Communicable Disease Disaster Management

by

Alexandra S. Ndiwalana
27094074

Study leader: Dr. Johan W. Joubert

submitted in partial fulfillment of the requirements for the degree of

Bachelors of Industrial Engineering

in the

Faculty of Engineering, Built Environment and Information Technology

University of Pretoria

5 October 2010

Executive Summary

Outbreaks of communicable diseases in the African context are considered disasters. As such, disaster management protocols and methodologies are applied to combat communicable diseases and the resulting consequences. Outbreaks escalate into epidemics and the prediction of outbreaks and epidemics is near impossible. However, it is possible to monitor outbreaks and the associated spread patterns. The intelligence gained can be used for proactive decision making that will facilitate the expedient execution of retrospective disaster management activities.

An interactive simulation of outbreaks and the disease transmission that will lead to epidemics was developed. The simulation was focused on two diseases namely H1N1 (Swine-flu) and Measles within the City of Johannesburg municipal area. Insight gleaned from this simulation would facilitate proactive decision making in the area and inform future simulation based epidemiology studies.

Mathematical epidemiology and Agent Based Modelling (ABM) are two techniques that in combination are expected to produce a realistic simulation. Mathematical epidemiology is the application of mathematics and related concepts in the study of disease. Compartmental epidemiology is a subset of mathematical epidemiology where individuals of the concerned population or location are grouped into one of three groups. Each group or compartment has one of the following states assigned to it and all its occupants: Susceptible, Infected and Recovered.

Measles and H1N1 both have a Susceptible Infected Recovered (SIR) compartmental infectious disease models. When exploring the SIR model in a stochastic context, a Markov Chain is an applicable tool to enable the modelling of inter-state transition of an individual within a population. ABM is used to study complex systems and to convey how macro phenomena emerge from micro level behaviour and interactions between agents in an environment. An epidemic (macro phenomenon) is the consequence of many lower level individual infections and the associated disease transmission (micro phenomena).Compartmental epidemiology is thus used to demonstrate disease transmission while ABM will be the interface that enables simulation of the interaction of humans within a population or environment.

Contents

List of Figures

List of Tables

List of Abbreviations

ABM	Agent Based Model
CoJ	City of Johannesburg
COST	Changing One Separate factor at a Time
CSIR	Council for Scientific and Industrial Research
EPAR	Epidemic and Pandemic Alert and Response
DOE	Design of Experiments
GIS	Geographical Information Systems
NICD	National Institute of Communicable Disease
SADC	Southern African Development Community
SD	System Dynamics
SI	Susceptible Infected
SIS	Susceptible Infected Susceptible
SIR	Susceptible Infected Recovered
SIRS	Susceptible Infected Recovered Susceptible
UN	United Nations
UNICEF	United Nations International Children's Emergency Fund
WHO	World Health Organisation

Chapter 1

Introduction

Outbreaks of communicable diseases (infectious diseases) in an African context are considered disasters. This is due to the ease with which an outbreak can reach epidemic or pandemic proportions, and the speed at which a disease can cross local and national borders. Hence, disaster management protocols and methodologies are applied to combat outbreaks and their consequences (Lusambo-Dikasa, 2008).

Disaster management is divided into four life-cycle activities (Green and Altay, 2006):

Mitigation Risk identification and the associated contingency planning. This activity is concerned with the long term reduction or elimination of risks.

Preparedness The activity where action plans for disasters are developed in preparation for disasters.

Response Responding to a disaster entails mobilisation of emergency services and other relevant stakeholders.

Recovery Once the actual disaster is deemed under control, steps must be taken to restore, replenish and repair all that was damaged or lost due to the disaster to its original state.

It is important to note that only the first two activities, Mitigation and Preparedness, are proactive or prospective, since they place an emphasis on the *possibility* of a disaster occurring in the future. The remaining two activities are reactive or retrospective since they occur with respect to a disaster that has *already occurred*. Preparedness however, can also be viewed as retrospective since it is still applicable when a disaster has already occurred.

The speed with which one reacts to a disaster is critical in the accomplishment of the retrospective activities. One should aim to contain the communicable disease in it's early stages to the smallest possible population and area. By limiting the affected area and population, one is able to limit the resulting devastation and death.

A communicable disease is defined as a disease that can be directly or indirectly transmitted between humans. Transmission occurs either directly (person-to-person) or indirectly through vectors (third parties) which have been infected by a pathogenic micro-organism. Pathogenic micro-organisms, of which there are four main types, are the causes of infectious diseases. The four main types are, viruses, bacteria, parasites and fungi. Once an individual contracts the infectious disease, they are referred to as a host, since their body is hosting the pathogenic micro-organism.

Outbreaks of communicable diseases are defined as the sudden occurrences of infections in many people. Outbreaks escalate into epidemics where the outbreak is considered widespread in a

particular community or population. Furthermore, epidemics escalate into pandemics; where the affected community or population spans a large geographical area.

Public health structures exist to monitor and react to epidemics and outbreaks as they occur. At an international level, the World Health Organisation (WHO), the health co-ordinating authority of the United Nations (UN), is tasked with global health issue surveillance and assessment (World Health Organisation, 2010a). Nationally, the *National Institute of Communicable Diseases* (NICD) is the designated public health body that acts in a supportive role to the South African government and SADC (Southern African Development Community) (National Institute of Communicable Diseases, 2010a). It is tasked with research and surveillance on a smaller scale. Furthermore, at a regional level, there exists an organisation for each province that is concerned with communicable diseases. Additionally, within each province there exists an organisation whose boundaries are defined by the municipality. Examples include the *City of Johannesburg Public Health Department* in Gauteng and the *Bojanala district Public Health Department* in the North West Province. Each public health structure makes use of the disaster management life cycle activities as relevant to their area of jurisdiction.

The Council for Scientific and Industrial Research (CSIR) approached the University of Pretoria seeking collaboration in research of communicable disease disaster management. The CSIR is a multi-disciplinary organisation that supports the South African government in matters of scientific and technological research, development and implementation. The CSIR‘s shareholder is the South African Parliament, held in proxy by the Minister of Science and Technology (CSIR, 2010).

The research project outlined in this report was performed in conjunction with the CSIR.

1.1 Problem identification

Generally, there are two forms of control where epidemics of communicable disease are concerned (apart from the inevitable mobilisation of medical treatment): Preventative immunisation and quarantining.

Preventative immunisation, a pro-active control, entails decreasing the chance of an epidemic by decreasing the number of individuals that are prone to infection. This is referred to as raising the herd immunity (Jones and Moon, 1987).

Quarantining, a reactive control, entails isolating the Infected individuals from those who are prone to infection. The aim is to reduce or eliminate the chance of an Infected individual infecting a healthy individual.

Sadly, due to the nature of some diseases and the means of vaccination cultivation, immunisation is not possible for all infectious diseases. Some pathogenic micro-organisms have the ability to mutate. Vaccines however do not, hence one vaccine may not be effective for a disease and keeping up with the mutations is a tedious practice. In some cases, the vaccine cultivation may require humans, animals or plants for cultivation and this may lead to scarce supply of the vaccine.

Even if vaccination were possible for all infectious diseases, there's no guarantee that all individuals within a community would have access to the vaccine.

Equally, quarantining is not always a practical endeavour since it is ineffective without enforceability. Copious amounts of man power and associated resources will be required to enforce the quarantine. Forcing Infected people into isolation can also be seen as an infringement on their human rights, especially if one considers that quarantining is not applicable to all infectious diseases. These controls are therefore not foolproof.

The previous paragraphs attest to the fact that not all infectious diseases outbreaks can be prevented through vaccination and immunisation. Furthermore, quarantining is not an apt way to

respond to an epidemic. Therefore, emphasis should be placed on limiting the spread of outbreaks and preventing the escalation into epidemics since *prevention is better than cure* .

Similarly, *forewarned is forearmed*, early warning or identification of a potential epidemic will result in the latter disaster management activities being initiated earlier. Earlier response will undeniably lead to fewer fatalities and quicker recovery. If one is able to use the surveillance of diseases and their spread patterns as key indicators of an early warning system (forewarned), one will be empowered (forearmed) to initiate the appropriate life-cycle activities earlier.

Three key means of improvement in which the ability of the African continent to limit the spread of outbreaks are listed below.

Need for early detection systems Currently, each country on the African continent has some form of a communicable disease outbreak surveillance system. These systems are retrospective and lack the capacity or capability to act as early warning systems. The WHO established a strategic plan, Epidemic and Pandemic Alert and Response strategic plan (EPAR). The chief objective of EPAR is to strengthen early warning capacity by allowing African nations to further their ability to identify, notify, verify and respond rapidly to communicable diseases (World Health Organisation, 2008). Early detection would lead to early execution of disaster management protocols and would undeniably lead to early disease containment and recovery.

Need to incorporate information technology systems *Knowledge is power.* Information technology systems would facilitate data capturing and sharing. Stakeholders will be empowered by data and information to execute informed decision making and the mobilisation of resources and expertise. Gaps in knowledge will be filled by allowing all stakeholders easy access to information.

Need to improve self sufficiency of African communicable disease intervention and disaster management Africans are inherently equipped with a better understanding of the problems relevant to the continent. For this reason, their solutions will be more practical, effective and culturally acceptable. Similarly, if the solutions are initiated from the home front, the execution and implementation will be less problematic, much quicker and less costly. Currently, many communities throughout the African continent don't have access to public health communicable disease intervention systems, as domestic health resources are insufficient and international assistance or efforts are often misaligned or impractical.

1.2 Solution proposition

This project proposes an early detection system based on the simulation of the reality of infectious disease transmission between humans in a community. The simulation will give insight regarding the spread and escalation of an infectious disease under certain conditions. The system aims to convey target areas for pre-emptive outbreak disaster management activities. Based on the intelligence from the simulation model, disaster management activities can be initiated prior to an outbreak reaching epidemic proportions and focus can be placed on disease containment.

Africa is a vast continent with many communities, each with its own population, and each population with its own set of demographic characteristics. Taking these variables into account, along with the myraid of variables linked to communicable diseases, results in a plethora of variables for consideration. This complexity is aggravated by the vast number of communicable diseases, each with it's own characteristics.

For any system to be of value it should be relevant. Therefore, it is proposed that the system initially be used for individual communities and for particular disease(s) affecting those communities. Johannesburg and its surrounding area have been selected as the region of concern. Data will be obtained from *The City of Johannesburg Public Health Department* and the NICD.

Two diseases, H1N1 (Swine-flu) and measles have been selected as the diseases of concern. Both diseases are caused by viral pathogenic micro-organisms. Viruses are simply chains of nucleic acid surrounded by a protein layer. As such, they are not living organisms and cannot reproduce themselves as they lack the living capacity and related mechanisms required to do so. Therefore, a viral strand can only multiply within a living host and the survival outside a host depends on the disease and the strand itself.

Measles is classified as a re-emergent disease due to the number of historical outbreaks recorded. However, it is also classified as a preventable disease due to the existence of a vaccine. In spite of the vaccine, the number of recorded cases remains unacceptably high. According to the World Health Organisation (2009), there were 281 972 world wide reported cases of measles in 2008. 164 000 of these cases were fatal, the majority of which occurred in children. Thus, a staggering 58% of the reported measles cases were fatal.

According to the NICD, 568 measles cases were confirmed between January 2009 and February 2010 (National Institute of Communicable Diseases, 2010b). The majority of these cases were prevalent in children. An estimated 25% of cases affected children between the ages of 6 and 11 months. Similarly, an estimated 18% of the affected children were between the ages of 1 and 4 years old. Furthermore, of the previously mentioned 568 South African cases, 82 (14%) of these cases were reported in the Gauteng Province. Measles can hence be seen as a relevant and apt choice of infectious disease for study.

H1N1 or Swine-flu is classified as an emergent disease. An emergent disease is described as a novel or new disease since there is an extreme element of the unknown concerning such a disease.

Since the emergence of Swine-flu in 2009, South Africa has had 12 640 recorded cases of the disease, 93 of which were fatal (National Institute of Communicable Diseases, 2010c). The Gauteng province alone accounted for 22 (24%) of the recorded fatal cases. The WHO estimates that, in 2009, a minimum of 213 countries had at least 1 reported case of the disease and there were an estimated 16 713 deaths (African Reginal Office, 2009) .

H1N1 can be considered as a prime example of an emergent communicable disease that quickly reached pandemic proportions. It spread globally, ravaging global populations and leaving many a health system to question its communicable disease response tactics. For this reason, H1N1 is an appropriate infectious disease choice for study.

1.3 Solution methodology

The proposed early detection system will take the form of a software application. Data obtained from *the City of Johannesburg Public Health Department*, will be used to generate behavioural algorithms for each disease. The behavioural algorithms are mathematical models that will be responsible for the transmission of the disease from person-to-person. A simulation will be developed where the initial conditions can be set by the modeller. This will permit model flexibility as the user will be able to augment certain environmental and disease characteristics such as the proportion of Infected individuals in relation to those who are healthy or immune. The effects of the initial condition

variations form part of sensitivity analysis. Sensitivity analysis entails investigating the volatility or variability of the model with respect to changes to the initial conditions.

Lao-tzu, a famed Chinese philosopher once said, *"A journey of a thousand miles begins with one step"*. Similarly, the implementation of the proposed early warning detection system requires the execution of the first task (first step): Literature study.

A literature study entails reviewing existing methods and practices used to solve similar problems or implement similar systems. The insight gained will be used as the basis for the formulation and implementation of the early warning detection system.

The second, Basic model development, is contained in Chapter 3 and entails generic model is development. The development of a generic model requires data gathering and basic model development. Data gathering involves analysing data obtained from the *City of Johannesburg Public Health Department* and substantiating this data through interviews conducted with medical professionals. Basic model development entails the development of generic or basic *AnyLogic* models for each disease. Data obtained in the data gathering phase will be used to develop behavioural algorithms for each disease.

The third step, model experimentation (Chapter 4) entails the design and conduction of experiments of each disease's generic model. Experiments will be conducted with the aim of validating the model and solving the identified problem for each disease.

The fourth and final step, entails the identification of prospective areas for research and improvement documented in Chapter 5.

Chapter 2

Literature Study

Simulation modelling is the practice where real-life scenarios are represented by quantitative relationships and executed within a prescribed environment. The environment must be constrained to allow ease of modelling, however, it must also be well defined and flexible where required to sufficiently represent reality. With regards to communicable disease and the spread of epidemics, two types of modelling will be discussed: Mathematical epidemiological modelling and agent based modelling.

2.1 Mathematical Epidemiology

Epidemiology is defined by MacMahon and Pugh as the study of the distribution and causes of disease frequency in man (Jones and Moon, 1987). Accordingly, mathematical epidemiology is the application of mathematics and related concepts in the study of disease. Daniel Bernoulli is credited with the inception of mathematical epidemiology since his use of mathematical methods regarding smallpox in 1790 (Anderson and May, 1992). Since then, various mathematical and statistical methods have been applied to the study of infectious disease.

Compartmental epidemiology (Brauer, 2002) models are used to group many individuals (possible hosts of an infectious disease) within a population into one of three distinct groups (compartments). Each compartment has one of three states assigned to it and all its occupants. In essence, the host population is divided into three states and the total population is equated to the sum of the individuals within each compartment:

1. Susceptible: The individual has yet to contract the disease and is at risk of contracting the disease.
2. Infected: The individual has contracted the disease and is now referred to as a host. This state accounts for those hosts in the incubation (Infected yet no symptoms evident) and latent (Infected yet not Infectious) periods.
3. Recovered: The individual has recovered from the infectious disease and may in certain cases have permanent or temporary immunity from re-infection.

These three states result in the following compartmental epidemiology infectious disease models, each of which can be applied in a deterministic or stochastic context.

1. SI : Susceptible Infected

- Individuals are in either of the two states. There is no recovery (no cure or treatment) for this infectious disease resulting in no acquired immunity.
- Example : HIV

2. SIS: Susceptible Infected Susceptible
   - Infected individuals recover from the disease however; they are almost immediately susceptible to re-infection. No immunity is acquired.
   - Infectious disease examples: Bronchitis and pneumonia, mostly bacterial infectious diseases.
3. SIR: Susceptible Infected Recovered
   - Infected individuals recover from the disease and acquire life-long (permanent) immunity for the disease.
   - Examples: Measles, H1N1 and chickenpox.
4. SIRS: Susceptible Infected Recovered Susceptible
   - Infected individuals recover from the disease and attain temporary immunity for the disease. Individuals are susceptible again once the immune period lapses.
   - Example: Malaria

Each model is concerned with the time and rate of transfer from one state (compartment) to another. The rate of transfer is expressed as the derivative with respect to time of the number individuals present in each respective compartment.

Furthermore, certain diseases require the inclusion of three more states: Exposed, Infectious and Symptomatic. These states are disease specific and depend on the nature of the disease. These states are the result of the decomposition of the Infected into its relevant stages.

1. Exposed The host has contracted the infectious disease but the latent period has yet to lapse and the host cannot transmit the disease.
2. Infectious The latent period has lapsed and now the host enters a period of communicability where they are able to transmit the disease. The incubation period however has yet to lapse
3. Symptomatic Upon the lapsing of the incubation period, the host becomes Symptomatic where the symptoms of the infectious disease become evident.

H1N1 (Swine flu) and measles are transmitted by viral pathogenic micro-organisms. Human beings gain what is termed acquired immunity once they have overcome a viral infection. This is due to the accumulation of antibodies that will respond to and attack the virus once it is detected. However, it is possible for a virus to mutate or undergo some form of metamorphosis where-by its structure is changed such that it is unrecognisable in its "new" form. Under these circumstances, the acquired antibodies are worthless and the individual is once again vulnerable to infection. In these rare cases the SIRS model would be applicable. Due to the aforementioned acquired immunity, the SIR compartmental epidemiology model will be applied to H1N1 and measles. A deterministic and stochastic SIR model will now be discussed. Both model formulations require a constant total population size (N) (Allen and Burgin, 2000)

2.1.1 Deterministic modelling

Deterministic modelling is accomplished using a system of difference equations (Allen and Burgin, 2000). The term deterministic implies that there is little or no uncertainty regarding the model formulation and the results. Strictly speaking, population proportions are used in deterministic models as opposed to finite population sizes in stochastic models. The population size parameter is described as an innocent parameter in the deterministic model opposed to a vital parameter in the stochastic model. This is due to the fact that the population parameter can be eliminated from the deterministic model by rescaling the state variables. The deterministic model is consequently a qualitative approximation of the stochastic model if and only if the population size is large enough. Furthermore, mathematical studies must be conducted on the stochastic model in order to discern what exactly is a sufficiently large population (Nasell, 2002a).

The primary application of the deterministic model regards the investigation of threshold behaviour (Allen and Burgin, 2000). Threshold behaviour is the extreme behaviour exhibited by an infectious disease in a population in extreme circumstances. Extreme circumstances are the complete infection of a population, or the end of an epidemic where only a certain proportion of the population became Infected (stabilisation). Threshold theory as suggested by McKendrick and Kermak (Anderson and May, 1992), states that the introduction of Infectious individuals into a population of Susceptible individuals will not result in an epidemic unless the number of sSusceptible individuals is above a certain critical level (threshold). The disease specific basic reproductive rate is said to be an indicator of this threshold.

The basic reproductive rate of a pathogenic micro-organism ($\mathscr{R}_0$) is defined as the average number of successful offspring that the organism is able to produce. It is also described as the average number of secondary infections an Infected host is expected to produce (Jones and Moon, 1987). In simple circumstances, $\mathscr{R}_0$ is linearly proportional to the total number of susceptible hosts. $\mathscr{R}_0 = \frac{N}{N_T}$, where N_T is the proportionality constant and N is the total population. However, in pragmatic circumstances, $\mathscr{R}_0$ is most likely to be a non-linear function where $\mathscr{R}_0 = f(N)$

In cases where $\mathscr{R}_0 > 1$, the infectious disease becomes endemic in the population. An endemic infectious disease is one where the overall population consists of Susceptible, Infected and Recovered individuals and sporadic outbreaks and epidemics occur when the appropriate proportions of individuals within the compartments is obtained. Similarly, in cases where $\mathscr{R}_0 \leq 1$,the infectious disease dies out as not enough Infected individuals will be produced in the long term to perpetuate the disease outbreak (Brauer, 2002).

With regards to deterministic mathematical epidemiological modelling, more than one version or model representation exists. However, for the sake of consistency, the model discussed by Allen and Burgin (2000) will be used for both the stochastic and deterministic models .

The following equation parameters are defined:

$S(t) =$ Number of Susceptible individuals at time t
$I(t) =$ Number of Infected individuals at time t
$R(t) =$ Number of Recovered individuals at time t
$N =$ Total population size (Constant population)
$\lambda(t) =$ Force of infection
$\beta\Delta t =$ Number of births or deaths in time period Δt
$\gamma\Delta t =$ Number of individuals that recover in time period Δt
$\alpha =$ Contact rate

SIR discrete time deterministic

Discrete time means that the time intervals are non rational integer values. In other words, there are no fractional time intervals ($t + \Delta t$ cannot be a fraction) since all time intervals are discrete.

$$N = S(t) + I(t) + R(t) \tag{2.1}$$
$$\lambda(t) = \frac{\alpha}{N} I(t) \tag{2.2}$$
$$S(t + \Delta t) = S(t)(1 - \lambda(t)\Delta t) + (N - S(t))\beta\Delta t \tag{2.3}$$
$$I(t + \Delta t) = I(t)(1 - \beta\Delta t - \gamma\Delta t) + (\lambda(t)S(t)\Delta t) \tag{2.4}$$
$$R(t + \Delta t) = R(t)(1 - \beta\Delta t) + \gamma I(t)\Delta t \tag{2.5}$$
$$\mathscr{R}_0 = \frac{\alpha}{\beta + \gamma} \tag{2.6}$$

subject to

$$S(0) > 0 \tag{2.7}$$
$$I(0) > 0 \tag{2.8}$$
$$R(0) \geq 0 \tag{2.9}$$
$$0 < (\beta + \gamma)\Delta t \leq 1 \tag{2.10}$$
$$\alpha \geq 0 \tag{2.11}$$

Equation (2.1) states that the total population is a summation of all the individuals within all three epidemiological compartments. The population size remains constant.

The force of infection, λ(t), denoted by equation (2.2), is the number of contacts that result in infection per Susceptible individual per unit time. In other words, the force of infection denotes the impact an infectious disease has on a population as it is seen as the measure of the strength of the infection. The contact rate, α, represents the number of successful contacts made by an individual during a unit time interval and can be seen as an approximation of the population density. The higher the population density the higher the contact rate and a greater number of successful contacts is expected.

Equation (2.3) is the recursive calculation that computes the number of susceptible individuals in the time interval $t + \Delta t$. It is comprised of two parts, a reduction and increase in the number of susceptible individuals. $S(t)(1 - \lambda(t)\Delta t)$ represents the reduction of Susceptible individuals, Susceptible individuals either die or progress to the next state and are classified as Infected (contract the infection). $(N - S(t))\beta\Delta t$ represents the increase of Susceptible individuals resulting from the birth of a new Susceptible individual. The logic of the increase and decrease in the number of individuals within the Infected and Recovered compartments follows similar logic.

Equation (2.4) is the recursive calculation that computes the number of Infected individuals in a time interval $t + \Delta t$. $I(t)(1 - \beta\Delta t)$ represents the reduction of Infected individuals, Infected individuals either die or recover from the infectious disease. $\lambda(t)S(t)\Delta t$ represents the increase of Infected individuals resulting from the infection of Susceptible individuals.

Similarly, equation (2.5 computes the number of Recovered individuals in a time interval $t + \Delta$t. $R(t)(1 - \beta\Delta t)$ represents a reduction in the number of Recovered individuals due to death while

$\gamma I(t)\Delta t$ represents an increase in the number of Recovered individuals due to iInfected individuals that become Recovered.

Equation (2.6) computes the reproductive rate $\mathscr{R}_0$. The reproductive rate is simply the quotient of the contact rate α and number of individuals that leave the model through death or recovery $(\beta + \gamma)$.

Equations, (2.7) to (2.11) denote the constraints of the deterministic model. Equations (2.7) to (2.9) ensure that all compartment sizes are positive as one cannot have a negative population size. Births (of Susceptibles) and deaths (of any individual) occur alternately. In other words, only one (either birth or death) can occur per time interval (Allen and Burgin, 2000). Accordingly, the summation of the birth and recovery rate must lie between zero and one as indicated by equation (2.10). Lastly, equation (2.11) ensures non negativity of the contact rate α.

2.1.2 Stochastic modelling

Stochastic modelling is performed using Markov chains to model a population process with either a continuous or discrete state space. State spaces represent the various states or conditions according to which individuals may be classified. Discrete state spaces have defined separate states while continuous state spaces are those where classification may occur between states. With respect to compartmental epidemiological models, the states are discrete since an individual may only be Susceptible, Infected or Recovered.

A Markov chain is defined as a stochastic process that obeys the following relation:

$$P(X_{t+1} = i_{t+1}|X_t = i_t, X_{t-1} = i_{t-1}, \ldots, X_0 = i_0) = P(X_{t+1} = i_{t+1}|X_t = i_t) \tag{2.12}$$

Essentially, (2.12) states that the probability of being in state i_{t+1} in time interval $t + 1$, depends only on the state in the previous time interval t irrespective of the states the chain progressed through to reach state i_t and i_{t+1} respectively (Winston, 2004). Equation (2.12) can be rewritten as

$$P(X_{t+1} = j)|X_t = i) = P_{ij} \tag{2.13}$$

Where P_{ij} is the probability of the system being in state j at time $t+1$ (one period or time interval earlier) given that at time t the system was in state i. P_{ij}s are referred to as transition probabilities. P_{ij} is therefore the probability of transitioning (moving) from state i to state j in one time step (Δt).

When more than one time step is considered, $P_{ij}(\mathrm{n})$ is referred to as the n-step transition probability from state i to j. The Chapman Kolomogorov equation is used to compute $P_{ij}(\mathrm{n})$ for any number of time steps where $n = \Delta t$.

$$P_{ij}(n) = \sum_{k=1}^{k=n} P_{ik}(m)P_{kj}(m-n), \qquad \text{where } m \leq n \tag{2.14}$$

(m) denotes m time steps while $(m - n)$ denotes $(m - n)$ time steps.

As previously mentioned, an important difference between the stochastic and deterministic model is they importance of a finite population size in the stochastic model (Castillo-Chavez and Yakubu, 2002). The deterministic model deals only with population proportions. The stochastic model accounts for what is referred to as demographic stochasticity (Nasell, 2002b). Demographic stochasticity refers to the changes in the population dynamics (size and ratio of each compartment) as a consequence of stochastic (random) events within the population.

The stochastic model takes the form of a Markov population (birth-death) process where the transition probabilities (the chance of progressing from one state to another) are modelled as a Poisson process and the time jump between each transition is modelled according to an exponential distribution (Nasell, 2002a). Such models are preferred for the SIR disease model as deterministic models have been found to yield qualitatively incorrect conclusions (Brauer, 2002). Due to dependence on the random variables I(t) and R(t), the stochastic SIR model is a bivariate process (one that depends on two variables) (Allen and Burgin, 2000). The subsequent model used is that of Allen and Burgin (2000).

SIR Markov chain with discrete time finite state space

The resulting SIR Markov chain infectious disease model is modelled in discrete time with an finite state space (only 3 states are present). The following must be taken into account:

1. At most one event (one birth, death, infection or recovery) occurs in each Δt.
2. The population size is constant and known with certainty. In other words, each death is accompanied by a birth. For example, a death of a Recovered individual is followed by the birth of a Susceptible.

The following equation parameters are defined:

$S(t) =$ s,where s is the number of Susceptible individuals at time t
$I(t) =$ i, where i is the number of Infected individuals at time t
$R(t) =$ r, where r is the number of Recovered individuals at time t
$N =$ Total population size (Constant population)
$\lambda(i) =$ Force of infection
$P_{ir}(t) =$ Joint probability function
$\pi_i \Delta t =$ Probability of newly Infected individual in time Δt
$(\beta + \gamma) i \Delta t =$ Probability of death or recovery in time Δt
$\gamma_i \Delta t =$ Probability of recovery of an Infected in time Δt
$\beta_i \Delta t =$ Probability of death of an Infected in time Δt
$\beta \Delta t =$ Number of births or deaths in time period Δt
$\gamma \Delta t =$ Number of individuals that recover in time period Δt
$\beta_r \Delta t =$ Probability of death of a Recovered in time Δt

$$N = S(t) + I(t) + R(t) \tag{2.15}$$

$$\lambda(i) = \frac{\alpha i}{N} \tag{2.16}$$

Transition Probabilities

$$P_{ir} = P(I(t) = i, R(t) = r) \tag{2.17}$$
$$\pi_{ir}\Delta t = \lambda(i)(N - i - r)\Delta t \tag{2.18}$$
$$P(I(t + \Delta t) = i + 1, R(t + \Delta t) = r | I(t) = i, R(t) = r) = \pi_{ir}\Delta t \tag{2.19}$$
$$P(I(t + \Delta t) = i - 1, R(t + \Delta t) = r + 1 | I(t) = i, R(t) = r) = \gamma_i \Delta t \tag{2.20}$$
$$P(I(t + \Delta t) = i - 1, R(t + \Delta t) = r | I(t) = i, R(t) = r) = \beta_i \Delta t \tag{2.21}$$
$$P(I(t + \Delta t) = i, R(t + \Delta t) = r - 1 | I(t) = i, R(t) = r) = \beta_r \Delta t \tag{2.22}$$
$$\tag{2.23}$$

Chapman Kolmogorov equation

$$\begin{aligned} P_{ir}(t + \Delta t) = {} & P_{i-1,r}(t)\pi_{i-1,r}\Delta t + P_{i+1,r-1}(t)\gamma\Delta t(i + 1) \\ & + P_{i+1,r}(t)\beta\Delta t(i + 1) + P_{i,r+1}(t)\beta\Delta t(r + 1) \\ & + P_{ir}(t)[1 - \pi_{ir}\Delta t - \gamma i\Delta t - \beta(i + r)\Delta t] \end{aligned} \tag{2.24}$$

Absorptive state

$$P_{0r}(t + \Delta t) = P_{0r}(t) \tag{2.25}$$

subject to

$$0 \leq i + r \leq N \qquad \forall i, r = 0, 1, 2 \ldots N \tag{2.26}$$
$$\pi_{ir}\Delta t + \gamma i\Delta t + \beta(i + r)\Delta t \leq 1 \qquad \forall i + r = 0, 1, 2 \ldots N \tag{2.27}$$
$$S(0) > 0 \tag{2.28}$$
$$I(0) > 0 \tag{2.29}$$
$$R(0) \geq 0 \tag{2.30}$$
$$0 < (\beta + \lambda) \leq 1 \tag{2.31}$$
$$\alpha \geq 0 \tag{2.32}$$
$$\tag{2.33}$$

Equations (2.15) to (2.16) and (2.28) to (**??**) are applicable to both the deterministic and stochastic models. However, the force of infection considered in equation (2.16) incorporates the number of Infected individuals into its formulation. This incorporation is necessary as the force of infection will fluctuate as the population of Infected individuals fluctuates.

Equations (2.17) to (2.23) represent variations of the joint probability function P_{ir}. P_{ir} in turn denotes the transition probabilities (see equation (2.13)) between states i (Infected) and (Recovered) respectively. Furthermore, these equations, (2.17 to (2.23 result in the computation of the Chapman Kolmogorov equation for the bivariate Markov chain.

Equation (2.24) represents the Markov population (birth-death) process. It consists of five cases:

1. An Infected individual recovers and a Recovered individual dies, $(i - 1, r)$.
2. A Susceptible individual becomes Infected and a Recovered individual dies, $(i + 1, r - 1)$.
3. A Susceptible individual becomes Infected and a new Susceptible individual is born, $(i + 1, r)$.

4. A Susceptible individual becomes Infected and an Infected individual recovers, $(i, r+1)$.

5. No birth, death, infection or recovery occurs, (i, r).

Equation (2.25) represents the absorbing state at the origin. The absorbing state is one where progression is not possible (Winston, 2004). For $P_{0r}(t+\Delta t)$, $I(t)=0$ and $R(t)=0$, therefore, one cannot progress beyond this state as one needs Infected individuals to generate Recovered individuals and Infected individuals are required to infect Susceptible individuals

Equations (2.26) and (2.27), ensure that the transition probabilities are positive and bounded by one.

2.2 Agent based modelling

An agent based model is defined as a system represented as a collection of autonomous decision making entities (agents) (Bonabeau, 2002). Agent Based Modelling (ABM) is therefore a tool used to study complex systems and is used to convey how macro phenomena emerge from micro level behaviour between agents in a heterogeneous (consisting of elements that vary in nature) environment. Simply put, an ABM enables one to observe how the interactions between individual lower level actions (micro phenomena) cause higher level reactions (macro phenomena).

According to Janssen (2005), AGMs consist of two main components

Cellular automation entails decomposing a complex system (cell lattice) into individual cells. Cells are represented by states, the complexity and number of which will depend on the system under study. Each cell has the ability to change its state according to defined transition rules which determine a cell‘s state with respect to time. Transition rules may affect the cell lattice on a “global” (system wide) scale or on a local neighbourhood scale. The main shortcoming of cellular automata is the limiting nature of the cellular states. Complex states are difficult to define and consequently will not be appropriately represented.

Agents are described as autonomous adaptive entities within an environment. Their adaptive (flexible) nature gives rise to the following characteristics:

- Agents are goal oriented and behave in such a manner to minimize or maximise some or other utility (value or convenience).
- Agents are reactive as they respond to environmental changes and they have the capacity to interact with other agents.
- Agents are autonomous. Autonomy is the capacity agents have to make decisions independent of human intervention. In other words, once the simulation is executed, the user doesn't affect the decisions of the agents.

Agents derive information from the environment that defines the perception they have about the state of the environment. Based on the goals and attributes an agent possesses, it makes decisions on actions to perform and these actions in turn affect the environment. The agents can interact indirectly, for example by affecting the common resource, or directly by communication.

Social studies have shown that humans exhibit a combination of reactive and goal orientated behaviour (Bonabeau, 2002). According to Ball (2003), in order to gain understanding or insight into society, it (society) must be decomposed into its constituent parts. The individual function of each constituent must be understood and then one can observe how they interact together to

compose the whole (society). It is for this reason that agents are deemed appropriate representatives of humans in complex systems. An agent assesses its situation and makes its decision based on a set of rules that define appropriate behaviour pertaining to the situation. Humans (ignoring the emotional capacity and so called "grey" areas) behave in a similar manner. In cases where state changes are applicable, the cellular automation aspect comes into play.

2.3 Conceptual Design-Modelling of epidemics using ABM

ABM, as previously discussed is a tool used to model macroscopic phenomena from the bottom up. In the case of communicable diseases, epidemics can be described as macroscopic phenomena that are a consequence of many singular infections (microscopic phenomena).

The behavioural rules that govern an agent's behaviour will effectuate the transmission of the infectious disease from person-to-person. This behaviour will be modelled by means of mathematical epidemiology using data from *the City of Johannesburg Public Health Department* and the NICD. Stochastic mathematical modelling in the form of a bivariate Markov chain, will be used to define the behavioural rules (algorithms) as opposed to deterministic modelling in the form of difference equations. An agent's behaviour is described as stochastic due to the element of decision making. Randomness and probabilistic mathematical formulations would be more inclined to represent this element and reality as opposed to aggregate and deterministic formulations. For this reason, the stochastic SIR mathematical epidemiological model will be applied as the rules of the agent's behaviour. Sensitivity analysis is the consequence of testing hypotheses by altering the behavioural rules and the agent's attributes. The effect of these changes on the macro phenomenon (epidemic) can be observed.

The complexity of the interactions that result in transmission will be more effectively modelled from an agent based perspective. This statement is further endorsed by Bonabeau (2002), who surmised that the nature of the repetitive interactions between agents requires the capability of computers to simulate the dynamics out of reach of pure mathematical methods.

Accordingly, it can be concluded that an effective simulation model of epidemics and the associated transmission of person-to-person infectious diseases is a combined mathematical epidemiological and agent based model. It is expected to comprehensively simulate reality and the numeric results will be appropriately displayed through animation and apt charting.

In conclusion, early outbreak detection will lead to early infectious disease containment effectively limiting the infectious disease‘s capability to escalate into an epidemic. By simulating the reality of infectious disease outbreaks, one is able to gain insight and knowledge which will be beneficial in the quest for disease containment. Through a literature study, it was established that the most effective means of simulating the spread (transmission) of infectious disease in a community or population is a combined mathematical epidemiological and agent based model.

Chapter 3

Basic Model Development

Model formulation or development is the stage where the conceptual model discussed in Chapter 2 is translated into an actual model. Translation requires the gathering of data from suitable sources and the incorporation of said data into a combined compartmental epidemiological SIR and agent based model. Epidemiological modelling has been described as more art than science as its effectiveness depends greatly on the discretion of the modeller (Hethcote, 2008). The modeller must make apt choices to ensure that the model formulation itself is as simple as possible yet; steps must be taken such that it is adequate for the problem being considered.

3.1 Data gathering

In order for the behavioural algorithm to be both relevant and effective, is gathered. According to Magnus (2007), one needs to take advantage of interview and medical record data since the combination will ensure the validity of model or investigative results.

By substantiating quantitative data with qualitative experience, one is able to accumulate constructive and valuable information that is pivotal in the development of a meaningful model. As such, medical record data was obtained from the *City of Johannesburg Public Health Department* (from here on referred to as the *CoJ*) and interviews were conducted with medical personnel.

It is a well known fact that data is readily available for naturally occurring epidemics and the associated re-emergent infectious disease. For this reason, the measles data obtained is more complete and comprehensive than the data obtained for H1N1. However, Hethcote (2008) stated that data is often incomplete due to under reporting or inaccurate record keeping. The data obtained from the *CoJ* is no exception to this statement. Incomplete data will inadvertently lead to an unreliable model based on incorrect or distorted parameter estimation. This reinforces the argument for the inclusion of the insights and reflections of medical personnel.

The concerned personnel selected have firsthand experience in dealing with communicable disease cases and consequent outbreak escalations. Thus, they are referred to as front-end medical personnel since they operate on the front-line

- Dr A.T.K. Ndiwalana: General Practitioner;
- Dr D.B.P. Ndiwalana: General Practitioner;
- Mrs N.M. Modibedi: Head Nurse and Manager of the Department of Infection Control at the Job Shimankana Tabana Public Hospital in Rustenburg in the North West Province (Bojanala district). Infection control within a hospital is a crucial element of public health infrastructure that is concerned with the identification, surveillance and control of communicable disease.

CoJ data is an aggregate of cases reported from various facilities or medical institutions throughout the *CoJ* municipal area. Case information is collected and used to produce line listings. Line listing is the process of organising or orientating the data according to time, place and person (Nelson et al., 2004). Organising data in this manner allows for easier reviewing and categorising. Line lists are disease specific and consequently, measles (see table 3.1) and H1N1 (see table 3.2) have their own format and inclusions.

3.1.1 Measles

The acquired measles data from the *CoJ* is in the form of a Microsoft Excel workbook with four distinct worksheets. Each worksheet accounts for one of four classifications of measles case:

1. Measles suspected;
2. Measles confirmed;
3. Measles negative;
4. Measles deaths.

The measles line listing form looks as follows:

Table 3.1: Measles format of line listing form

Case no.	Facility	I=In-patient O=Out-patient	Address	Sex	Age	Date seen at facility	Date of onset of disease	Fever	Rash	Cough, coryza or conjunctivitis	No. of measles vaccines recieved	Date of last measles vaccine	Source of vaccination status	Date lab specimen taken	Results of lab testing	Outcome A=Alive D=Dead	Comments
01c0j01																	

Data is collected concerning the patients and infection's particulars. Patient particular information includes the patient's gender or sex, address and age. The infection or measles' particulars include a case number, the patient's symptoms, and the facility which they reported to. The case number is the unique identification number given to each measles case. The facility is classified as the medical institution such as a hospital or clinic where the patient was seen and recorded. Auxiliary facility information includes the date when the patient was seen and the date when the lab specimen was taken for testing. The results, measles negative or positive, are recorded once they are received. Vaccination particulars are also recorded concerning the last date of measles vaccination and an indication of how the vaccination status was confirmed (verbally, or was a vaccination card seen?).

Symptomatic information pertains to the symptoms that the patient displays upon arrival at the facility. The indicator symptoms are fever, maculopapular rash (rash that is the combination of small flat discoloured spots on the surface of the skin-macules, and small raised bumps-papules) cough,

coryza (inflammation or swelling of nasal mucous membranes) and conjunctivitis (inflammation of the outer most layer of the eye, the conjunctiva, whereby the eye appears red or pink in colour) Benenson (1985).

In-patient or Out-patient records whether or not the patient was admitted into the facility or whether they were sent home. Admittance depends on the severity or degree of infection. Follow up information is recorded at a later date and is concerned with the status of the patient (alive or dead) and associated comments.

As per the Measles suspected worksheet, between October 2009 and June 2010, there were 3607 suspected cases of measles recorded. A suspected measles case presents itself in an individual that experiences the symptoms of measles. The case is suspected until laboratory testing is performed to verify the individuals infection status. Urine or blood and in some cases a throat swab are sent to the NICD for laboratory testing.

Of the aforementioned 3 606 suspected cases, 921 were verified as positive measles cases while only 94 cases were verified as negative measles cases. Furthermore, 15 deaths were recorded. Of the recorded deaths, six correspond to confirmed cases and the remaining seven to suspected measles cases.

Age groupings

Its no secret that those with weaker constitutions or immune systems are more vulnerable to infection. However, measles is classified as an infectious disease whose primary victims are those aged eight months to five years (Kassner, 1985). As such, it is appropriate that age is recorded with each measles case. Children younger than one year have their age recorded as a fraction of 12 months. For example, a child who is three months old will have their age recorded as $\frac{3}{12}$ months which expands into 0.25 years.

According to *CoJ* data, 378 confirmed measles cases were in children younger than 1 year. Furthermore, 523 children aged five years and younger tested positive for measles accounting for 57% of confirmed measles cases. (there were nine cases where age was not recorded) With respect to measles deaths, 50% of the confirmed measles deaths occurred in children younger than one year.

Latency and Incubation periods

Latency is the time from infection (disease contraction) to Infectious (ability to infect others) and incubation is the time from infection to symptoms. The incubation period for measles lasts between 10 and 14 days while individuals take at most 72 hours (three days) to become Infectious. Therefore, the compartmental state Infected is decomposed into three states, namely Exposed, Infectious and Symptomatic. The SIR model is hence transformed into a SEISR model. Figure 3.1 depicts the transformation.

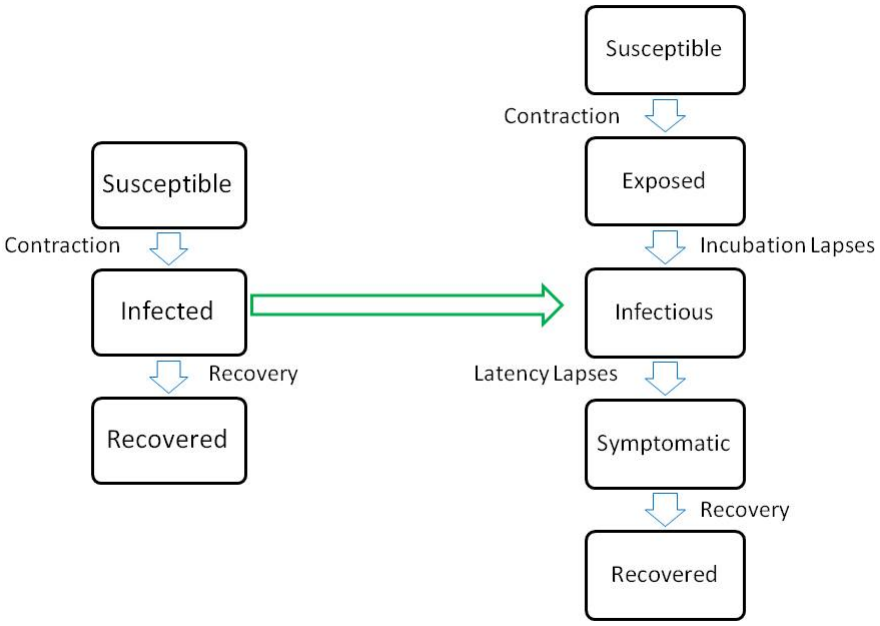

Figure 3.1: Diagram depicting SIR transform to SEISR

It is important to note that it is possible for one to become Infectious without displaying symptoms of infection. This has grave consequences on the ability for the disease to be transmitted within a population as one may be ignorant of the fact that they are potentially spreading the disease. Measles however is a disease where infectiousness transcends the Infectious period as one may still be Infectious even when Symptomatic. Hethcote (2008) argues that assuming constant infectivity for measles is valid due to the nature of the disease and the shortness of the Infectious period. In reality, it is possible for the degree of infectiousness (also referred to as infectivity) of an individual to fluctuate with respect to time. However, measles infectivity is assumed to be constant.

CoJ data gave no bearing on incubation and latency periods and information was gleaned through literature, (Kassner, 1985) and interviews.

The Markov Chain discussed in Chapter 2 can be said to be probabilistic since it requires rates and probabilities to subsequently calculate future rates and probabilities. Thus, allowances need to be made in model development to ensure that the Markov Chain and time based transitions function in harmony.

In-patient vs. Out-patient

According to Modibedi, measles is an infectious disease whose treatment is of a symptomatic nature. In other words, measles itself is incurable; one can only treat the symptoms until the disease itself runs its course. That said, the chances of recovery without symptomatic treatment or positive reinforcement treatment to ones immune system is very low. As such, out-patients are prescribed the necessary symptomatic or immune reinforcement medication and follow up appointments are made. Out-patients typically have uncomplicated measles and recovery (phasing out of symptoms) is expected in 2-3 days if the medication is taken and the treatment plan is followed. Conversely, in-patients are afflicted with measles that is complicated by an opportunistic disease or infection

such as pneumonia. Other In-patients include those who have severely compromised or weakened immune systems such as infants and those suffering from chronic ailments such as heart or lung illnesses.

Of the 921 confirmed measles cases, 149 were in patient cases and 80 (54%) were found in children younger than five years. Additionally, two in-patients died, both of whom are children younger than five years.

Figure 3.2 depicts the inclusion of the In-Patient and Out-Patient recovery times. This incorporation is expected to result in a more realistic model.

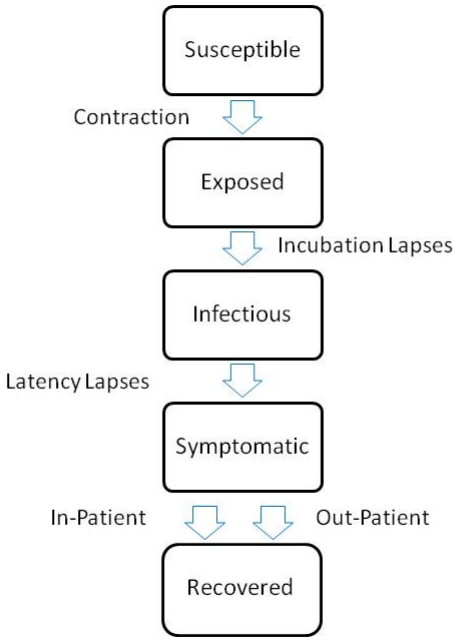

Figure 3.2: Diagram depicting SEISR with In-Patient and Out-Patient inclusions

3.1.2 H1N1

The acquired H1N1 data is in the form of a Microsoft Excel spreadsheet consisting of a single workbook. The line listing form is illustrated in Table 3.2.

Table 3.2: H1N1 format of line listing form

No.	Date of onset	School	Travelling History	Tests	Symptoms	Date seen at facility	Treatment	Results	Separations	District	Hospital	Nationality	Age	Gender	Address	Contacts	Comments
1																	

Again, data is collected concerning patient and infection particulars. Patient particulars include

the patient's age, gender and address. Auxiliary information pertaining to the patient's nationality, travelling history, and school are also recorded.

Infection particulars include the date of H1N1 onset, the symptoms present at the time of recording and the hospital where the patient was seen. H1N1 presents with a combination of generic (applicable to all types of influenza) and H1N1 specific symptoms (Kassner, 1985). Generic symptoms include fever, sore throat, cough, coryza, nasal congestion and head and body aches. H1N1 specific symptoms are diarrhoea and vomitting.

Information pertaining to lab testing of each patient is also recorded. Treatment denotes the medication or treatment plan undertaken to combat positive H1N1 cases. Contacts are the people that the Infected individual has come into contact with and must be tested for H1N1.

Types of influenza

There are three types of influenza: A, B and C. Influenza A and B are found to be the two that infect humans. Due to the nature of the influenza virus and its ability to mutate, the naming convention is based on the presence of certain proteins and the respective codes or sequences. Influenza A is associated with widespread epidemics and pandemics. Influenza B is associated with localised epidemics and type C influenza is associated with sporadic cases and minor localised outbreaks (Benenson, 1985).

Influenza A has many subtypes, one of which is H1N1. According to the World Health Organisation (2010b), the particular strain of H1N1 that emerged in 2009 and caused a pandemic, is seen as an evolved Influenza A strain that combined genes and proteins from various sources. Sources include human, pig (swine), and avian (bird). This corresponds with a statement made by Nelson et al. (2004) where the virus's ability to mutate is affirmed as the key to its ability to cause annual epidemics and periodic pandemics since susceptibility becomes universal.

Data incompleteness

The *CoJ* H1N1 received data is classified as incomplete. An indication of the incompleteness of the H1N1 *CoJ* data is the fact that only positive cases are recorded on the received line listing form. This is not a true reflection of reality since it is not uncommon for individuals that present with H1N1 symptoms to test negative for the disease. This is due to the symptom similarities between H1N1 and regular flu or the common cold.

Furthermore, of the 520 cases recorded on the *CoJ* line list form, 427 have "Unknown" travelling history recorded. Similarly, 458 were recorded as receiving Unknown treatment while a staggering 397 of these patients received their "Unknown" treatment at an "Unknown" hospital. Through sporadic comments, some light is shed onto the well-being of the patients at a later date. However, theres no explicit capturing of the deaths or status of recovery of most patients.

CoJ data gave no bearing on the outcome of the patient. In other words, no death or recovery statistics for the number of the confirmed H1N1 cases were recorded on the line listing form. However, the NICD has death statistics for H1N1. According to the NICD, during the peak (20 July 2009 20 September 2009) of the H1N1 pandemic, 93 deaths resulted from 12 447 confirmed cases.

The lack of regional H1N1 (*CoJ*) death statistics the existence of national H1N1 (NICD) speaks volumes and further endorses the statement that the data collected is incomplete.

School and Travel

Nelson et al. (2004) states that families with school aged children have the highest rates of infection. Schools provide a mechanism that enables the uncomplicated transmission of infectious disease.

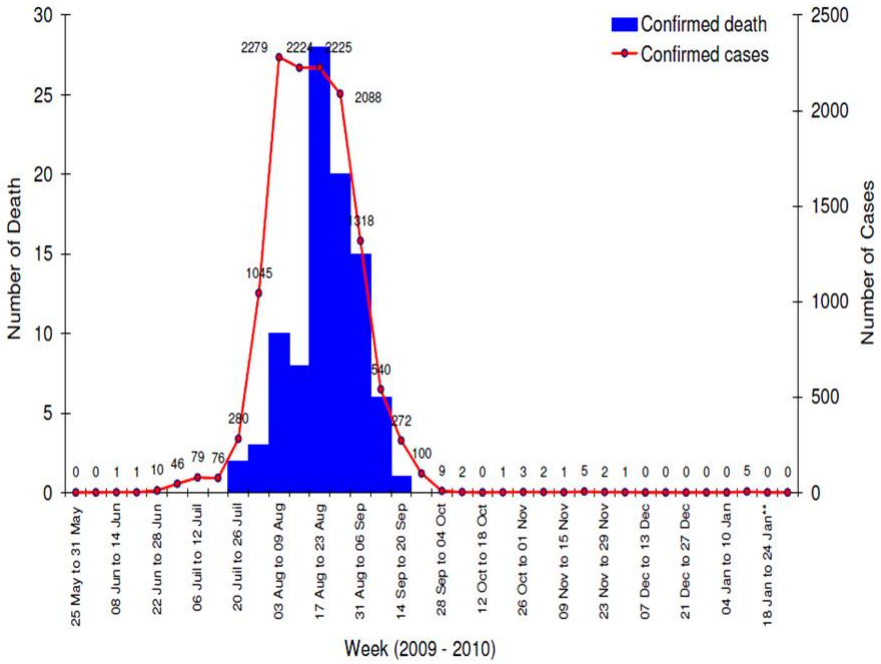

Figure 3.3: Diagram depicting National Institute of Communicable Diseases (2010c) national H1N1 data

Simply put, students are in contact with their fellow students as well as their families and friends outside of school. Furthermore, teachers and teaching staff are contact with a plethora of students on a daily basis as well as their own families, peers and seemingly inconsequential encounters with strangers outside of school. If one considers the that H1N1 can be confused with regular flu due to the similar symptoms, this further strengthens the possibility of an H1N1 Infected student infecting their peers. Students and teachers will continue to attend school (unless complications arise) even when Infected with H1N1 as one would normally expect of anyone with the common cold or regular flu. *CoJ* data conveys that 71 Students were confirmed to have H1N1. In the context of the *CoJ* data, students comprise of primary, secondary (high school) and tertiary (university) level students.

Influenza is described as a seasonal infectious disease as the majority of outbreaks and epidemics are recorded in winter months. However, the H1N1 virus behaved atypically and caused high numbers of summer infections.

The speed and ease of travel are cited by Nelson et al. (2004) as mechanisms that facilitate infectious disease transmission that leads to epidemics and pandemics. An Infected individual upon a domestic or international flight exposes the travellers and flight attendants to the infectious disease. Travellers will interact with people upon arrival to their destination and they in turn may expose anyone they come into contact with. Similarly, flight attendants will now perpetuate the probable cycle of infection by interacting with a different set of travellers on another flight. The same logic can be applied to all modes of public transport. International travelling will result in global spreading of the infectious disease. Consider the case where an international traveller is a student. Both the school and the travelling mechanism are now probable.

CoJ data conveys that 94 patients had a confirmed history of travel. Modibedi describes a recent history of travel as a key indicator of the possibility that regular flu symptoms may in fact be those of H1N1. According to Modibedi and corroborated by Dr Ndiwalana, the majority of H1N1

cases involved an individual who either travelled internationally or came into contact with someone who travelled internationally.

Latency and Incubation periods

H1N1, like most other types of influenza has a short incubation period of 1 to 3 days (24 to 72 hours). Unlike measles however, the incubation period and latent period happen in tandem. In other words, one is Symptomatic and Infectious simultaneously. Therefore, in the case of H1N1, the SIR model transforms into a SEIR model where the Infected state incorporates both Infectious and Symptomatic. Figure 3.4 indicates this transformation

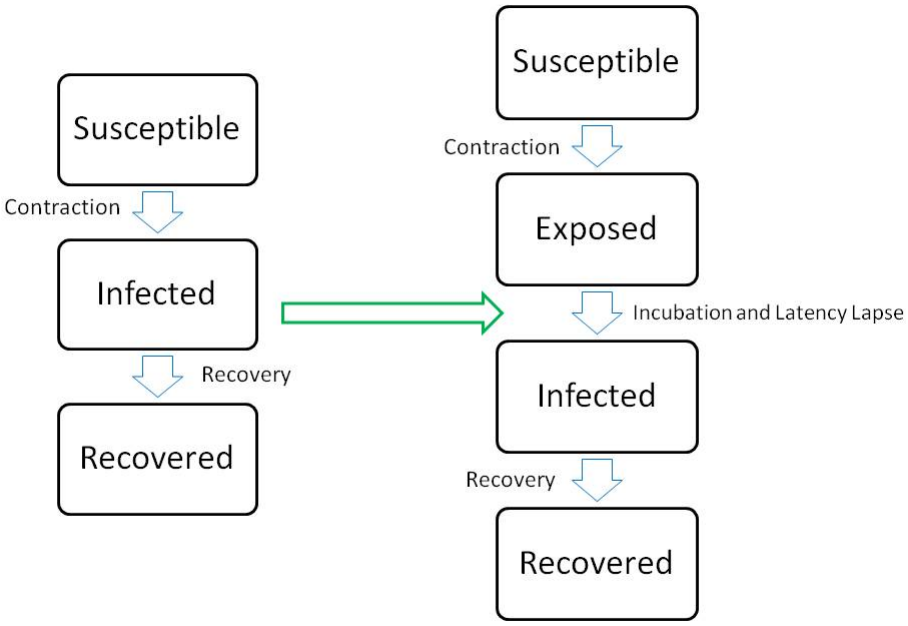

Figure 3.4: Diagram depicting SIR transform to SEIR

Treatment and recovery

Once positive confirmation of H1N1 is received, Tamiflu is prescribed. Tamiflu is an orally (pill or capsule form) administered form medication for the treatment of H1N1. Tamiflu must be administered taken within 48 hours (two days) after an individual becomes Symptomatic for effective response and recovery from the actual influenza takes between 36 and 48 hours.

CoJ data shows that Tamiflu was administered to 35 patients while 458 patients had their treatment recorded as "Unknown"; 29 of the remaining patients were recorded as receiving no treatment ("Nil") and two patients were prescribed symptomatic medication consisting of anti-congestants and cough mixture.

3.2 Generic model development

Initially, a generic or non-specific model was developed. The purpose of this model is to develop the skeleton of the final model. The generic model will be expanded and augmented as necessary for circumstantial sensitivity or scenario analysis.

3.2.1 State chart development

A generic state chart named Illness is defined for the SEISR compartmental model. Due to the replication nature of *AnyLogic*, the state chart need only be defined once and it will be replicated to exist for all agents.

A trigger is that which initiates the transition from one state to another. Subsequently, transitions are based on a particular trigger, of which there are five types.

1. Rate: Occurs according to a defined rate.
2. Timeout: Occurs after a defined time period lapses.
3. Message: Occurs upon the reception of a particular message.
4. Arrival: Occurs upon the arrival of a specific agent
5. Condition: Occurs according to a certain condition holding or being true

The generic model makes use of three of the five types of triggers, namely message, timeout and rate. Furthermore, there are three distinct transition categories.

Message based transition: Susceptible to Exposed A message based transition is a consequence of agents communicating and passing messages between each other. A Susceptible individual can only become Exposed to the infectious disease through contact with an Infectious individual. Accordingly, the *AnyLogic* trigger is hence the message "Exposed" that is passed from an Infectious agent to a Susceptible agent.

Timeout based transitions: Exposed to Infectious, Infectious to Symptomatic, Symptomatic to Recovered The lapsing of the latent period is responsible for the state transition from Exposed to Infectious while, the lapsing of the incubation period is responsible for the state transition from Infectious to Symptomatic. Lastly, recovery time is accountable for the transition between Symptomatic and Recovered. All three timeout transitions take the form of uniform distributions that utilise the minimum and maximum times obtained through literature and interviews in the data gathering phase.

Measles

- Exposed to Infectious: one to three days
- Infectious to Symptomatic: nine to twelve days

H1N1

- Exposed to Infected: one to three days

Time opposed to Rates

As previously mentioned, the Markov Chain requires probabilities and rates in order to function. Unfortunately, data obtained from the *CoJ* didn't facilitate the use of the SIR model in its current form. Probabilistic data is required for the transition Infected to Recovered and the data received didnt allow for extrapolation or calculation of appropriate recovery probabilities.

In order to calculate the probability of recovery, one requires the number of people that recover in a time period in relation to those that don't recover. However, *CoJ* data merely recorded whether the the individual was alive or dead. No record was kept of the date upon which this "status" was discerned. Even if one were to assume that the individuals classified as living have Recovered from the infectious disease, the time between the infectious diseases's onset and the individual's final status is not recored. In essence, data noted whether or not an individual Recovered but not how long it took them to recover.

For this reason, time-based transitions will be used opposed to rates or probabilities. Time-based state transitions are documented times that an infectious disease takes to progress from one state to another. Furthermore, all of the consulted front-end medical personnel reinforce the inclusion of time based transitions since they are considered documented evidence per disease.

Internal transitions: Infectious and Symptomatic Within the Infectious and where applicable the Symptomatic states, an internal transition occurs. These transitions enable Infectious and Symptomatic agents to infect Susceptible agents. This transition is cyclic since it recurs in each agent that is in the specified state. The Markov Chain discussed in Chapter 2 is responsible for this transition. The Markov Chain returns a value, probability of new Infected individual ($\pi_i \Delta t$) which functions as a rate trigger.

It must be noted that in order for the message based transition to occur, the internal transition must occur first. This means that there must be at least one agent that is Exposed for state transitions to occur. This of course makes perfect sense since all epidemics must start with an initial outbreak or initial case of infection, commonly referred to as case or person zero. More light will now be shed on the Markov Chain.

3.2.2 Markov Chain

Within the Illness state chart (see state chart development), only the internal transition is controlled by the Markov Chain since the remainder are triggered by timeouts. The internal transition is hence controlled by the Probability of New Infected, $\pi_i \Delta t$, aspect of the Markov Chain.

As previously discussed in Chapter 2, the probability of a new Infected individual, equation 2.18, depends on the force of infection and the number of Susceptible individuals. The force of infection, equation 2.16, depends on the contact rate, α, and the relation of Infected individuals with the total population. The contact rate and total population size are defined by the modeller to ensure that the generic model is able to conform to desired specifications resulting in an *exact or precise* scenario.

$$\lambda(i) = \frac{\alpha i}{N} \tag{3.1}$$

$$\pi_{ir} \Delta t = \lambda(i)(N - i - r)\Delta t \tag{3.2}$$

However, due to the decomposition of the Infected state into three states (Exposed, Infectious and Symptomatic) the Markov Chain needs to be augmented accordingly. The total population, is

now a summation of 5 states (SEISR) or 4 states (SEIR) opposed to the originally defined 3 states (SIR).

The Markov Chain is executed once every time interval or once every day. Consequently, the force of infection and the probability of a new Infected are calculated daily. This is an apt recursion since the number of occupants within each compartment changes with respect to time and the time interval is defined as one day.

3.2.3 Environment

The environment of an *AnyLogic* ABM is the interface where the agents interact with each other. Technically speaking, it's a construct that is used to define the properties common to a group of agents. Properties include the space, layout and network type. The space type refers to the nature of the environment. The environment could be continuous or discrete. A continuous environment is one where events can occur at any point in time while a discrete environment is one where events occur at specific defined intervals.

The layout refers to the manner in which the agents are arranged upon the simulation's execution. There are three layout options: random, user defined and arranged. Lastly, the network type refers to the manner in which the agents are connected to each other. There are six network alternatives: random, user defined, small world, ring lattice, scale free and distance-based.

The generic model has both the layout set as random while the network is distance based. Furthermore, the environment is set as continuous and 2-dimensional. A distance based network means that agents are connected to other agents that are a defined distance away from them. The implications of a distance-based network and the continuous environment will be elaborated on within the Mobility section.

3.2.4 Mobility

Mobility is the ability for the agents to move and be mobile within their environment. Mobility has implications on the spread of infectious disease as the more mobile an Infected individual is, the greater the number of Susceptible people they encounter and can hence infect. The *AnyLogic* environment must be continuous to facilitate agent movement.

The generic model agent movement can be defined as aimless or directionless as the agents are programmed to wander or meander throughout the environment hence the continuous environment and the random agent arrangement. Furthermore, the distance-based network ensures that agents become connected to other agents near them during their meandering.

Essentially, each agent is programmed to have a connection range (radius) of 100. Whenever an agent enters the connection range of another agent, the agents are connected. If one of those agents happens to be in the Infected state and the other in the Susceptible state, the Infected agent may infect the Susceptible agent according to the probability of infection that exists in that instant.

3.2.5 Birth-Death events

In Chapter 2, it was established that the stochastic Markov Chain model takes population demographic information into account in it's formulation. Therefore, the following statements were made:

1. At most one event (one birth, death, infection or recovery) occurs in each Δt.

2. The population size is constant and known with certainty. In other words, each death is accompanied by a birth. For example, a death of a Recovered individual is followed by the birth of a Susceptible.

Two cyclic events, Birth and Death, are defined. They seemingly occur in tandem, however, they actually occur in series, one after the other. The intervals between each execution and subsequent birth or death are small and thus inconsequential to ensure that the total population size remains constant with 500 agents.

The Birth event only gives rise to Susceptible individuals while the Death event results in the death of an individual in any state. Deaths of Susceptible, Infected and Recovered individuals are described as collectively exhaustive since all death alternatives are considered. This means that all the states in which an individual may die are considered and the sum of all death probabilities equals 1.

CoJ and NICD data only sheds light on the likelihood on the death of confirmed patients. No light is shed on the likelihood of a Recovered or Susceptible individuals death. For this reason, the probability of an Infected individual dying will be that calculated from the obtained data. Since no data exists for the death of Susceptible or Recovered individuals, their death probabilities will be assumed to be an equal proportion of the remaining probability.

A random number ($0 <$ random number < 1) is generated with each execution of the Death event. This is the mechanism that controls which particular state will be facing a death during any instant. Conditions are in place that determine which state loses a member according to a defined probability.

3.2.6 Creation of Exposed individuals

In order for the Markov Chain to function, there must be at least one Exposed individual within the population. Thus, upon execution of the simulation, at least one Exposed individual is created at start up. However, due to the inclusion of time based transitions, one Exposed individual is no longer sufficient.

Initially, when only one Exposed individual existed, the probability of a new Infected individual was too low to facilitate disease spread. In other words, the time transitions would lapse and the individual would recover before infecting another individual. This results in the existence of no Exposed individuals within the population. The simulation would hence enter an absorptive state since it cannot progress beyond this point.

For this reason, experiments must be conducted to investigate the number of Exposed individuals required to ensure that the absorptive state is avoided. Experiments will be explained and conducted in Chapter 4, Model experimentation.

The internal transition is responsible for an Infected individual infecting a Susceptible individual. The H1N1 state chart has two states (Infectious and Symptomatic) that exhibit this type of transition while the measles state chart has only one state (Infected) that exhibits the internal transition.

Chapter 4

Model experimentation

A model, in simplistic terms is described as an approximation of reality. A model is never 100% realistic thus, experiments must be conducted to validate the model. According to Eriksson et al. (2000) there are three experimental objectives.

Screening Entails determining which factors within the system or model under study are influential and discerning their relevant ranges.

Optimization Concerned with the identification of the optimum solutions and investigating whether the optimum is unique or variant to meet conflicting demands.

Robustness testing Concerned with ensuring that the model well-rounded enough to survive certain conditions and to effectively represent reality in more than one scenario. It involves the adjustment of experimental factors with the aim of ensuring a robust model.

Eriksson et al. (2000) goes on to state that experiments are performed with variables of which there are two fundamental types:

Factors Tools for manipulating the system, process or model. Factors exert influence on the system and one aims to investigate or map the nature of said influence.

Responses That which informs the experimentor about the properties and general conditions of the system, process or model. Responses reveal the nature of the system's behaviour, be it healthy or unhealthy.

Screening experiments are performed in the beginning of a model's life. The aim is the exploration of specified factors and the influence exhibited by the model's responses. Essentially, screening is performed as a type of model validation to ensure that the model behaves as it should and under what conditions undesirable behaviour detected. Thus, in the context of the problem of communicable disease disaster management and the developed ABM described in Chapter 3, screening is the type of experiment that must be conducted to discern what conditions facilitate disease spread.

Optimization experiments are conducted upon the completion of screening experiments. Optimization experiments aim to predict response values for all possible combinations of factors within the experimental region (Eriksson et al., 2000). Thus, in the context of the problem of communicable disease disaster management, optimization experiments are those that will be performed with a fully validated model to monitor outbreaks and their escalation patterns.

It was established in Chapter 1, that outbreaks escalate into epidemics where the number of those Infected with an infectious disease spreads beyond a few people and throughout a population. Therefore, optimization experiments are those that will enable proactive decision making and will hence form part of the Mitigation and Preparedness phases of disaster management. Scenario analysis discussed in the latter stage of Chapter 4 is one means of optimization experimentation.

Lastly, robustness testing is the final form of model examination that is performed before the model is released for practical use. It aims to ensure that the model will be able to endure use and testing within a practical environment. Technically speaking, the aim is to ascertain that the model is robust to small fluctuations in the factor levels.

The generic model developed in Chapter 3 enters an absorptive state under certain conditions. In fact, certain conditions require the unrealistic inception of 200 or more Exposed individuals before the Markov Chain becomes functional and disease spread can thus be facilitated. For this reason, before the model can be of use, screening must be applied to investigate the factors that are required to facilitate disease spread. Without this step, the model is a purposeless computer simulation.

According to Eriksson et al. (2000) Design of Experiments (DOE) entails the design and tailoring of representative experiments with regards to a given question. With respect to the ABM, the question is concerned with the conditions required to facilitate disease spread through the Markov Chain's implementation.

4.1 Screening experiment design

The instinctive means of experimentation is that where one factor is changed at a time and the effect is monitored and tracked. This form of testing referred to as "Changing one separate factor at a time" testing or COST testing. Different implications or results are revealed by experiments with different starting points. For this reason, COST testing is described as an inefficient form of testing or experimentation (Eriksson et al., 2000).

A preferred approach entails the construction or design of carefully chosen experiments in which all relevant factors are varied. Such experiments also require that the input conditions (stipulations that define the experiment) are specified. Input conditions include the experimental objective and the number of factors and their ranges.

Experimental objective The experimental objective is the investigation of the conditions required to facilitate disease spread through the Markov Chain's implementation. Due to the inclusion of time based transitions owing to the lack of rate or probabilistic data, the conditions under which the simulation begins (initial conditions) determine whether or not the Markov Chain will play a role in the model and allow for the spread of the disease. Therefore, investigating the combination of initial conditions that will allow for an effective representation of disease spread is critical in the process of model validation.

Factors The conflict exhibited between the Markov Chain and time based transitions is the reason behind the model entering an absorptive state. If the probability of an Infected individual infecting a Susceptible individual is too low, the time based transitions lapse before another Infected individual can be born into the model.

The time based transitions can't be altered if the model is to be realistic. The duration of the incubation and latent periods are documented actualities regarding each communicable disease. Thus, the Markov Chain is the conflicting party that is open to investigation.

The probability of a new Infected is the recursive component of the Markov Chain (see equations (4.1) and(4.2)) that is used as an internal transition within the Infectious and Symptomatic states pertaining to the measles model and, within the Infected state pertaining to the H1N1 model. It determines the likelihood of a Susceptible individual contracting the infectious disease. Thus, progression into an absorptive state depends on the probability of a new Infected recursion.

$$\lambda(i) = \frac{\alpha i}{N} \tag{4.1}$$

$$\pi_{ir}\Delta t = \lambda(i)(N - i - r)\Delta t \tag{4.2}$$

Regarding the recursive component, the equation (4.1) is affected by the contact rate α and the number of Infected individuals in each time interval. Similarly, equation (4.2) is affected by equation (4.1)and the number of Susceptible individuals $(N - i - r)$ in each time interval. For these reasons, the following two factors for experimentation have been identified:

Contact Rate The contact rate α is a measure of the number of successful contacts an individual may have with another individual. As such, it is a sensitive or variable value that may change for many reasons, one of which being the population density.

Number of initially Infected individuals The number of initially Infected individuals affects the demographic of the population. Since the population size must remain constant, the subsequent number of individuals within the other compartments will alter if one changes. By altering the number of initially Infected individuals, one will consequently alter the number of Susceptible individuals which will in turn alter equation (4.2) and hence the probability of a new Infected. The measles and H1N1 model both have the same Infected state, Exposed, thus the number of initially Exposed individuals will be factor under investigation.

The following ranges are defined per factor:

$$\begin{aligned} \text{Contact rate} = &\quad 5 \leq \alpha \leq 25 \\ \text{No. of initially Exposed individuals} = &\quad 100 \leq \text{“E”} \leq 300 \end{aligned}$$

Response The previously identified factors in themselves dictate the responses. The function of the probability of a new Infected π_{ir} and the population demographic will be responses measured with respect to time.

The probability of a new Infected will be measured according to the time that it takes to reach and stay zero. This is a measure of the length of an outbreak which is a measure of the ability of the disease to be spread within the population. If disease spread is facilitated, the outbreak will last for a longer period of time before it's eradicated due to the recovery of all Infected individuals.

Similarly, the population demographic with respect to time is a measure of the the number of individuals within each state. The time till the number of the individuals within the Exposed is zero and constantly so will be the response under study. If no new Exposed individuals are created, disease spread is no longer facilitated.

Experiments will be conducted using *AnyLogic* and Microsoft Excel. *AnyLogic* contains connectivity functionality that permits data to be translated to and from a Microsoft Excel format. Simulation and experiment data will be collected and captured as time based data-sets that will be written to a Microsoft Excel spreadsheet. The data will then be arranged and manipulated as required. Furthermore, graphs will be made to better present the data and the findings.

4.1.1 Control experiment

A control is the base experiment which is the basis for comparison for all other subsequent experiments. Brown and Hollander (1977) state that one must first consider a sample-case when conducting medical research experiments.

With regards to communicable disease, a sample-case is referred to as "case-zero". The instance where the first positively Infected individual is identified. The circumstances surrounding the individual's contraction of the communicable disease and secondary infections they are likely to have induced will hence be investigated. This, is the beginning of case surveillance.

The aim of the control experiments is to substantiate the claim made in Chapter 3 (Generic model development) that more than one Exposed individual is required in order to facilitate disease spread. Accordingly, the responses to measure this are the population demographic regarding number of individuals in each state and, the probability of a new Infected.

The following population parameters are defined:

$$\begin{aligned} \text{Population size} &= 500 \text{ agents} \\ \text{Contact Rate} &= 5 \\ \text{No. of initially Exposed individuals} &= 1 \end{aligned}$$

Measles results

Table 4.1: Table depicting measles control experiment results

			π_{ir}		No. of Exposed
Experiment number	α	No. of initial Exposed	Max. value	Days to zero	Days to zero
1	5	1	0.00002	15	3

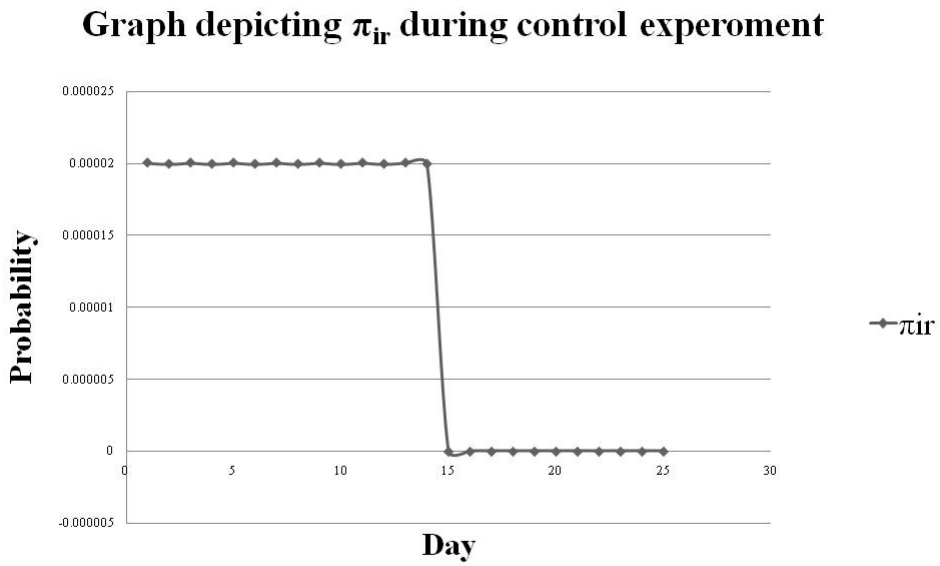

Figure 4.1: Graph mapping the probability of a new Infected during the measles control experiment

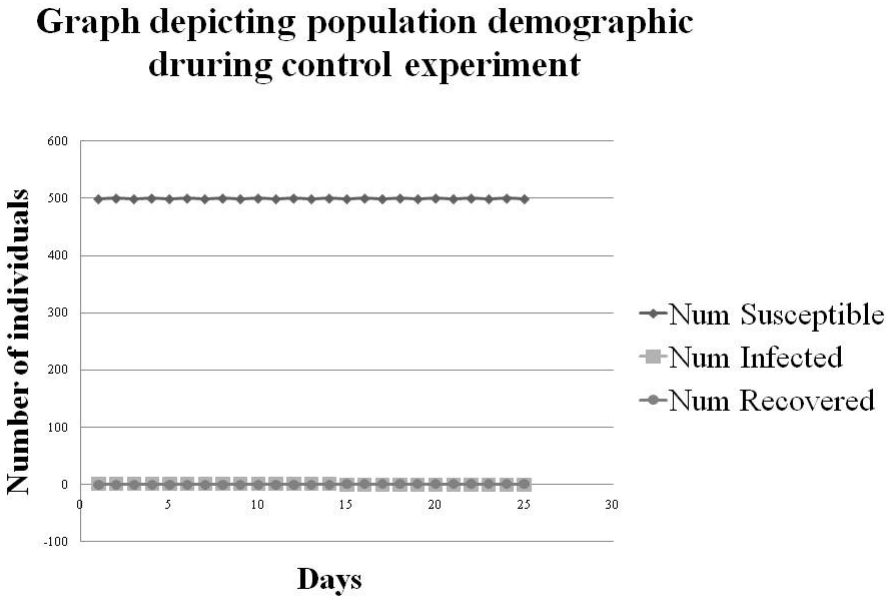

Figure 4.2: Graph mapping the population demographic during the measles control experiment

Observations Table 4.1.1 indicates that the maximum probability of a new Infected individual is 0.00002. However, the difference between the maximum and minimum probabilities is 0.00000008 as

depicted by Graph 4.1. Thus, it can be deduced that no significant change occurs to the probability of a new Infected individual as it remains minute until it eventually becomes zero after 15 days.

Furthermore Graph 4.2 indicates, the population demographic remains constant throughout the simulation. Thus there are no individuals besides the initial Exposed agent that undergo state changes which in turn means that disease spread or transmission is not facilitated. If no Infected individuals are "born" then no Susceptible individuals have contracted the infectious disease.

H1N1 results

Table 4.2: Table depicting H1N1 control experiment results

			π_{ir}		No. of Exposed
Experiment number	α	No. of initial Exposed	Max. value	Days to zero	Days to zero
1	5	1	0.00002	6	3

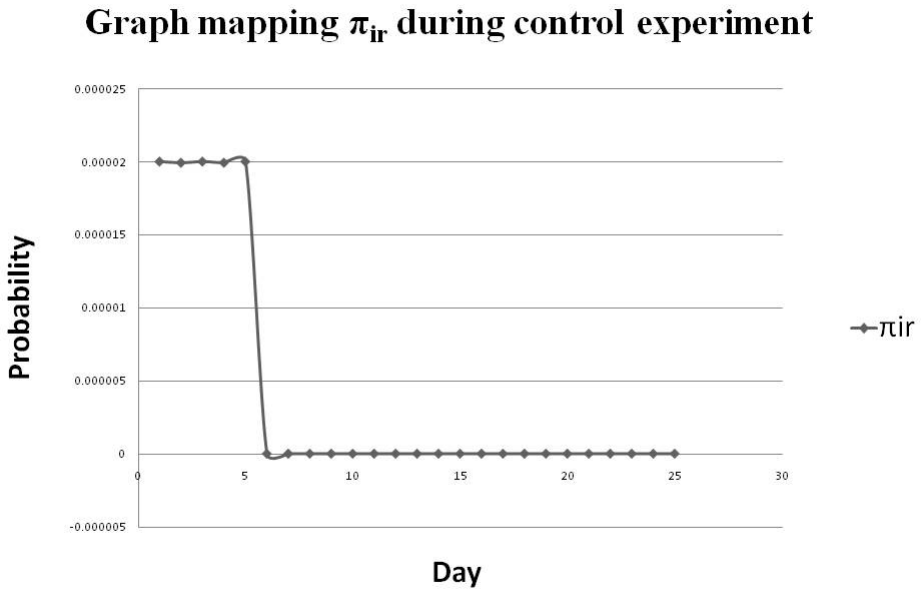

Figure 4.3: Graph mapping the probability of a new Infected during the H1N1 control experiment

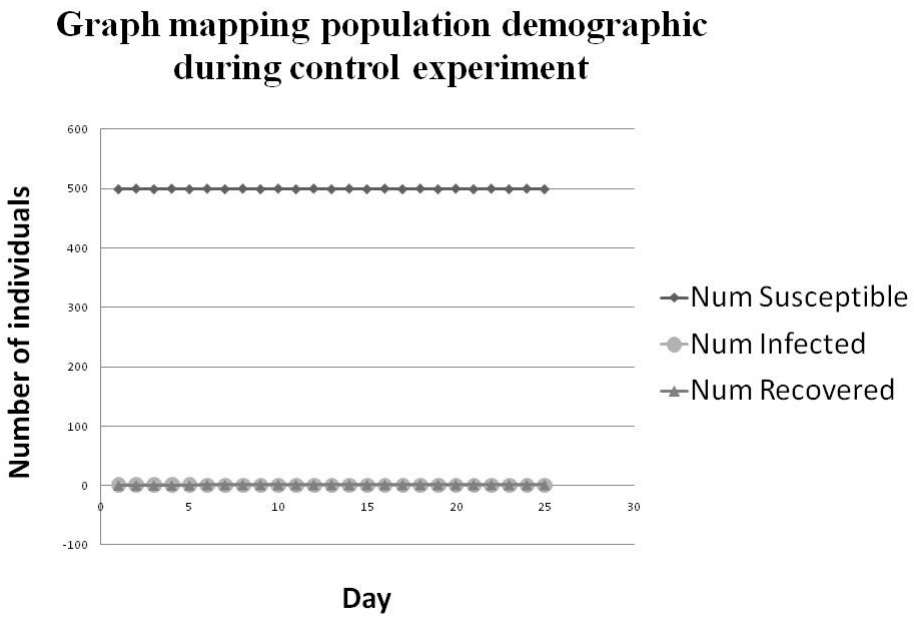

Figure 4.4: Graph mapping the population demographic during the H1N1 control experiment

Observations Since the Markov Chains of each (measles and H1N1) control experiment are identical and so too are the initial conditions, similar observations are made regarding the control experiments.

Graph 4.4 reveals that population demographic of the H1N1 control experiment is identical to that of the measles experiment (see Graph 4.2).

The probability of a new Infected individual requires six days to reach zero opposed to the 15 days required for the measles control experiment. Nevertheless, the H1N1 control experiment also corroborates the previous claim that one Exposed individual is not sufficient to facilitate disease spread.

4.1.2 Screening experiments

A total of nine experiments will be conducted per infectious disease model. Experiments will make use of the minimum, median (middle) and maximum value of the range for each factor. The three factors will then be arranged into nine ($3^2 = 9$) unique conditions for experiment and each experiment corresponds to a number from one to nine.

Measles results

Table 4.3: Table depicting measles screening experiment results

			π_{ir}		No. of Exposed
Experiment number	α	No. of initial Exposed	Max. value	Days to zero	Days to zero
1	5	100	0.0025	19	9
2	15	100	0.007843	30	18
3	25	100	0.013939	28	15
4	5	200	0.006687	31	16
5	15	200	0.023191	34	19
6	25	200	0.047936	38	25
7	5	300	0.017027	31	17
8	15	300	0.064239	32	20
9	25	300	0.159735	36	25

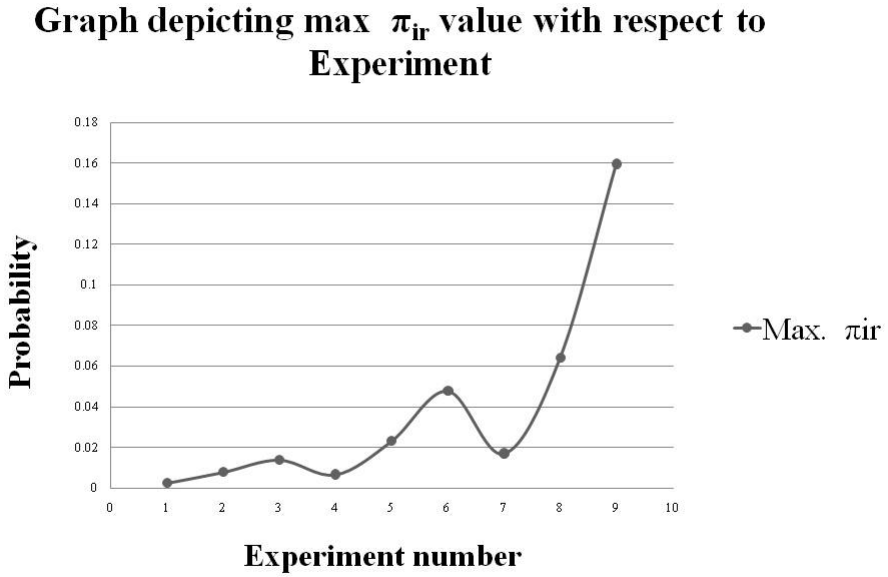

Figure 4.5: Graph depicting measles maximum probability of a new Infected with respect to experiment

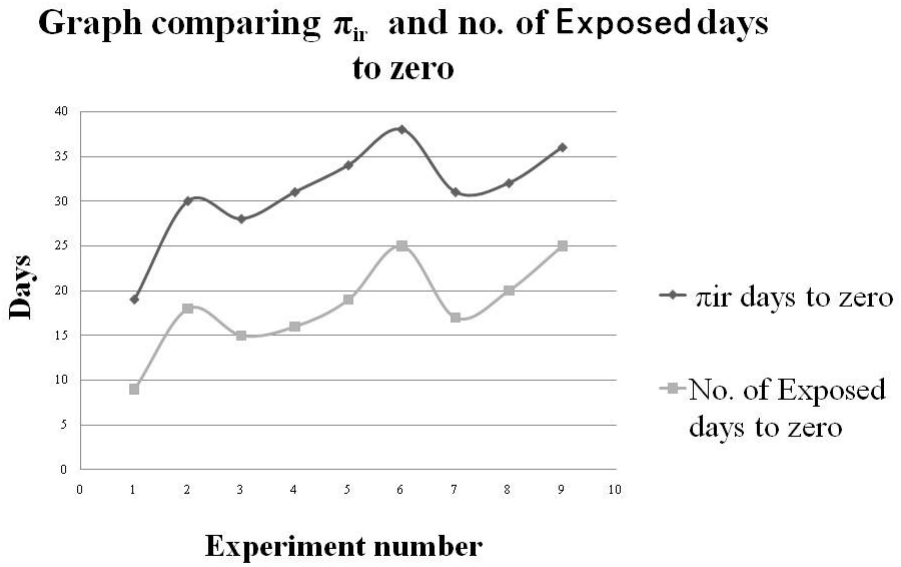

Figure 4.6: Graph comparing measles days to zero probability of a new Infected and no. of Exposed

Observations Table 4.1.2 and Graph 4.6 indicate that experiment number six takes the longest time before both the probability of a new Infected and number of Exposed individuals are zero. Consequently, experiment number six exhibits the longest sustained spell of disease spread. Thus, for the measles simulation, the preferred initial conditions are:

$$\begin{aligned} \text{Contact Rate} &= 25 \\ \text{No. of initially Exposed individuals} &= 200 \text{ (40\% of total population)} \end{aligned}$$

Furthermore, Graph 4.5 indicates that the longest sustained spell of disease spread doesn't necessarily correspond to the greatest maximum π_{ir} value. In fact, experiment six experiences a local opposed to a global maximum value.

H1N1 results

Table 4.4: Table depicting H1N1 screening experiment results

			π_{ir}		No. of Exposed
Experiment number	α	No. of initial Exposed	Max. value	Days to zero	Days to zero
1	5	100	0.0025313	9	6
2	15	100	0.00745	9	7
3	25	100	0.012657	10	8
4	5	200	0.006667	7	4
5	15	200	0.02	11	8
6	25	200	0.033893	6	9
7	5	300	0.015	10	10
8	15	300	0.04693	10	7
9	25	300	0.082275	11	8

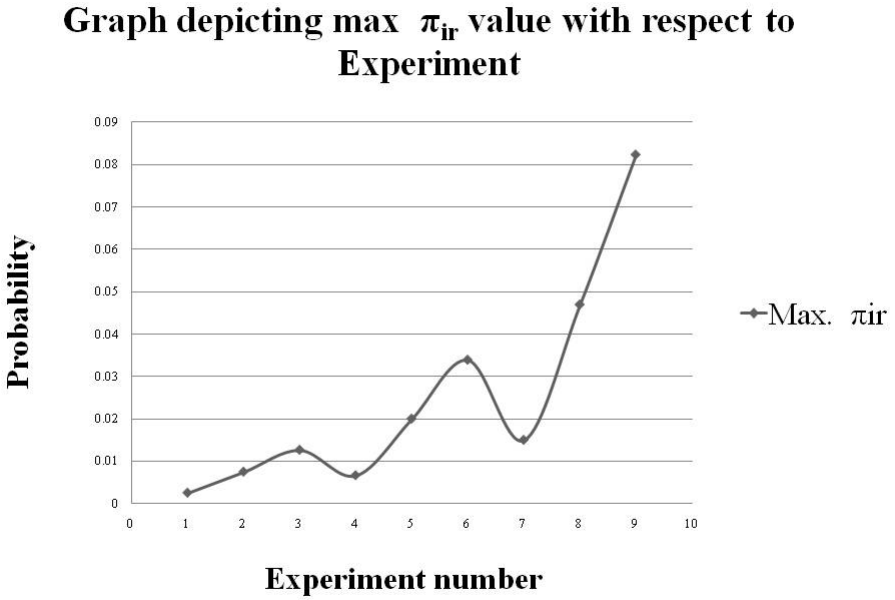

Figure 4.7: Graph depicting H1N1 maximum probability of a new Infected with respect to experiment

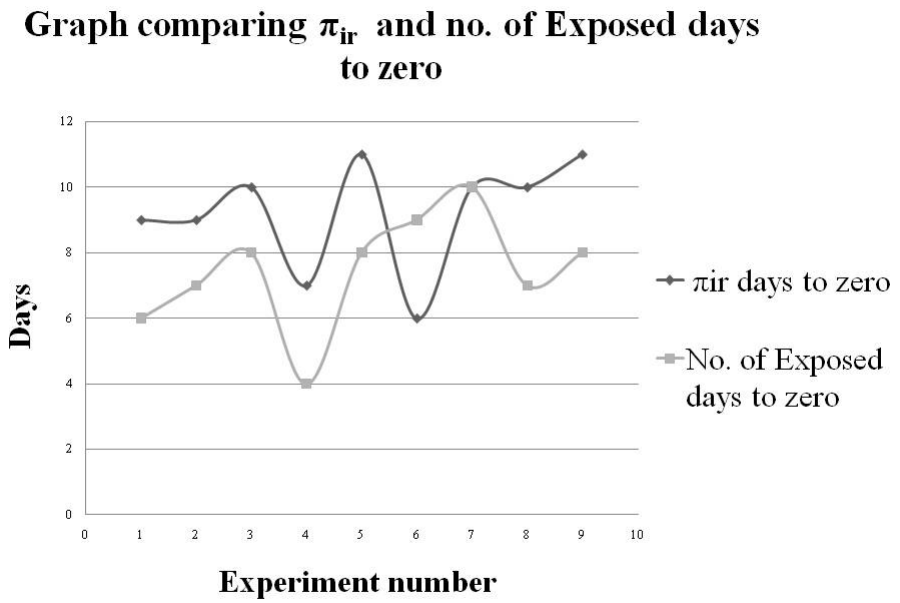

Figure 4.8: Graph comparing H1N1 days to zero probability of a new Infected and no. of Exposed

Observations The no. of days to zero Exposed individuals opposed to the no. of days to a zero probability of a new Infected is a truer reflection on the respective experiement's ability to facilitate disease spread. If no new Exposed individuals are "born", disease spread is no longer facilitated. The probability of a new Infected may not be zero due to the existence of Infected individuals. However, the probability of a new Infected may be too low to allow for disease transmission.

Therefore, experiment seven exhibits the longest sustained spell of disease spread. Graph 4.8 indicates that experiment seven experiences takes the longest time before the number of Exposed individuals is zero. However, experiment five takes the longest time before the probability of a new Infected is zero. Nevertheless, since the time to a zero probability of a new Infected is a truer reflection of an experiement's ability to facilitate disease spread, experiment seven exhibits the longest sustained spell of disease spread. Thus, for the H1N1 simulation, the following initial conditions are preferred:

$$\begin{aligned} \text{Contact Rate} &= 5 \\ \text{No. of initially Exposed individuals} &= 300 \text{ (60\% of total population)} \end{aligned}$$

4.1.3 Experimental conclusion

Comparisons of the measles and H1N1 control and screening experiments reveal that the time-based transitions play a large role in the effectiveness of the Markov Chain and the enablement of disease transmission or spread.

Graphs 4.5 and 4.7 depict the maximum probability of a new Infected value with respect to screening experiment. The trends or gradients of both graphs are identical. This is due to the identical behaviour of the Markov Chain during each experiment under the same conditions. However, the values (disease spread duration) differ according to the respective disease.

Thus the Markov Chains behave the identically yet, H1N1 conveys lower time to exhaustion values since disease transmission lasts for a shorter period of time.

Comparisons between Graphs 4.6 and 4.8 reveal that measles is transmitted for a longer period of time since even the lowest times taken for measles are greater that the maximum times taken for H1N1.

The time based transitions are implicated in this fact due to their duration. A single individual who has contracted measles, will take at most 18 or 22 days to transition from Exposed to Recovered. Conversely, a single individual who has contracted H1N1 will take at most six days to transition from Exposed to Recovered.

Thus, the Infected state for measles lasts longer than that for H1N1. The longer the Infected state lasts, the greater the probability of a new Infected since it is possible for more individuals to exist in this state. The greater the probability of a new Infected, the greater the number of Susceptible individuals that will contract the infectious disease subsequently increasing the number of Infected individuals which will impact the greater the probability of a new Infected.

Measles screening experiment six presented with the best conditions for transmission of measles. Similarly, H1N1 experiment seven presented with the best conditions for spread of H1N1. However, comparison of the responses for each experiment reveal that measles better facilitates disease spread through the Markov Chain. Figure 4.9 substantiates this statement.

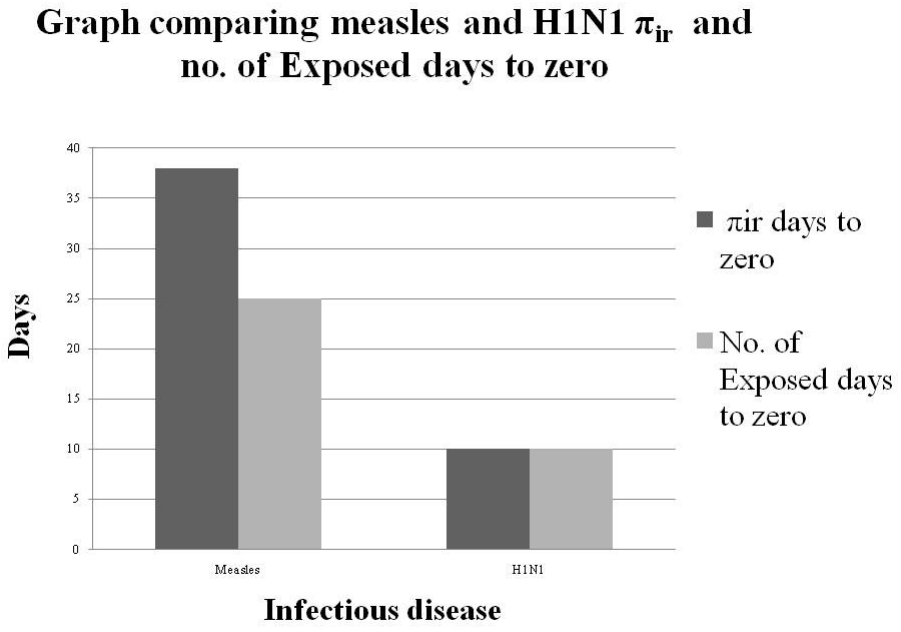

Figure 4.9: Graph comparing measles and H1N1 days to zero probability of a new Infected and no. of Exposed for established initial conditions

Thus, as conducted experiments will attest to, the Markov Chain is more effective in the simulation of measles. Furthermore, it can be concluded that the Markov Chain is better suited to infectious diseases that experience long Infected periods. The transmission or spread of the infectious disease is dependent on the length of the Infected period and hence dependent on the length of the simulation's time-based transitions.

4.2 Optimization experiment design

As previously mentioned, optimization experiments aim to predict response values for all possible combinations of factors within the experimental region (Eriksson et al., 2000). Furthermore, optimization experiments are concerned with the identification of the optimum solutions and investigating whether the optimum is unique or variant to meet conflicting demands.

With regard to the spread of infectious disease, conflicting demands refer to the different circumstances or scenarios that surround the experimental region. Scenario analysis is proposed where the testing and investigation of defined situations or circumstances concerning communicable disease and its transmission are conducted.

Hethcote (2008) defines communicable disease models as a practical approach to answering questions about communicable diseases and the strategies used to combat them. The challenge lies in the identification of specific valuable questions and the level of information required to answer them.

Literature, data and interviews will be the indicators that forge the path towards realistic scenarios that are worthy of testing. Scenario analysis is expected to convey intelligence that can be used in the fight against communicable diseases and their consequences. As previously stated, one should aim to reduce the impact of communicable disease by preventing or reducing the escalation of outbreaks into epidemics.

Unfortunately, optimization experiments require a realistic model in order for them to be of any value. Models that require 40% or 60% of the total population to be Exposed to the infectious disease are not realistic.

Therefore, due to the conflict exhibited between the Markov Chain and the time-based transitions, optimization experiments can't be performed as the time where disease transmission is active is too short for the experiments to be true reflections of reality.

In light of the observations made during screening experiments, a plausible scenario will now be suggested per communicable disease model. The reasoning and rationale behind each scenario will be given.

Hethcote (2008) argues that comparisons lead to a better understanding of the spread of disease. For this reason, two scenarios, a base and secondary scenario will be compared for each disease. Each scenario entails the adaptation of the generic infectious disease model to suit the relevant circumstances or situation.

4.2.1 Measles

Measles is classified as a preventable communicable disease. A vaccine in injection form is administered to children at 9 months and a booster is administered at 18 months. This vaccine is free at all public health facilities however, a vast number of people remain unvaccinated and susceptible to infection. According to United Nations International Children's Emergency Fund (2008) (UNICEF) health statistics, 62% of South African children receive at least one dose of the measles vaccine before they are 1 year old. Accordingly, measles optimization or scenario experimentation should incorporate the existence of immunised or vaccinated individuals in the model. Upon the commencement of the simulation, Recovered and Exposed agents should be created.

Scenario: Test immunisation strategies

Base scenario The nature of measles transmission in an environment where immunisation is performed before the outbreak. Essentially, before the introduction of Exposed individuals into the

population, a certain percentage of the population will be set to Recovered as they will already be immunised.

The base scenario will be modelled using the generic SEISR model. Before the model is executed, the number of Recovered individuals with respect to the number of Susceptible will be set. This will incorporate the determined percentage of immunised individuals while ensuring that the initial conditions obtained through screening experiments hold. The simulation will be executed and notes will be taken regarding the time it takes for the outbreak to escalate and whether or not it dies out.

Secondary scenario Nature of measles transmission in an environment where immunisation strategies are implemented during an outbreak. Once a measles outbreak has been confirmed, its human nature to then proceed to attempt to educate and immunise the masses. Essentially, once the problem is identified, only then do immunisation and education drives begin.

The secondary scenario will also be modelled using the generic SEISR model with the inclusion of a Vaccinated state within the state chart. A separate Vaccinated state is warranted since it will enable the modeller to distinguish between an individual that is immunised before and during the potential epidemic.

According to literature from Kassner (1985) and Benenson (1985) and corrobarated with interviews from front-end medical personnel, Susceptible and Exposed people are the main targets for vaccination. Susceptible people are clear targets when one aims to increase the immunity of the herd population. Similarly, Exposed people are those who have contracted the disease and in some cases it is believed that the vaccine can prevent the individual from progressing to the Infectious state.

The condition that dictates transition into the Vaccinated state consists of two sub-conditions. The first sub-condition is concerned with the time period where vaccinations occur while the second is concerned with the percentage of the population that has access to healthcare and thus the vaccine. Consequently, in order for vaccination to occur, both sub-conditions conditions must be valid.

The time period condition is based on the analysis of the base scenario. The base scenario will be monitored to find the instant where the outbreak is at it's peak. This instant will be identified by exhibiting the largest $\pi_{ir}\Delta t$ value. When $\pi_{ir}\Delta t$ is the largest, the most Infected (Exposed, Infectious and Symptomatic) individuals exist than in any other time period. The identified time period will be dictate the the time based sub-condition.

The second sub-condition is depends on statistics obtained from Statistics South Africa (2004). Currently, it is estimated that 62.7% of South African's have access to healthcare. Healthcare includes access to clinics and hospitals within a distance of 2km. This statistic is controlled by an event, "VaccinationTrigger" that executes each instant during the designated vaccination period. A random number is generated with each execution and if the random number is less than or equal to 0.627 (62.7%) then this condition is valid.

4.2.2 H1N1

Before H1N1 scenario analysis can be performed, the SEIR model must be augmented to include the arrival of Exposed individuals into the population. The Exposed individuals that arrive are a consequence of travelling as discussed in Chapter 3. An Arrival cyclic event will be programmed that models the cyclic arrivals of Exposed individuals. The Arrival event will be programmed to occur at a time period according to the modeller's discretion.

H1N1 is an emergent disease. For this reason an effective response hinges on the speed of the official recognition of the disease. H1N1 symptoms are similar to those of the regular flu and as such

the general public will respond to treat their supposed regular flu. H1N1 however, is an advanced influenza strain that doesn't respond to regular influenza treatment. All the while, the virus is multiplying within the host and the host in turn is exposing people they come into contact with.

Base scenario H1N1 transmission in a population according to the generic SEIR model and with the inclusion of Exposed travelling arrivals.

Secondary scenario H1N1 transmission in a population following the Base scenario with the inclusion of the element of emergent disease recognition time. While the disease remains unidentified or unrecognised, people will recover from the disease through basic treatment or the disease will run its course. However some people will not recover and unfortunately die from the disease. This statement is corroborated by Figure 3.3, more specifically the period of 20 July 2009 - 20 September 2009.

In Figure 3.3, the number of deaths recorded in this period is notably greater than in any other period. Thus a new state Dead will be incorporated to account for those individuals that die from the disease within this period.

Initially, the model will have an extra state, Dead for examination purposes. Essentially, while the disease is unidentified, some Infected individuals will recover and some will die. The inclusion of the Dead state allows one to explicitly see the effect disease recognition has on the number of individuals that die or recover from the disease. After the disease's official recognition, the likelihood of death drops to almost zero and the recovery time increases.

Official recognition will be modelled by a user controlled event. In other words, the modeller decides when the disease has been officially recognised and programs the event accordingly. The time interval for the event's commencement and its duration are thus at the discretion of the modeller. As such, sensitivity analysis can be performed to investigate the model's response to fluctuations in either factor.

Chapter 5

Future reference

Future reference is concerned with opportunities for the model's improvement and future topics of investigation . As the model increases in complexity, it will grow to be a better approximation of the actual disease and its transmission within an environment. Assumptions were made in the development of the generic model that resulted in the model being less realistic. One such assumption is concerned with the population. It was assumed that the population is uniform and homogeneous. This means that individuals within the population are the same and don't differ according to age, race, gender or genetic makeup and thus, everyone has the same probability of contraction. Areas for improving the realistic nature of the model will be identified in this chapter.

5.1 Incorporation of non-uniform heterogenous population

In reality, no population is completely uniform. In fact, Eileen Caddy, a Scottish spiritual teacher once said that *"A human being is a single being. Unique and unrepeatable."*. By correcting the false assumption that all human beings are the same, a more realistic model is expected. Three means of correction were identified

Incorporation of age structure In a uniform population, all agents are said to be the same age. However, as Chapter 3 data will attest two, this is not realistic. Incorporating the age within an *AnyLogic* model could take place in many ways, two of which will be explained.

Inclusion of a second state chart Inclusion of a second "Age" state chart where each state represents and distinct age compartment. Each compartment represents a particular age group. Rates of aging will function as the transitions that determine an agent's progression from one age group to the next.

Inclusion of age by considering age specific contact rates Individuals of a school or university going age can be said to have more contacts per day due to the nature of the school environment. The elderly can be said to have fewer contacts due to the likelihood of their age decreasing their mobility.

A third alternative could incorporate both previously discussed means of age incorporation. An age specific contact rate could be applied to each age compartment or age group.

Incorporation of host genetic factors Genetic factors of the host can affect who is pre-disposed and thus more likely to contract an infectious disease. Modelling host genetic factors in *An*yLogic

first requires the establishment of which genetic factors should be considered. Secondly, data concerning the genetic factors within the population should be gathered and analysed. Next, the host population should be divided into subgroups according to genotypes those who have specific genes opposed to those that don't. Lastly, the modeller should decide whether to model the genetic factors individually or to aggregate their presence with respect to the host population

Incorporation of medical history and chronic illnesses An individual's history of illness or surgery has been known to lower their immune system and hence increase their chances of infectious disease contraction. Chronic illnesses such as asthma, diabetes, HIV, haemophilia, hypertension and sickle cell anaemia are also known to lower an individuals immune system.

Host generic factors and medical history could be incorporated into *AnyLogic* by assinging a pre-disposition variable to each agent. This variable will indicate the probability that an agent has of contracting the infectious disease. Further data gathering with respect to genetic factors and medical history will indicate the proportion of individuals within a population that are positive for certain genetic factors or medical histories. An example of data that can be used is Figure 5.1. Figure 5.1 is obtained from the NICD (National Institute of Communicable Diseases, 2010c).

Table 5.1: Table displaying clinical characteristics of H1N1 recorded deaths, October 2009 - February 2010

Factor	Factor Frequency:Number of cases with data available	%
HIV infected	19:38	50
Pregnant	26:91	28
Diabetes	11:83	13
Obese	18:84	21
Cardiac disease	8:82	10
Active Tuberculosis (TB)	9:83	11

5.2 Incorporation of spatial and geographic effects

Spatial and geographic effects can influence the transmission of infectious disease. Spatial effects refer to the proximity that individuals have to certain vectors (water, insects, plants etc.) and the access that they have to healthcare or means of transportation. Infectious diseases that are transmitted by vectors, for example malaria, require the presence of said vector in order to cause an outbreak. Access to healthcare will determine whether or not individuals have access to vaccines to prevent infection or treatment to combat infection. Lastly, access to means of transportation (airports, bus or train stations, taxi ranks etc.) will determine the likelihood of an Exposed or Infectious individual inadvertently introducing the infectious disease into a population.

Geographic effects refer to the landscape and phenomena such as climate and seasons. Haggard (1994) describes the historical geography of infectious disease as a pattern that is affected by strong population growth in the host population, by worldwide environmental changes associated with that growth, and by increased spatial mobility for both the disease causing micro-organisms and for the human host. Geographical changes include climate change (global warming), deforestation and damming.

Spatial and geographic effects can be modelled in *AnyLogic* with the incorporation of a geographical environment. *ArcGIS* is a software package that allows one to make use of Geographical Information Systems (GIS) thus enabling them to map, model and manipulate data on a geographic landscape or map. With regards to the current model, the environment in which the simulation is executed can be changed to an ArcGIS environment as opposed to the current 2-dimensional discrete environment.

A constructive scenario that incorporates *ArcGIS* could be one where Exposed or Infected individuals are placed at predetermined locations within the geographical environment. The model can hence be used to track the spread of the infectious disease within an geographical area (region) an furthermore, to track the escalation of the infectious disease from one area to another.

Furthermore, one should also consider mobility with purpose. In reality, people have the ability to move within their environment with a defined purpose. This is seen with people as they travel to school, work and other places to pursue and meet an objective. In doing so, Infected people have the ability to infect people that they come into contact with. Mobility with purpose could be depicted and programmed in an *ArcGIS* environment.

5.3 Access to health care

The measles secondary scenario used the Statistics South Africa statistic for the access to health care as a condition for the Vaccinated state trigger. However, access to health care is not necessarily an a true reflection on the access that an individual has to the vaccine. It is possible for an individual to have access to a health care facility that doesn't in fact have access to the measles vaccine. Thus, the vaccination trigger should be linked to the availability of the measles vaccine. This can be effectively modelled by incorporating a System Dynamic (SD) element that will depict the vaccine levels with respect to time. If there are no vaccines available, then no one can be vaccinated and no one can subsequently enter the Vaccinated state. The SD element can be included such that it incorporates both out-flow (use leading to potential exhaustion) and in-flow (replenishment).

5.4 Future topics of investigation

Investigation and analysis of compartmental epidemiological models The inclusion of time-based transitions results in the model entering an absorptive state under certain conditions. This is due to the conflict between the probability required to transmit disease and the length of time that individuals spend in the Infected state. The H1N1 model in particular conveyed through experiments that the Markov Chain did not effectively allow for the spread of the H1N1 infectious disease.

Thus, it is proposed that research into other probabilistic compartmental epidemiological models is performed with the aim of identifying a model that is better suited to the transmission of disease for infectious diseases that do not have long Infected time periods. The H1N1 model can be used as a baisis for experimentation.

Furthermore, research into other probabilistic compartmental epidemiological models can be performed to identify models that will require the inception of fewer Exposed individuals before disease transmission is enabled.

Threshold behaviour ($\mathscr{R}_0$) Threshold behaviour (see Chapter 2) can be explored and investigated using a combination of agent based and mathematical epidemiological modelling. However,

consideration must be given to the current formulation of the behavioural Markov Chain. Upon considering threshold behaviour, one must utilise the appropriate equations and formulae.

The following is said to be true for the situation where $\mathscr{R}_0 > 1$, where disease elimination requires a vast amount of time and the disease tends to become endemic (always present) in the population under study.

$$(\frac{1}{\mathscr{R}_0})^a \qquad \text{where a is the number of Infected individuals at t=0} \tag{5.1}$$

Similarly, for $\mathscr{R}_0 \leq 1$, where the disease is said to become extinct as all Infected individuals recover.

$$1 - (\frac{1}{\mathscr{R}_0})^a \qquad \text{where a is the number of Infected individuals at t=0} \tag{5.2}$$

Equation (5.1) calculates the probability that the epidemic fades out while equation (5.2) calculates the probability that the epidemic persists. Both equations depend on the initial number of Infected individuals. Thus, iterative calculations and graphs can be used to depict how the threshold is affected by the number of initial Infected individuals.

Passive immunity Newborns that are breast fed receive various antibodies and the associated immunity from their mothers. This immunity is temporary and hence expires. Thus the logic behind vaccinating children at nine months. The compartmental epidemiology state that contains newborns with passive immunity is denoted by the letter "M". Thus, the simple SIR model would now become the MSIR model.

However, mother's that never gained immunity through immunisation or through recovery do not possess the necessary antibodies to pass onto their children. For this reason, their children are Susceptible from birth.

Variable total population size The Markov Chain developed by (Allen and Burgin, 2000) requires that the population size remain constant. In fact, during model development, when the total population size varied, the Markov Chain and more specifically the Probability of a new Infected individual behaved erratically. Values for π_{ir} would jump around and become unstable. Situations arose where a π_{ir} value greater than one and in some cases greater than 200 was identified.

Therefore, future investigation can be performed to discover the reasons for the erratic behaviour and an appropriate solution where found can be documented.

Chapter 6

Conclusion

Communicable diseases in the African Context are considered disasters. Speed is crucial in the response to communicable diseases to limit their impact and prevent outbreak escalation. By virtue of simulation using mathematical epidemiology and agent based modelling, one can model infectious disease transmission and gain insight which can be used in the fight against communicable disease.

A Markov Cahin developed by Allen and Burgin (2000) as well as time based transitions were used to model the infectious disease's progression from state to state within an individual. However, the conflicting nature between the two lead to the investigation of the best initial conditions that allowed for the best representation of disease spread. Experiments were conducted which revealed that the Markov Chain was more effective in the simulation of measles due to its longer Infected state.

The model currently developed is a framework that can be expanded and reinforced to improve upon its representation of reality. From here, optimization experiments can be conducted upon which pre-emptive disaster management protocols and techniques can be observed.

Bibliography

African Reginal Office (2009). Pandemic(h1n1) 2009 in the african region. Technical report, World Health Organisation.

Allen, L. J. and Burgin, A. M. (2000). Comparison of deterministic and stochastic sis and sir models in discrete time. *Mathematical Biosciences*, 163:1–33.

Anderson, R. M. and May, R. M. (1992). *Infectious diseases of humans-Dyanmics and Control*. Oxford Science Publications.

Ball, P. (2003). The physical modelling of human social systems. *ComPlexUs*, 1:190–206.

Benenson, A. S. (1985). Contol of communicable diseases in man. Technical report, American Public Health Association.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99:72807287.

Brauer, F. (2002). Basic ideas of mathematical epidemiology. *Mathematical approaches for emerging and reemerging infectious diseases*, 125:31–65.

Brown, B. W. and Hollander, M. (1977). *Statistics: A biomedical introduction*. John Wiley and Sons.

Castillo-Chavez, C. and Yakubu, A.-A. (2002). Discrete time sis models with simple and complex population dynamics. *Mathematical approaches for emerging and reemerging infectious diseases*, 125:165–182.

CSIR (2010). Csir-about us.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Wilkstrom, C., and Wold, S. (2000). *Design of experiments: Principles and Applications*. Umertrics Academy.

Green, W. G. and Altay, N. (2006). Or/ms research in disaster operations management. *European journal of Operations Research*, 175:475–493.

Haggard, P. (1994). Geographical aspects of the emergence of infectious diseases. *Human Geography, 2, The Changing Geography of Disease Distributions*, 76:91–104.

Hethcote, H. W. (2008). *Mathematical understanding of infectious disease dynamics*, chapter 1:The basic epidemiology models: Models, expressions for Ro, parameter estimation and applications, pages 1–33. World Scientific.

Janssen, M. A. (2005). Agent based modelling. *International Society for Ecological Economics*.

Jones, K. and Moon, G. (1987). *Health disease and society-An introduction to medical geography*. Routledge and Kegan Paul Inc.

Kassner, C. (1985). *Communicable disease in the African continent*. Shuter and Shooter, 9 edition.

Lusambo-Dikasa, P. S. (2008). Strategic orientations for health emergency management in the african region. *African Health Monitor*, 8:2–3.

Magnus, M. (2007). *Essentials of infectious disease epidemiology*. Jones and Barlette Publishers.

Nasell, I. (2002a). Edemicity, persistence and quasi-stationarity. *Mathematical approaches for emerging and re-emerging infectious diseases*, 125:199–227.

Nasell, I. (2002b). Stochastic models of some endemic infections. *Mathematical Biosciences*, 179:1–19.

National Institute of Communicable Diseases (2010a). About us.

National Institute of Communicable Diseases (2010b). Case based rash surveillance. Technical report, National Institue of Communicable Diseases.

National Institute of Communicable Diseases (2010c). Influenza a/h1n1 ("swine-flu") situation report. Technical report, National Institute of Communicable Diseases.

Nelson, K. E., Williams, C. M., and Graham, N. M. (2004). *Infectious disease epidemiology: Theory and Practice*. Jones and Bartlett Publishers.

Statistics South Africa (2004). Perceived health and other health indicators in south africa. Technical report, Statistics South Africa.

United Nations International Children's Emergency Fund (2008). South africa statistics. Technical report, United Nations International Children's Emergency Fund.

Winston, W. L. (2004). *Operations research volume two: Introduction to probability models*, volume two. Thompson Brooks/Cole, fourth edition.

World Health Organisation (2008). Epidemic and pandemic alert and response strategic plan 2009 - 2013.

World Health Organisation (2010a). About who.

World Health Organisation (2010b). What is the pandemic (h1n1) 2009 virus? Technical report, World Health Organisation.

World Health Organisation (December 2009). Measles fact sheet report. Technical report, World Health Organisation.