

The Macro

```
Sub macro1()

'OHsick and GLsick count the number of times disease hits each settlement
Dim OHsick As Integer, GLsick As Integer, Bothsick As Integer

Dim OHsickBool As Boolean
Dim GLsickBool As Boolean
'A is the array that holds the sums of the columns
'L is a line from the file
't is the right side of the line
Dim A(200, 60) As Long, L As String, t As String
'i, j, k are loop variables.  i goes throught the number of columns
'j counts the number of runs
'k is the number of lines in a run
'x is the length of right side of a line
'y is the length of the individual column entry
Dim i As Integer, j As Integer, k As Integer, x As Integer, y As Integer
'#1 is the name of the file we're opening
Open "C:\Documents and Settings\carrie\jbproject\NHOHGL\testdata.txt" For Input As
#1
j = 0
OHsick = 0
GLsick = 0
Bothsick = 0
```

```
OHsickBool = False
GLsickBool = False
Do While Not EOF(1)     'do this for the whole file
Line Input #1, L        'put Line 1 into L, to throw it away
'break up line 2(labels) and put it in SS:
Line Input #1, L        'Put line 2 into L
If j = 0 Then           'Only need to label stuff once
  x = Len(L)            'x is the length of the line
  For i = 1 To 54       'loop through the columns.  this is 54 because that's the number of
commas (delimiters) between columns
   t = Right(L, x)      'extracts the right part of the line that hasn't been processed yet
   y = InStr(1, t, ",", 1)      'gets the length from the beginning of t to the end of the column
   Cells(1, i) = Left(t, y - 1)      'adds the column into the array
   x = x - y            'subtracts that column out of the line
  Next i                'continue i loop
  Cells(1, 55) = Right(t, x)        'add the last column into the array
End If

For k = 1 To 200        'loop through the next 200 lines in the file
Line Input #1, L        'get a line in L
  x = Len(L)            'x is the length of the line
  For i = 1 To 54       'loop through the columns.  this is 54 because that's the number of
commas (delimiters) between columns
   t = Right(L, x)      'extracts the right part of the line that hasn't been processed yet
   y = InStr(1, t, ",", 1)      'gets the length from the beginning of t to the end of the column
   A(k, i) = A(k, i) + Val(Left(t, y - 1))      'adds the column into the array
   'implement two counters: count the occurances of disease in OH and GL
   If k = 200 And i = 25 And Val(Left(t, y - 1)) < 330 Then 'check for oxford house
getting sick
       OHsickBool = True
   ElseIf k = 200 And i = 41 And Val(Left(t, y - 1)) < 300 Then 'check for god's lake
house getting sick
       GLsickBool = True
   End If
   x = x - y            'subtracts that column out of the line
  Next i                'continue i loop
  A(k, 55) = A(k, 55) + Val(Right(t, x))        'add the last column into the array
 Next k                 'continue k loop
 j = j + 1              'increment the number of runs
 If OHsickBool = True Then
    OHsick = OHsick + 1
 End If
 If GLsickBool = True Then
   GLsick = GLsick + 1
```

```
 End If
 If OHsickBool = True And GLsickBool = True Then
    Bothsick = Bothsick + 1
 End If

 OHsickBool = False
 GLsickBool = False
Loop                  'continue the while not EOF loop
line1:                'label?
Close #1              'close the file
For k = 1 To 200      'loop through the lines again
 For i = 1 To 55      'loop throught the columns again
  Cells(k + 2, i) = A(k, i) / j   'Put the avg in the Cells in the Spreadsheet
 Next i               'continue i loop
Next k                'continue k loop
Cells(204, 1) = "Number of Runs:"
Cells(204, 2) = j     'Put the number of runs into the next line
Cells(204, 4) = "Number of times OH got sick:"
Cells(204, 5) = OHsick
Cells(204, 7) = "Number of times GL got sick:"
Cells(204, 8) = GLsick
Cells(204, 10) = "Number of times OH and GL got sick:"
Cells(204, 11) = Bothsick

End Sub
```

Bibliography

Appenzeller T. 2005. Tracking the next killer flu. National Geographic 208:2-31.

Axtell RL, Epstein JM, Dean JS, Gumerman GJ, Swedlund AC, Jason H, Chakravarty S, Hammond R, Parker J, and Parker M. 2002. Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. Proceedings of the National Academy of Science 99:7275-7279.

Barrett CL, Eubank SG, and Smith JP. 2005. If smallpox strikes Portland. Scientific American 292:54-61.

Bishop CA. 1974.The Northern Ojibwa and the fur trade: an historical and ecological study. Toronto: Holt, Reinhart and Winston.

Borland. 2005. JBuilder: accelerate your Java™ development. Borland Software Corporation. Accessed: Nov 28, 2005. Last update: 2005. http://www.borland.com/us/products/jbuilder/index.html

Burghes DN, Huntley I, McDonald J. 1982. Applying mathematics. In: Burghes DN, Huntley I, McDonald J, editors. Applying mathematics. New York: John Wiley and Sons. p 140-160.

Campione M, and Walrath K. 2005. About the Java technology. Sun Microsystems. Accessed: Nov 28 2005. Last update: 2005. http://java.sun.com/docs/books/tutorial/getStarted/intro/definition.html

Carpenter C. 2004. Agent-based modeling of seasonal population movement and the spread of the 1918-1919 flu: the ffect on a small community. Master of Arts thesis, Department of Anthropology. Columbia, MO: University of Missouri.

Centers for Disease Control and Prevention (CDC). 2004. Prevention and control of influenza: recommendations of the advisory committee on immunization practices (ACIP). Morbidity and Mortality Weekly Reports 53:1-40.

Corish JL, editor. 1920. Health knowledge. New York: Domestic Health Society.

Couch RB, and Kasel JA. 1983. Immunity to influenza in man. Annual Review of Microbiology 37:529-549.

Cox NJ, and Subbarao K. 1999. Influenza. Lancet 354:1277-1283.

Crawford DH. 2000. The invisible enemy: a natural history of viruses. New York: Oxford University Press.

Crosby AW. 2003. America's forgotten pandemic. Cambridge: Cambridge University Press.

Cross M, and Moscardini AO. 1985. Learning the art of mathematical modelling. New York: John Wiley and Sons.

Dean JS, Gumerman GJ, Epstein JM, Axtell RL, Swedlund AC, Parker MT, and McCarroll S. 1999. Understanding Anasazi culture change through agent-based modeling. In: Kohler T, Gumerman G, editors. Dynamics in human and primate societies. London: Oxford University Press. p 179-205.

Epstein JM, and Axtell RL. 1996. Growing aritificial societies: social science from the bottom up. Washington D.C.: Brookings Institution Press.

Flannery R. 1995. Ellen Smallboy: glimpses of a Cree woman's life. Montreal: McGill-Queen's University Press.

Gilbert N, and Troitzsch KG. 1999. Simulation for the Social Scientist. Buckingham: Open University Press.

Hallowell AI. 1992. The Ojibwa of Berens River, Manitoba: ethnography into history. Brown JSH, editor. Fort Worth: Harcourt Brace.

Hanks C. 1982. The Swampy Cree and the Hudson's Bay Company at Oxford House. Ethnohistory 29:103-115.

Hart T. 2004. Microterrors: the complete guide to bacterial, viral and fungal infections that threaten our health. Buffalo: Firefly Books.

Herring DA. 1994. "There were young people and old people and babies dying every week": the 1918-1919 influenza pandemic at Norway House. Ethnohistory 41:73-105.

Herring DA. 2000. Mundane diseases can kill: the 1918 influenza Pandemic in Canada. Journal of the Ontario Occupational Health Nurses Association summer 2000: 6-11.

Herring DA, and Sattenspiel L. 2003. Death in winter: Spanish flu in the Canadian subarctic. In: Phillips H, Killingray D, editors. The Spanish iInfluenza pandemic of 1918-1919. London: Routledge. p 156-172.

Herring DA, Abonyi S, and Hoppa RD. 2003. Malnutrition among northern peoples of Canada in the 1940s: an ecological and economic disaster. In: Herring DA, Swedlund AC, editors. Human biologists in the archives. Cambridge: Cambridge University Press. p 289-310.

Holmes EC, Taubenberger JK, and Grenfell BT. 2005. Heading off an influenza pandemic. Science 309:989.

James DJG, and McDonald JJ. 1981. An introduction to modeling. In: James DJG, McDonald JJ, editors. Mathematical modeling. New York: John Wiley and Sons. p 1-16.

Johnson NPAS. 2003. The overshadowed killer: influenza in Britian in 1918-19. In: Phillips H, Killingray D, editors. The Spanish influenza pandemic of 1918-19: new perspectives. New York: Routledge. p 132-155.

Karlen A. 1995. Man and microbes. New York: Touchstone.

Kohler TA. 1999. Putting social sciences together again: an introduction to the volume. In: Kohler TA, Gumerman GJ, editors. Dynamics in human and primate societies. London: Oxford University Press. p 1-18.

Kohler TA, and Gumerman GJ, editiors. 1999. Dynamics in human and primate societies: agent-based modeling of social and spatial processes. New York: Oxford University Press.

Kohler TA, Kresl J, Van West C, Carr E, and Wilshusen RH. 1999. Be there then: a modeling approach to settlement determinants and spatial efficiency among late ancestral Pueblo populations of the Mesa Verde region, U.S. southwest. In: Kohler TA, Gumerman GJ, editors. Dynamics in human and primate societies. London: University Press. p 145-177.

Lazzari S, Stöhr K. 2004. Avian influenza and influenza pandemics. Bulletin of the World Health Organization 84:242.

Long JS. 1995. Historical context. In: Flannery R, editor. Ellen Smallboy: glimpses of a Cree woman's life. Montreal: McGill-Queen's University Press. p 65-75.

Longini IM, Jr., Halloran ME, Nizam A, and Yang Y. 2004. Containing pandemic influenza with antiviral agents. American Journal of Epidemiology 159:623-633.

Lytwyn VP. 2002. Muskekowuck Athinuwick: original people of the great swampy land. Winnipeg: University of Manitoba Press.

McNeil WH. 1976. Plagues and peoples. Garden City, NY: Anchor Press.

Open Source Technology Group (OSTG). 2005. Repast: recursive porus agent simulation toolkit. Open Source Technology Group. Accessed: Nov 28, 2005. Last update: 2005. http://repast.sourceforge.net/

Ray AJ. 1974. Indians in the fur trade: their role as trappers, hunters, and middlemen in the lands southwest of the Hudson Bay 1660-1870. Toronto: University of Toronto Press.

Ray AJ. 1976. Diffusion of diseases in the western interior of Canada, 1830-1850. Geographical Review 66:139-157.

Reid AH, Janczewski TA, Lourens RM, Elliot AJ, Daniels RS, Berry CL, Oxford JS, and Taubenberger JK. 2003. 1918 influenza pandemic caused by highly conserved viruses with two receptor-binding variants. Emerging Infectious Diseases 9:1249-1253.

Sattenspiel L. 1990. Modeling the spread of infectious disease in human populations. Yearbook of Physical Anthropology 33:245-276.

Sattenspiel L. 2003. Infectious diseases in the historical archives: a modeling approach. In: Herring DA, Swedlund AC, editors. Human biologists in the archives: demography, health, nutrition and genetics in historical populations. Cambridge: Cambridge University Press. p 234-265.

Sattenspiel DL, and Dietz K. 1995. A structured epidemic model incorporating geographic mobility among regions. Mathematical Biosciences 128:71-91.

Sattenspiel L, Herring DA. 1998. Structured epidemic models and the spread of influenza in the central Canadian Subarctic. Human Biology 70:91-115.

Sattenspiel L, Herring DA. 2002. Human mobility and the spread of the 1918-1919 influenza epidemic in the Norway House region of Manitoba. In: Allum C, Kahn J, Cluney C, Peuramaki-Brown M, editors. Ancient travellers: proceedings of the twenty-Seventh annual conference of the archaeological assocaition of the University of Calgary. Calgary: The Archaeological Association of the University of Calgary. p 245-258.

Sattenspiel L, Herring DA. 2003. Simulating the effect of quarantine on the spread of the 1918-1919 flu in central Canada. Bulletin of Mathematical Biology 65:1-26.

Sattenspiel L, Mobarry A, and Herring DA. 2000. Modeling the influence of settlement structure on the spread of influenza among communities. American Journal of Human Biology 12:736-748.

Skinner A. 1909. Some remarks on the culture of eastern near-arctic Indians. Science 29:150-152.

Smith JGE. 1981. The Western Woods Cree: anthropological myth and historical reality. American Ethnologist 28:133-156.

Stevens J, Corper AL, Basler CF, Taubenberger JK, Palese P, and Wilson IA. 2004. Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. Science 303:1866-1870.

Taubenberger JK, and Reid AH. 1997. Initial characterization of the 1918 'Spanish" influenza virus. Science 275:1793-1797.

Taubenberger JK, Reid AH, and Fanning TG. 2005a. Capturing a killer flu virus. Scientific American 292:62-71.

Taubenberger JK, Reid AH, Lourens RM, Wang R, Jin G, and Fanning TG. 2005b. Characterization of the 1918 influenza virus polymerase genes. Nature 437:889-893.

Tumpey TM, Basler CF, Aguilar P, Zeng H, Sorlorzano A, Swayne DE, Cox NJ, Katz JM, Taubenberger JK, Palese P and Garcia-Sastre A. 2005. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. Science 310:77-80.

World Health Organization (WHO). 2003. Influenza fact sheet number 21. World Health Organization. Accessed: Dec 6, 2005. Last update: March 2003. http://www.who.int/mediacentre/factsheets/fs211/en/print.html

World Health Organization (WHO). 2004. Pandemic preparedness. World Health Organization. Accessed: Dec 6, 2004. Last update: 2004. http://www.who.int/csr/disease/influenza/pandemic/en/print.html