Studies on Epidemic Control in Structured Populations with Applications to Influenza

by

Romarie Morales

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved October 2015 by the
Graduate Supervisory Committee:

Carlos Castillo-Chavez, Co-Chair
Anuj Mubayi, Co-Chair
Sherry Towers, Member

ARIZONA STATE UNIVERSITY

May 2016

ABSTRACT

The 2009-10 influenza and the 2014-15 Ebola pandemics brought once again urgency to an old question: What are the limits on prediction and what can be proposed that is useful in the face of an epidemic outbreak?

This thesis looks first at the impact that limited access to vaccine stockpiles may have on a single influenza outbreak. The purpose is to highlight the challenges faced by populations embedded in inadequate health systems and to identify and assess ways of ameliorating the impact of resource limitations on public health policy.

Age-specific per capita constraint rates play an important role on the dynamics of communicable diseases and, influenza is, of course, no exception. Yet the challenges associated with estimating age-specific contact rates have not been decisively met. And so, this thesis attempts to connect contact theory with age-specific contact data in the context of influenza outbreaks in practical ways. In mathematical epidemiology, proportionate mixing is used as the preferred theoretical mixing structure and so, the frame of discussion of this dissertation follows this specific theoretical framework. The questions that drive this dissertation, in the context of influenza dynamics, proportionate mixing, and control, are:

I. What is the role of age-aggregation on the dynamics of a single outbreak? Or simply speaking, does the number and length of the age-classes used to model a population make a significant difference on quantitative predictions?

II. What would the age-specific optimal influenza vaccination policies be? Or, what are the age-specific vaccination policies needed to control an outbreak in the presence of limited or unlimited vaccine stockpiles?

Intertwined with the above questions are issues of resilience and uncertainty including, whether or not data collected on mixing (by social scientists) can be used effectively to address both questions in the context of influenza and proportionate mixing. The objective is to provide answers to these questions by assessing the role of aggregation (number and length of age classes) and model robustness (does the aggregation scheme selected makes a difference on influenza dynamics and control) via comparisons between purely data-driven model and proportionate mixing models.

DEDICATION

To my parents Roberto and Marta, for not pressuring me to marry, have 20 kids or to drop out of school.To my brother Richard for always reminding me to keep the eye on the ball. To my sisters, Liz and Zulicka for being a great example of an independent female warrior.

To my children (in the future); Mami, did it...so can you!

To Kehinde Salau; Cheers, I conquered!

ACKNOWLEDGMENTS

It is hard to put in to words the appreciation for my advisors and mentors. Dr. Carlos Castillo-Chavez,Dr. Anuj Mubayi, and Dr. Sherry Towers thanks for being the pillars that held me during the last few years. I could have not done it without you. Special thanks to all my friends and colleagues that worked with me side by side. Thanks to Rina, Claudia, Bichara, Kamal, and Komi for stepping to help me when I needed the most. To my partner in life, Kehinde Salau, may God give you all the years in the world so we can continue our journey together. Thanks for always making me smile :)

This dissertation has been partially funded by the Alfred P. Sloan Foundation, the Mathematical, Computational and Modeling Science Center (MCMSC) and More Graduate Education @ Mountain State Alliance (MGE @ MSE). The following published, open access, paper is included in this dissertation: Chapter 1:Lee, Sunmi, Romarie Morales, and Carlos Castillo-Chavez. "A note on the use of influenza vaccination strategies when supply is limited." Math. Biosci. Eng 8.1 (2011): 171-182.

TABLE OF CONTENTS

CHAPTER Page

APPENDIX Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Chapter 1

BACKGROUND ON INFLUENZA, SIR MODEL, VACCINE IMPLEMENTATION AND RESEARCH FOCUS

1.1 Influenza Background and SIR Model

An influenza pandemic occurs when a non-human (novel) influenza virus steadily and effectively enters a human population and quickly leads to a global outbreak [1, 33, 45]. The World Health Organization (WHO) provides an influenza pandemic alert system, with a scale ranging from Phase 1 (no known infections to humans) to Phase 6 (a full-blown pandemic) [19, 68, 82]. The influenza viruses are classified into types A, B and C on the basis of their core proteins. The virus types differ in their level of spread and impact within its host populations. Types A and B can only lead to cases in humans with a possibility of a full-blown pandemic [20, 47, 82]. The focus of this dissertation is tied in to influenza A outbreaks in human populations. The subtypes of influenza A viruses are determined by their membrane glycoproteins which could be either haemagglutinin (HA) or neuraminidase (NA) [4, 82]. High mutation rates and frequent genetic re-assortments of these viruses contribute to even great variability of the HA and NA antigens. All of the currently identified 16 HA and 9 NA subtypes of influenza A viruses are maintained in wild, aquatic bird populations. Viruses of the subtypes H1, H2 or H3, and N1 or N2 generally infect humans. The primary transmission pathways of influenza are disseminated in the environment by unprotected coughs and sneezes. Short-distance direct airborne transmission of influenza viruses between individuals may occur, particularly in crowded enclosed spaces. Hand contamination and direct inoculation of virus is another possible source

of transmission. Influenza pathogenicity may vary from asymptomatic to fatal infection. Typical influenza symptoms include fever with abrupt onset, chills, sore throat, cough, and are often accompanied by headache[5, 49, 52, 66, 64, 70, 76]. Influenza weakens the immune system and can cause patients to suffer from pneumonia. In this theoretical study the A-H1N1 2009 Pandemic is used to formulate the scenario that highlight the results of this dissertation.

Pandemics influenza are different from Seasonal flu outbreaks. Seasonal influenza is caused by influenza viruses that are similar to those already affecting people unlike in the case of pandemic influenza, which is the result of emergent new influenza viruses; that is, viruses that have had not been isolated in human populations. Seasonal influenza generally causes an acute viral infection. It mainly affects children younger than 2 years old, adults 65 years or older, pregnant women, and people of any age experiencing chronic heart,lung kidney,liver blood or metabolic diseases (such as diabetes), or weakened immune systems. The death rate for seasonal influenza can be high, for example, around 56,979 deaths were attributed to influenza in the United States of America [87].

The A-H1N1 2009 pandemic Influenza resulted in 12,469 deaths (8,868-18306) in the United States, a surprisingly low number [44]. However unlike seasonal influenza, the virus responsible for the A-H1N1 pandemic affected young adults. Relative protection from prior influenza exposure helps our understanding of the transmission dynamics of influenza [14]. Key factors associated with the transmission dynamics of influenza include variations in individuals' susceptibility, variability in mixing among the population, differences in infectiousness of infected people and changes in behavior due to an infection outbreak [13, 72, 71]. Variations in susceptibility are often a function of individuals' immune system responses, which depends on the individuals' prior history of exposure to related influenza strains [71]. Variability in mixing are

the result of variations in the interactions of individuals in a population [12].

Changes in behavior can be due to the levels of information (education) on the disease and decisions to avoid or not activities that might lead to infection during an outbreak [30]. Mixing of individuals is directly related to infection patterns in the population [14, 13, 25]. Hence, understanding and modeling the mixing structure of a population is one of the important aspects associated with the study of the dynamics of a disease. In this study, our goal is to understand the impact of the variability in a population's mixing structures on the transmission dynamics of influenza, using mathematical models and available mixing data. The SIR epidemic model, where S stands for (susceptible) individuals not yet infected at the current time; I represents individuals that are infectious, in other words, individuals that have the diseases and can spread it to others; and R refers to individuals who were infected but no longer have and cannot transmit the infection, underlies our modeling framework first introduced in 1927 by Kermack and McKendrick [51]. Historically, the SIR continuous models have been used to model infectious disease dynamics and provide public health policy guidance [10]. A marked increase in the use of SIR transmission models, continuous and discrete, persists [2, 3, 11, 16, 17, 40, 75].

Optimal control methodology, a powerful tool, is used to make decisions involving complex situations. The behavior of the system, as described by the state variables, is continuously steered via suitable control functions. The goal is to 'adjust' the control in order to maximize or minimize a given objective functional. The costs associated with a disease are incorporated in the objective function; and so, the goal to achieve maximum reductions at a minimal cost.

There has been many studies in the literature on the dynamics of pandemic or seasonal influenza. Ball et al.(2002) considers stochastic epidemics among a population partitioned into households; mixing locally within households and globally

throughout the population. In this study, the local transmission rate was assumed higher than the global. Their optimization analysis led them to the conclusions that vaccinating at the very beginning of the infection is, not surprisingly, the best strategy if the goal is to reduce the transmission of a fast spreading disease[6]. Merl et al. (2009) searched for optimal vaccination strategies for influenza outbreak within a homogeneously mixing population within a Bayesian framework. Their vaccination strategies are adjusted to reflect the anticipated trajectory of the epidemic given the population's current state and approximately updated parameter estimates. That is, the transmission rate from the previous time step is used and the vaccination strategy is adjusted accordingly [62]. Ludkovski et al. (2010) used the work of Merl et al. in the context of (closed) communities where medical officers can monitor each individual [59]. Their model though important- assumed homogenous mixing even though age structure rates may have been critical [59].Finally, a volume of papers connected to the 2009 pandemic influenza that focusses on the role of mobility, optimal control, and more is available online [53].

Castillo-Chavez et al. (1988,1989) have studied the role of age structure on single and multiple strain influenza models in the context of cross-immunity, a partial "somewhat" effective natural vaccine [14, 13]. Mylius et al. (2008) focus on analyzing an age structure influenza model and verifying what vaccination policy is best to implement in order to reduce the transmission of the disease. These researches proposed the introduction of vaccines at the beginning or at the peak of the infection [69]. Lee et al. (2012) introduced an age structure vaccination within an optimal control framework and addressed the question of vaccines optimality when resources are limited.

Since the dynamics of influenza are strongly correlated with age dependent mixing [13], formulating an age-structured continuous in time model that divides the

population into a finite number of age groupings.

In this thesis, we study the impact that variability in the contact matrix and mixing assumptions have on the size of an influenza outbreak and on the development of optimal vaccination strategies.

1.2 Contact Matrix

Researchers have used several methods to provide information about social contacts, some of them include the use of synthetic data, while others used population surveys [26, 28, 31, 46, 67, 96]. Lozzi et al. (2010) used an individual-based model (IBM) to compute a variety of synthetic age specific contact matrices. Their scenarios used data from a survey that collected contact data, supported by the Italian National Statistical Agency (INSA). Their sample size was 55,773 individuals, grouped into 21,075 households. Their data included socio-demographic data (e.g., school and workplace attendance, household structure, etc.) [46]. This Italian survey considered the difference between weekday and weekend information by dividing the sample in three groups (first group was asked to fill the diary on a given workday (18,085 diaries collected), the second group collected only Saturday activity (16,828), and the third only Sunday information (16,293)). Wallinga et al. (2006) collected similar data via a large cross sectional survey in the Netherlands. The total number of participants that met the criteria for their analysis was 59 % (1,819 participants of 3084). Participants were randomly selected from population records. Survey participants were asked to provide information on how many individuals they conversed with during a typical week, excluding household members. They divided the contacts by age using the following six age classes: [0-5], [6-12], [13-19], [20-39], [40-59], and $\geq$60, groupings in years.

Studies using specialized populations have also been carried out. Edmunds et al.

(2006) sampled a population of undergraduate students at the University of Warwick with the goal of estimating the number of physical contact, the stability and level of assortative sexual encounters as a function of age. The participants were asked to record the number of contacts and contact details over a period of three days. Contact constituted a physical touch without conversation to an intimate contact of sexual nature [28].

Mossong et al. 2008 carried out a large scale contact pattern survey across Europe with the goal of parametrizing the mathematical models used to design control strategies [67]. They define contact as a skin to skin physical contact (kiss or handshake) or a two way conversation, with 3 or more words, in the physical presence of another person but no physical contact.

Del Valle et al. (2007) used census data from the city of Portland Oregon as input to generate a social network of synthetic population[26]. Their goal was to estimate the age-dependent forces of infection in age-structured, compartmental models on fast spreading diseases.

1.3 Research Focus

In the second chapter of this dissertation we identify vaccination policies aimed at reducing the spread of disease. The policies are implemented via a simple model and unlimited and limited vaccine stockpiles. The focus shifts to the role of vaccines in structured age-populations in later chapters.

Publications geared towards increasing our understanding of the dynamics of a system within age structure model include [23, 56, 25, 39, 90]; These publications, though useful, used only an arbitrary number of age group divisions. Here, we test, in a limited way, age group divisions and test how well mathematical formulations like proportionate mixing does when compared to a data-driven mixing formulation[13].

Researchers have conducted regional, city or state surveys to gather information on how people interact with each other [39, 67, 96]. There are only a few comprehensive surveys, and so, Mossong et al. (2008) has become the survey to "go". Mossong et al. (2008) goal was to define and measure explicit contacts of individuals by age groups via the survey analysis carried out on eight European countries [67].

Computer based simulations aimed at tracking the behavior of the individuals on a daily basis have also been carried out using, synthetic census-generated populations [26]. Theoretical general age-specific models have been introduced and analyzed to study disease dynamics, for example proportionate mixing, preferential mixing, standard mixing, and reduced mixing [9, 13, 25, 90]. Castillo-Chavez et al.(1989), Busenberg et al. (1991), and Blythe's et al. (1991) research introduce key definitions, derivations, and properties to model mixing [9, 12, 13].

Glasser et al. (2011) tested the impact of age structured population model on disease dynamics using different contacts data sets(Mossong et al. 2008, Wallinga et al. 2006, Del Valle et al 2007); the focus on the kids-parents contact relations [39]. Towers et al. (2012) goal identified groups (based on age peers or family classification) to address where social distancing would be most effective in controlling an outbreak using an epidemic model and Mossong et al. data[90]. Del Valle et al. (2013) used mixing patterns based on both empirical and theoretical studies to determine the impact that different mixing functions have on disease spread. Their study tested several mixing assumptions, including proportionate mixing . They evaluated the role of contact data variability in an SIR type model [25]. For comparison purposes; they divided their population on four age groups albeit no group re-arrangements were tested and no distributions to construct in mixing assumptions were used to address variability. Our aim, is use, to compare the mathematical proportionate mixing approach against empirical using age-stratification to find how robust the

mathematical formulation is to misspecification in the contact matrix. As part of our comparison we will increase the number of age groups, consider only proportionate mixing approach, using point estimates and variability in the point estimates.

Also, some researchers have suggested to do some form of comparison using optimization tools, in our case we introduce the vaccination as a form of control [25]. We do a relative cost difference analysis to asses how robust proportionate mixing is when compared to the model output using directly the Mossong et al. (2008) survey contact data. The goal is to assess how accurate the proportionate mixing formulation is, and see if it is a good alternative to solving age-structure problems.

The overall objective of this thesis is to focus on the role of age stratification, mixing, resources, and cost of disease outbreaks. The thesis involves three studies:

- Study I: The analysis of influenza vaccination strategies when supply is limited
- Study II: The study of the impact of mixing patterns and assumptions on the dynamics of age structured populations
- Study III: The role of age-related contact structure and cost of an outbreak on the development of optimal vaccination control policies

The third chapter will introduce the age structure and contact matrix formulations, both making use of the survey data and contact definitions found in Mossong et al. (2008). We will compare how well the approaches do when simulating the fast spreading disease, influenza. In chapter 4, we will introduce vaccines as a form of control and compare the relative cost outputs and relative cost of infection using proportionate mixing and directly data-driven mixing models.

Chapter 2

ROLE OF VACCINATION USE STRATEGIES WHEN SUPPLY IS LIMITED

2.1 Motivation

In 2009 Health Agencies from the US declared an influenza Pandemic on june 2009 [41, 80, 97] . The fear of a drastic increase in infection and mortality due to transmission drove significant efforts to develop and distribute vaccines to protect the population [92]. Some current research allocate resources assuming that we have unlimited stockpiles of vaccines [60, 86, 91] . However only 900 million vaccine dosages were available in 2009-2010 [34, 36, 77, 83].The challenge posed by vaccine production restrictions and logistic distribution limitations (experienced over the course of this pandemic) raised a key question: What is the impact of having access only to a limited number of dosages? This work aim to shed light on finding methods to minimize the transmission of a fast spreading disease when supply is limited.

2.2 Introduction

The World Health Organization (WHO) declared an influenza A(H1N1) pandemic on June 11, 2009 [35, 19], that is, just a few months after a novel strain had been identified from an infectious individual in Mexico [95]. The subsequent spread of this new strain of H1N1 in the southern hemisphere increased the fear that a more virulent strain would return to the northern hemisphere in the fall. Hence, efforts to develop a vaccine that could be distributed prior to the return of H1N1 to the northern hemisphere became a global priority. The uncertainties associated with the magnitude and nature of this health emergency dramatically increased the demand

for a vaccine that was still in the design phase.

It soon became apparent that no more than 900 million vaccine dosages would be available in 2009-2010 [74]. Wealthy nations soon purchased most of the expected vaccine production leaving an inadequate supply available for the rest of the world. Canada for example ordered more than 50 million dosages while the USA secured 200 million dosages [81, 88]. On the other hand the WHO was able to secure only 37 millions dosages that had to be judiciously distributed to poor nations from December to February [29]. The first batches arrived in Canada on October 21 and in the USA in October 19, 2009 [73, 18]. Limited supplies were used to vaccinate primarily pregnant woman, young children, and medical personnel [7, 73, 18]. In other words, nations that had managed to secure large supplies still had to wait several months for full delivery, with most dosages arriving well within the second wave. The challenge posed by vaccine production restrictions and logistic distribution limitations (experienced over the course of this pandemic) raised several questions: What is the impact of having access only to a limited number of dosages? What is the impact of delays in accessing the available vaccine supply? What is the role of a large percentage of H1N1 asymptomatic infectious individuals?

In this note, we only address the first question but its relationship to the others is addressed briefly in the conclusion. A single-outbreak epidemiological model is used to study the impact of a limited vaccine (not 100% effective) supply on the fall 2009-wave in the northern hemisphere. It is assumed (a highly conservative assumption) that the vaccine is available from the beginning of the outbreak. The role of this vaccine is first analyzed under the assumption that the population has access to an unlimited supply. Next, the case when the vaccine supply can only protect a small proportion of the total population at risk is studied numerically. In both cases, optimal control theory is used to find the optimal vaccine-strategy (unconstrained

and constrained cases) in situations where the vaccine fails to protect a fraction of the vaccinated populations.

2.3 Pandemic Influenza Model with Vaccination Control

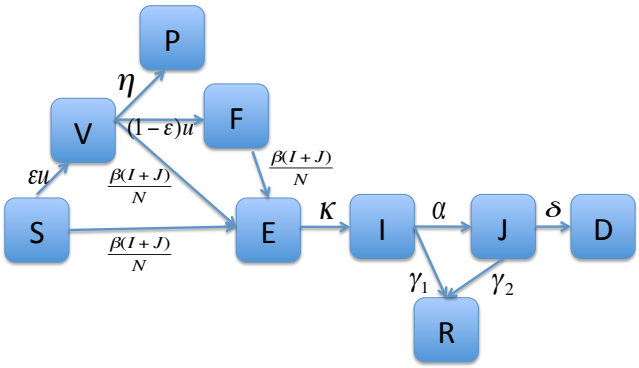

Figure 2.1: Flow Chart of the Single Outbreak Influenza Transmission Model with Vaccination

In this section, we use an existing single outbreak model [21, 23] modified through the incorporation of a control function, the time-dependent vaccination rate. Two optimal control problems are formulated and their impact on numerically derived optimal vaccination strategies is evaluated under distinct vaccine coverage scenarios and various values of the basic reproduction number R_0 [57]. The model classifies individuals as susceptible (S), effectively vaccinated (V), ineffectively vaccinated (F), protected by vaccination (P), latent (E), infectious (I), hospitalized (J), recovered (R), and dead (D) (see Figure 1). Susceptible individuals are exposed to the influenza

virus with a force of infection given by $\beta\frac{I(t)+J(t)}{N(t)}$ where β is the transmission rate and $N(t)$ denotes the total population size ($N(t) = S(t) + V(t) + F(t) + P(t) + E(t) + I(t) + J(t) + R(t)$). The control function (denoted by $u(t)$), the vaccination rate of susceptible individuals, is calculated under a particular set of assumptions. Susceptible individuals (S) go to the vaccination class at the rate $\epsilon u(t)$, where ϵ in (0,1) is a measure of the vaccine efficacy. Therefore, $(1-\epsilon)$ denotes the fraction of ineffectively (failure) vaccinated individuals per unit of time. Vaccinated individuals go to the P (perfectly protected) class at the rate η while infected individuals enter either the Recovered class $R(t)$ at the rate γ_1 or the Hospitalized class ($J(t)$) at the rate α. Individuals in the Hospitalized class enter the Recovered class at the rate γ_2 or the Death ($D(t)$) class at the rate δ [23]. These definitions and assumptions lead to the following system of nonlinear differential equations that models the dynamics of a single epidemic outbreak.

$$
\begin{aligned}
\dot{S}(t) &= -u(t)S(t) - \beta\frac{I(t)+J(t)}{N(t)}S(t) \qquad (2.1)\\
\dot{V}(t) &= \epsilon u(t)S(t) - \eta V(t) - \beta\frac{I(t)+J(t)}{N(t)}V(t)\\
\dot{F}(t) &= (1-\epsilon)u(t)S(t) - \beta\frac{I(t)+J(t)}{N(t)}F(t)\\
\dot{P}(t) &= \eta V(t)\\
\dot{E}(t) &= \beta\frac{I(t)+J(t)}{N(t)}(S(t)+V(t)+F(t)) - kE(t)\\
\dot{I}(t) &= kE(t) - (\alpha+\gamma_1)I(t)\\
\dot{J}(t) &= \alpha I(t) - (\gamma_2+\delta)J(t)\\
\dot{R}(t) &= \gamma_1 I(t) + \gamma_2 J(t)\\
\dot{D}(t) &= \delta J(t)\\
\dot{Y}(t) &= u(t)S(t)
\end{aligned}
$$

The model's basic reproduction number R_0 (using the next generation operator approach [27, 93]) is given by:

$$R_0 = \beta(\frac{1}{\alpha + \gamma_1} + \frac{\alpha}{(\alpha + \gamma_1)(\gamma_2 + \delta)}). \tag{2.2}$$

R_0 is a measure of the transmissibility of the infectious disease when the population is completely susceptible. Specifically, R_0 accounts for the average number of secondary cases generated by a primary case during his/her infectious period given that $S(0) \approx N(0)$. Typically, we have that when $R_0 > 1$ an outbreak takes place while $R_0 < 1$ indicates that an outbreak cannot be sustained. We take $S(0) = S_0$, $E(0) = E_0$, and $I(0) = I_0$ with $S_0 + E_0 + I_0 = N_0$ [$I_0 >$ and $E_0 > 0$ and $N_0 >> I_0 + E_0$]. Using the approach in [11], we derive the final size relationship

$$ln\frac{S_0}{S_\infty} < R_0[1 - \frac{S_\infty}{N_0}]$$

where the percentage of individuals recovered or dead from the infection is given by $[1 - \frac{S_\infty}{N_0}]$.

The aim of this work is to minimize the number of infected individuals over a finite time interval $[0, T]$ at a minimal cost. This outcome would be the result of implementing an optimal vaccine policy during the course of an influenza epidemic outbreak. These unconstrained (unlimited vaccines) and constrained (limited vaccines) optimal control problems are handled using the approach illustrated in [42, 50, 54, 57]. The objective functional to be minimized is therefore

$$\mathcal{F}(u(t)) = \int_0^T [I(t) + \frac{W}{2}u^2(t)]dt \tag{2.3}$$

where the control effect is modeled by a quadratic term in $u(t)$. The weight constant W is a measure of the *relative* cost of vaccination over a finite time period. The constrained optimal problem with the isoperimetric-constraint (limited vaccine) consists

of finding an optimal control function $u^*(t)$ such that

$$\mathcal{F}(u^*(t)) = min_{\Omega}\mathcal{F}(u(t)) \tag{2.4}$$

$$\int_0^T u(t)S(t)dt = B \tag{2.5}$$

where $\Omega = \{u(t) \in L^1(0,T) \| 0 \leq u(t) \leq b, t \in [0,T]\}$ and subject to System (1). The equality constraint (isoperimetric constraint) represents the total amount of vaccines available B over the time interval $[0,T]$. Constraint (5) can be reformulated in terms of the differential equation $\dot{Y}(t) = u(t)S(t)$ with the initial condition $Y(0) = 0$ and final condition $Y(T) = B$ (added to System (1)). This way of including the isoperimetric constraint lets us apply the standard Pontryagin's Maximal Principle in our search for an optimal solution of the constrained optimization problem (additional necessary conditions are derived in Appendix). The solution of the *unconstrained* optimal control problem excludes the $\dot{Y}(t)$ equation while the solution of the *constrained* problem requires the inclusion of the $\dot{Y}(t)$ equation which is equivalent to (5). The default values for initial conditions and model parameters are in Table 1. Units are per day for all rates. These baseline values are used throughout the manuscript unless otherwise indicated.

2.4 Simulations and Results

We present numerical simulations under two scenarios: the unconstrained optimal vaccination and the constrained optimal one. The first case assumes that there exists a large vaccine supply to protect the (almost) full population while the later assumes that there is limited access to the vaccine. Our focus is on understanding the effects of optimal vaccination strategies on the dynamics of pandemic influenza. The impact of such controls is evaluated under different values of the basic reproductive number, R_0, and under pre-selected levels of vaccine coverage. A sensitivity analysis is carried

out on the weight constants, on the upper bounds of the controls, and a mean vaccine efficacy in the unconstrained optimal vaccination strategy.

2.4.1 Results in the Unconstrained Vaccine Supply Case

First, the unconstrained optimal control problem is solved under two distinct transmissibility levels modeled by R_0. Figure 2 illustrates the optimal control functions computed (as a function of time) when $R_0 = 1.3$ and $R_0 = 2.0$ (top). The corresponding daily incidence of the infected class with/without vaccines is displayed in Figure 2 (bottom). The implementation of an optimal vaccine strategy reduces the number of infected cases significantly (almost no outbreak) when R_0 is low ($R_0 = 1.3$). We observe that maximum vaccination effort must be allocated at the beginning of the pandemic for both values of R_0. The optimal vaccination must be applied for a longer period of time when R_0 is high (R_0=2.0) since there are much more infected individuals (red solid curve in the top graph). In the absence of vaccines, a higher peak in the incidence of infected cases over a shorter time period is detected when $R_0 = 2.0$ (red dotted curve in the bottom graph). Even though a considerable increase in the number of vaccines is put in place, there is still an outbreak (red solid curve in the bottom graph).

2.4.2 Sensitivity Analysis

We explore the effect of changes in model parameters on the dynamics of influenza pandemics by carrying out a sensitivity analysis. Our sensitivity analysis focuses on studying the role of varying the control weight constants (W), the control upper bounds (the maximum vaccination rate), and the vaccine efficacy. The impact of these parameters on the fraction of cumulative infected cases is compared in the presence and the absence of vaccines.

Weight constant

We choose four values of the weight constants, $W = 1, 10^2, 10^4$, and 10^6 when $R_0 = 1.3$. A comparison of results implementing optimal vaccination policies is shown in Figure 3 under different weight constants on the control. Time series of optimal control functions are in the top of Figure 3 (left) while the corresponding incidence of infected individuals are in the bottom graph of Figure 3 (left). The general shapes of the control functions are similar (monotonic decreasing in time) with large changes in magnitude. As the weight constants are increased, the cost of vaccinations also increase. These changes result in increases in the height of the epidemic peak that are the result of reductions in the number of vaccines available to individuals. The impact of the weight constants on the fraction of cumulative infected cases and vaccinated cases is displayed as a function of R_0 in the right graph. Under higher vaccination coverage (almost 90% using the weight constants in the range of 10^2-10^6), the fraction of infected cases is reduced significantly (less than 10% of the total infected cases up to values of $R_0 <= 1.6$).

Upper bounds of the control

We carry out a sensitivity analysis on pre-selected control upper bounds starting from their impact on the fraction of cumulative infected cases. Four different upper bounds on controls (a priori maximum vaccination rates for susceptible individuals) are chosen in the range (0.05, 0.1, 0.2, 0.5). Figure 4 shows the results of implementing optimal vaccination controls constrained by these upper bounds when $R_0 = 1.3$. Optimal controls are graphed at the top while the corresponding incidence of infected individuals are plotted at the bottom (left). The cumulative fraction of infected and vaccinated cases are shown in the right graph of Figure 4 under different values of R_0. The larger upper bound, the better impact of vaccines in reducing the number of infected cases for all ranges of R_0 values. For example, the use of the largest

upper bound ($b = 0.5$) generates dramatic reductions in the number of infected cases, regardless of the value of R_0.

Vaccine efficacy

We explore the impact of changes in vaccine efficacy on the control functions as well as on the fraction of cumulative infected cases. Figure 5 presents the results of implementing optimal vaccination strategies using different values for vaccine efficacy ($\epsilon = 0.4, 0.6, 0.8, 1$) when $R_0 = 1.3$. The optimal controls and the corresponding incidence of infected individuals are given in Figure 5 (left). It is observed that the vaccination strategy with larger vaccine efficacy uses less number of vaccines. For example, the amount of vaccines with an efficacy, $\epsilon <= 0.4$ is almost double of those required when $\epsilon >= 0.8$. The role of vaccine efficacy variation becomes clearer when one looks at the cumulative fraction of infected cases as R_0 is varied (Figure 5 right). The optimal vaccination strategy using a higher value of vaccine efficacy ($\epsilon = 0.8$) manages to keep the fraction of infected cases under 20% for most ranges of R_0 (< 2) even though it uses less vaccines.

2.4.3 Results Under Isoperimetric Constraint

We solved a constrained (isoperimetric) optimal control problem under pre-specified isoperimetric constraints (amount of vaccines). We consider three vaccine coverage levels (15%, 30%, and 50%). The numerical method used to solve this constrained optimal control problem turned out to be quite sensitive to the levels of vaccine available. As a result of the boundary conditions in the adjoint system (from the isoperimetric constraint), convergence issues had to be addressed.

Figure 6 plots time series of the optimal vaccinations under three different vaccine coverages (15%, 30%, and 50%). These results are contrasted with those obtained in the absence of vaccines. The left graph displays the control functions as well as

the resulting incidence (infected cases) when a smaller value of the control upper bound (b=0.05) was in use. The right graph shows the results of using a larger value for the control upper bound (b=0.2). We see that maximum vaccination rates must always be implemented at the beginning until all the available vaccines are depleted in both cases. The effect of two upper control bounds on the incidence of infected cases does not seem to be significant provided that the maximum vaccination rates are put in place at the start of the pandemic and that R_0 lies in the low range (e.g $R_0 = 1.3$). The impact of vaccinations with 30% and 50% vaccine coverages on the fraction of cumulative infected cases is illustrated in Figure 7. When R_0 is low enough ($<= 1.4$), the optimal strategy with 30% still generates significant reductions ($<= 10\%$). However, the benefits of the application of controls under higher vaccine coverage ($>= 50\%$) increase as R_0 increases ($R_0 >= 1.5$).

2.5 Conclusion

An existing influenza transmission model is used to explore the impact of optimal vaccination policies under limited vaccine supplies (isoperimetric constrained model). Under a rather set of optimistic assumptions (vaccines available at the beginning of a pandemic), constrained and unconstrained optimal control problems are formulated. The constraint imposed (a limited vaccine supply) is incorporated in the isoperimetric optimal control problem through the addition of a differential equation with two "boundary" conditions.

Results suggest that both optimal vaccination policies (constrained or unconstrained) must be implemented at the maximum vaccination rate for all ranges of R_0 and at the beginning of the outbreak. There are no significant differences under the constrained or unconstrained strategies when R_0 is low ($<= 1.3$) and more in line with transmissibility estimates for seasonal influenza [22]. Hence, in some sense, this

result indicates that there is no "need" for a nation to have more than 15 % of the vaccines needed to cover its total population. The pre-selected values of the upper bounds on controls and vaccine efficacy levels have a significant impact on the final size of infected cases. Increases on the upper bound of the optimal control and the efficacy level result in a decrease in the amount of vaccines that must be administered.

In this study, the focus has been exclusively in the study of constrained and unconstrained vaccine availability scenarios at the start of an epidemic outbreak. In other words, the study has been quite limited as noted in the introduction with important critical questions that should be simultaneously addressed being ignored. The time delays associated with the process of vaccine preparation, production and delivery of a new vaccine (around 6 months) just after a new strain was identified must be factored in. The role of age-specific risk, asymptomatic individuals (which can infect others) and the impact that these individuals have through their "accidental" use of vaccine supplies, must also be considered. These ignored factors make it difficult to mitigate the impact of super-fast spreading diseases like influenza. In summary, for super-fast spreading diseases limitations in supply can make all the difference particularly when R_0 is large.

Chapter 3

WHAT IS THE EFFECT OF AGGREGATION ON DISEASE DYNAMICS

3.1 Introduction

The lack of precise information on how individuals interact with each other increases our difficulty in assessing the impact of a fast spreading diseases like influenza. Over the past 30 years scientist have suggested various ways of estimating contact rates given the limited and non-precise data. Collecting data via surveys on samples of particular populations and using synthetic data and simulations [9, 13, 43, 65, 67, 96].Contact mixing data collected through surveys is expensive, whereas synthetic data, generated via models, has built in limitations [26, 67]. Researches have attempted to connect limited survey with mathematical models in order to improve the value of contact rates estimates. The modeling variations involves mixing assumptions including proportionate mixing and data driven mixing [9, 13, 25, 43, 65, 90].

The models that used contact data *directly* from the field are referred here as empirical-data-driven mixing models, whereas the models built under theoretical mixing assumptions, are named after their standard nomenclature including proportional or preferred mixing models [39, 90].

The use of data driven mixing models have helped identify the presence of groups, within the population of interest that have drastic impact on the model-simulated disease dynamics. Glasser et al. (2012) used various mathematical formulations of mixing within a Susceptible-Infected-Recovered (SIR) infection model comparing the outcomes of theoretically driven models to those generated via empirical data-driven

mixing formulations.Towers et al.(2012) extended the models of Glasser et al. (2012) by incorporating interactions between grandparents and grandchildren; identifying the role of these interactions on the transmission of a disease like influenza.

While Glasser et al.(2012) focused on the study of the use of several data-motivated mixing formulations on incidence outcomes, Del Valle et al. (2013) instead concentrated on identifying the difference in final size and peak of epidemic infection outcomes under various theoretical mixing formulations (proportionate mixing, preferential mixing, reduced mixing, and data driven mixing). Del Valle et al. (2013) used an age-structured mixing model to test the role of different age-specific mixing assumptions (dependence on population size, average,contacts, group preference and more) on disease dynamics using simulated data. These studies used a fixed number of age groups and age-boundaries. They did not include vaccination. The impact of contacts on disease dynamics within these modeling was measured using output metrics that included final size outbreak, peak outbreak, the average number of secondary cases generated by an infected individual, and outbreak duration.

Our simulations show that this type of models are extremely sensitive to the use of age-boundaries. Therefore, we study the role of uncertainty associated with the number of age groups selected and age aggregation boundaries on disease dynamics. We develop and analyzed a mathematical framework that addresses the effect of age aggregation, a step needed in validating model results using empirical data. The focus is on the role that proportionate mixing models have in the study of influenza outbreaks. Our study begins to address the following challenges:

1. How do we develop robust age-structured models? What is the appropriate number of age-groups? How do we make such choices given the limited survey data? What impact do our aggregation choices have on the value on interventions?

2. How do we quantify the uncertainty generated by variations in the estimates of age-specific contact rates on disease patterns?

The prototype selected theoretical model, proportionate mixing in this dissertation is tested against the empirical data-driven model, generated using Mossong et al. (2008) data. We carry this analysis in the context of a simulated influenza outbreak. Contact matrices are constructed using the average number of contacts (from data) within the proportionate mixing model. The generated disease dynamics are compared to those generated when the average number of contacts data is used *directly* under no specific theoretical model (data-driven mixing model) within the influenza simulation.

3.1.1 Contact Definition

In short, we constructed a contact matrix following the information given in the study by Mossong et al. (2008) [67], a large scale survey involving Belgium, Germany, Finland, Great Britain, Italy, Luxembourg, the Netherlands and Poland, carried out from May 2005 to September 2006. The sampling procedure guaranteed that sample used in the survey was representative of the whole population, in terms of geographical spread, age, and gender. The researchers oversampled children and adolescents given their (assumed) greater role on the transmission dynamics of influenza. A contact was defined as a skin to skin physical interaction (kiss or handshake) or a two way conversation with 3 or more words in the physical presence of another person but no physical contact. The researchers gave the participants diaries to fill with the following information: employment level, sociodemographic information, level of completed education, household composition, age, and gender.

Participants were assigned a random day of the week and asked to record all the contacts they made for a period of 24 hours starting at 5 am. The contacts made that day were recorded once. The record included the location (home, work, school,

park, ice cream shop, etc.), the duration (5 mins, 5 to 15 mins, 15 mins to 1hour, 1 to 4 hours or more), and the frequency (e.g daily, almost daily, once or twice a week, once or twice a month, less than once a month, once a year or first time). Participants were asked to write down the age and gender of the person they had a contact with. If they did not know the age of the person, they were asked to provide an estimate of the person's age range. The data from this survey is used to construct a 100 x 100 contact matrix that takes into account all the relevant information from the participants and the contacts that they made in a given day. We matched the participants with their corresponding age and with the individuals that they made contacts with (including their age).

3.1.2 Reciprocity

Assumption of closed population means, within the contact matrix C_{ij}, that reciprocity holds, that is , that the total number of contacts between age group i and age group j must be equal to the total number of contacts between age group j and age group i [12, 15**?** , 90, 84, 85, 96]. We implemented reciprocity using the approach found in Medlock et al. (2009) [61]. Specifically, the elements of the contact matrix are generated by dividing the number of contacts per person within age group i with people in age group j (C_{ij}) by the proportion of the population in age group j, that is,

$$\hat{\phi}_{ij} = \frac{NC_{ij}}{N_j}$$

where N is the total population and N_j is the total population for age group j.

We correct for reciprocity by using:

$$\phi_{ij} = \frac{\hat{\phi}_{ij} + \hat{\phi}_{ji}}{2}$$

3.2 Methods

We use the age-structured SIR Model ([10, 51]) given by the following set of nonlinear differential equationsThe age structure model is as follows:

$$\begin{aligned} \dot{S}_i &= -qS_i \sum_{j=1}^{m} f_{ij} \frac{I_j}{N_j} \\ \dot{I}_i &= qS_i \sum_{j=1}^{m} f_{ij} \frac{I_j}{N_j} - \gamma_i I_i \\ \dot{R}_i &= \gamma_i I_i \end{aligned} \tag{3.1}$$

where i and j $\in \{1, 2, ..., m\}$, m are the number of age classes and f_{ij} is used for different mixing formulations. We used the following two:

$$f_{ij} = \begin{cases} C_{ij}, \text{ directly from data driven mixing} \\ C_i \bar{p}_j, \text{ under proportional mixing assumption} \end{cases} \tag{3.2}$$

where C_{ij} represents the average number of contacts that an individual in age group i with an individual in age group j has per unit of time. C_i represents the average number of contacts of an individual in age group i per unit of time [43] and p_{ij}, defined below, denotes the probability of an individual in age group i contacting an individual in age group j; It is worth observing that C_{ij} and C_i are estimated from the Mossong et al. (2008) data. Since we are dealing with proportionate mixing p_{ij} is independent of i, that is $p_{ij} = \bar{p}_j$. The parameters for this model are listed and defined in Table 3.1:

Table 3.1: Parameter Definitions

Variables	Definition	Value	Source
i	Age group index	n	[67]
S_i	Susceptible population in age group i	Varies	–
I_i	Infected population in age group i	Varies	–
R_i	Recovered class in age group i	0	–
Parameters	**Definition**	**Value**	**Source**
q	Probability of Infection given a contact	0.67	[8, 58]
$\bar{p}_j$	Proportionate mixing	Varies	[43, 67]
f_{ij}	Effective mixing rates between age group i and age group j	Varies	[67]
γ_i	Recovery Rate	$\frac{1}{1.5}$	[55]
w_i	Statistical weight adjustment $\frac{PopulationPercentage}{SamplePercentage}$	Varies	[67]
N_j	The number of people in group j	Varies	–
μ_i	Contact mean from the survey data	Varies	–
σ_i^2	Variance of contacts from the survey data	Varies	–
$\hat{\mu}_i$	Estimate of contact mean used as parameter for the TN Distribution	Varies	–
$\hat{\sigma}_i^2$	Estimate of variance used as parameter for the TN Distribution	Varies	–

A mixing matrix from empirical data [67] is estimated and used on the data driven model and the proportionate formulations [13, 38, 26, 25, 43] to generate the values used for C_{ij} and C_j, $i, j = 1, 2, ..., m$. We varied the age group divisions from two up to eight groups. For comparison purposes we concentrated on the use of proportional mixing as the 'null' model so that we are able to understand the role of variations in contact rates between the age groups under proportionate mixing and an empirically data-driven model.

3.2.1 Proportionate Mixing

The general mixing formulation [9, 12, 13, 48] satisfy the following properties:

1. $0 < p_{ij} < 1$ all i, j, and t
2. $\sum_{j=1}^{m} p_{ij} = 1$

3. $C_i p_{ij} N_i = C_j p_{ji} N_j$

$C_i p_{ij} N_i = C_j p_{ji} N_j$ accounts for the reciprocity demand; the total number of contacts made by members of group i with members of group j per unit of time must be equal to the total number of contacts made by members of group j with members of group i per unit of time.

Proportionate mixing is the particular solution given by:

$$p_{ij} = \bar{p}_j = \frac{N_j C_j}{\sum_{v=1}^{n} N_v C_v} \tag{3.3}$$

where N_j is the number of people in group j, $C_i = \sum_{j=1}^{m} C_{ij}$ is the average number of contacts that a typical individual in group i has with anyone in the population per unit of time; C_{ij} is the average number of contacts of a typical individual in group i with a typical individual of group j per unit of time [38]. We use data from the Mossong et al. (2008) study to estimate N_j and C_{ij} [67]. The data can be used to generate estimates for each of the ages, from 0 to 100 years.

3.2.2 Basic Reproductive Number

The Basic reproductive number (R_0) denotes the average number of secondary infections generated by one infected individual during his/her infections period, assuming that he/she is embedded in a fully susceptible population [13, 38, 26, 25, 43, 55]. R_0, a function of demographic and epidemiological quantities, is a measure of the transmissibility of the infectious disease in a naive population. When $R_0 > 1$ an outbreak is expected to take place while whenever $R_0 < 1$ an outbreak is unlikely to be sustained in a population. We use the methods in [94] to compute R_0, under proportionate mixing (Model 3.1). It is given by

$$R_0 = q \frac{\sum_{i \neq j}^{m} p_i C_i \prod \gamma_j}{\prod_{n=1}^{m} \gamma_n} \tag{3.4}$$

The details involved in calculating R_0 can be found in the Appendix. Estimates of R_0 under the data driven mixing formulation are presented later on.

3.2.3 Final Epidemic Size

Comparing how well proportionate mixing does against a data-driven mixing formulation must be carried out under some metrics. In this case, we use final epidemic size.The final epidemic size for our model, under proportionate mixing, is a solution of the following system:

$$\ln\left(\frac{S_i(0)}{S_i(\infty)}\right) = \sum_{j=1}^{m} q\frac{C_i p_j}{N_j \gamma_j}(S_j(0) - S_j(\infty)$$

for, i = 1,2,3,...,m

3.3 Numerical Results

The focus of this section is on assessing the relative difference in epidemic burden under a data driven vs. proportionate mixing models.

We proceed as follows: First we assume that the mixing and its variability are captured in the survey or modeled explicitly. We compute the role of mean estimates of average contact rates within each age group; secondly, we simulate the transmission dynamics of influenza, single outbreak, comparing the outcomes under the data-driven vs. the proportionate mixing formulations. We evaluate the role of aggregation by considering two to eight age groupings.

The probability of infection parameter (q) is kept constant at .3334.

Case 1: We use mean estimates and variances of the estimated average contact age-group rates using the Mossong et al. (2008) survey data; the estimates are plugged in under the assumptions that the population mixes according to proportion-

ate mixing. The variability in mean estimates was added using a truncated normal distribution of the age groups. We used our metric, final epidemic size, in calculating relative differences from the mean under proportionate mixing (using Mossong et al. data (2008)).

The relative difference is calculated by subtracting the final epidemic size outcome obtained under the model with mean estimates of average contacts of age-groups and variability in mean estimates, both under proportionate mixing. Our simulations (using 1000 samples) showed that the duration of the outbreak, on average, did not last more than 30 days in all age groups. The goal was to see how robust the age group divisions are when we considered having error in the mean estimates under proportionate mixing.

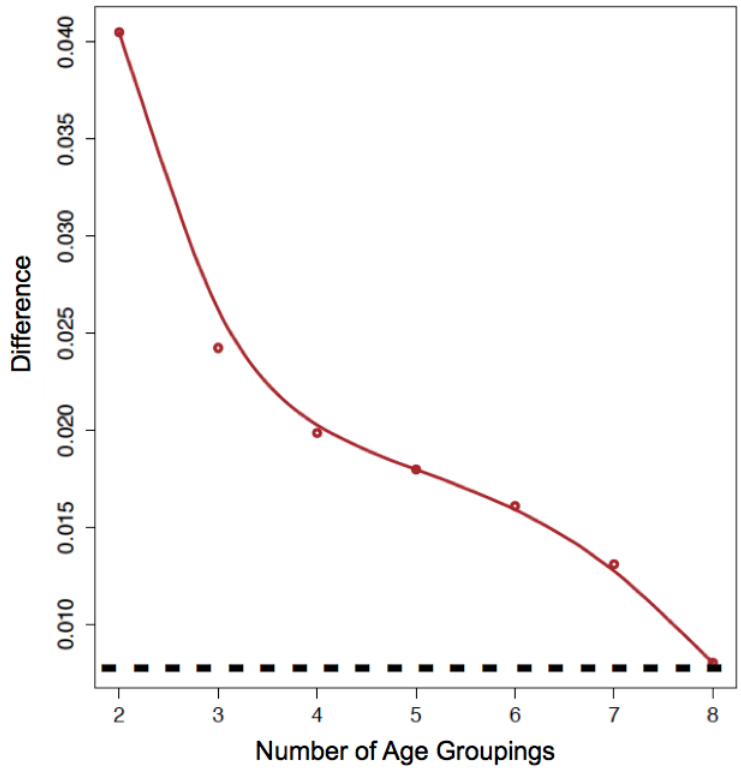

Figure 3.1: Variability in Mean Estimates vs. Mean Estimates under Proportionate Mixing.

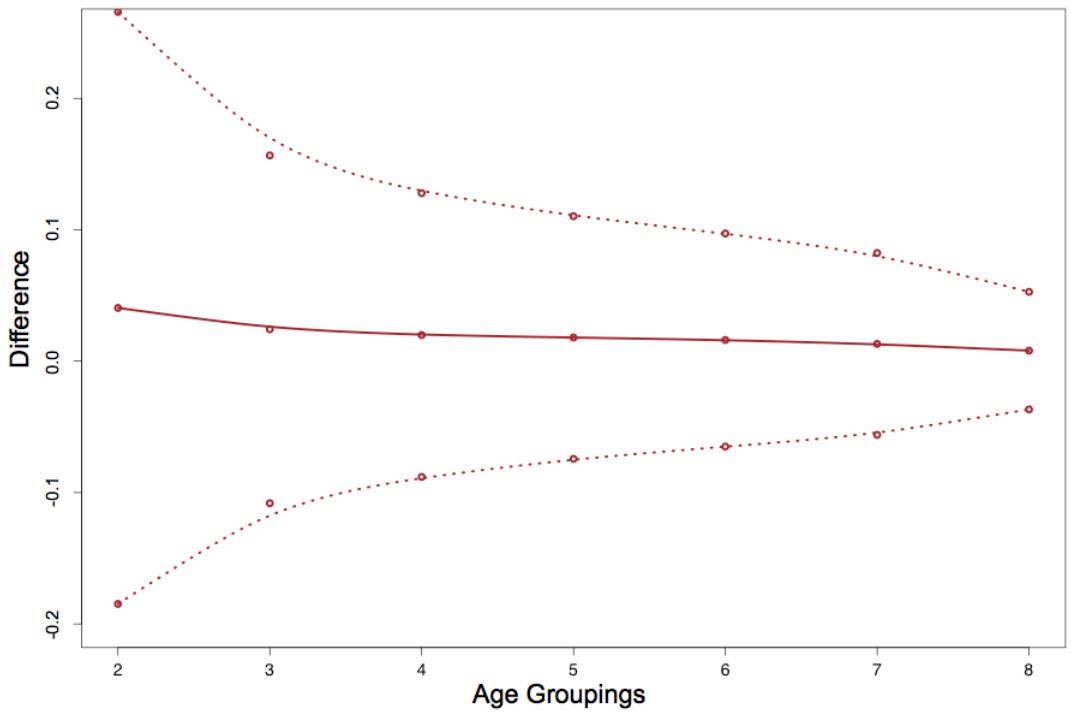

Figure 3.2: Two Standard Deviations Away from the Mean Estimates under Proportionate Mixing.

Using the relative difference in final size of the disease as a metric, we notice that the greater the number of age groupings the smaller the impact of variations on the average number of contacts (Figure 3.1). That is, increasing the number of age groupings makes the model results more robust (the difference gets closer to 0 regardless of variation on the mean estimates). Figure 3.2 shows the output when we include two curves with 2 standard deviations away from the mean estimated under proportionate mixing. In this case we notice that as we increase the number of age groupings the standard deviations decreases.

Considering variation in the age-boundaries of the groups (for example, in two age groups cases age 0-20, 21-100 vs 0-25,26-100), leads to similar qualitative outcomes to those seen in the final size relative differences simulations. We observed this same outcome as we changed from two to eight age groups. Hence, we only considered less

than 5 changes in variations on the age boundaries. As part of our future work we will consider carrying out extensive and systematic variations.

Case 2: The end goal is to see how well the proportionate mixing formulation in a SIR model with parameters from a fast spreading disease like influenza does against the same epidemic model simulated using the data driven mixing formulation, as we increase the level of aggregation (i.e. as we increase the number of age-groupings). The comparisons are carried out using relative final epidemic size difference between the outputs of both models. In this case, the baseline model for computing the relative final size difference analysis used, is the model generated via the data- driven mixing assumption. Figure 3.3 shows the data driven model vs mean estimates under proportionate mixing formulation. We notice that as we increase the number of age groupings, the difference between the final size also increases. We notice, using influenza outbreak parameters, that there exist a threshold, 5 age groupings in our simulations, after which, the relative differences in the final epidemic size (Figure 3.3) begin to increase.

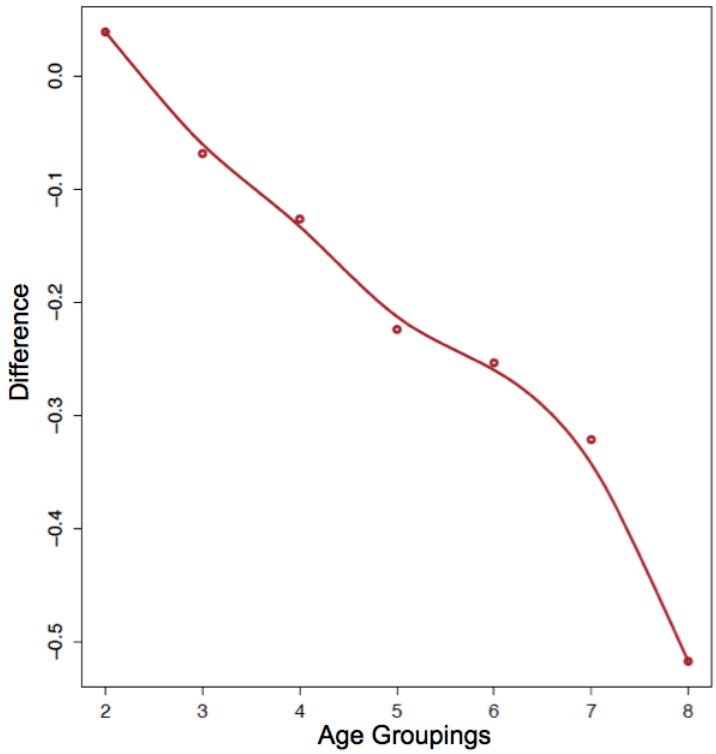

Figure 3.3: Data Driven Model vs Mean Estimates under Proportionate Mixing Formulation.

3.4 Conclusion

The goal of this chapter has been that of setting a process for evaluating the effect of aggregation on disease dynamics. It is driven by the question, how many age groupings must we consider in modeling an outbreak? And, what is the effect of changing age boundaries, number of age grouping, on disease burden? In order to answer these questions we use parameters and age structure formulations on a fast spreading disease like influenza. We have carried out our analysis using exclusively contact data from the survey of Mossong et al. (2008) and have applied it to the simulation of influenza outbreaks. We have observed that proportionate mixing does well in "matching" influenza dynamics up to a certain threshold (number of age groupings) when compared with the data-driven mixing model. We found, that the number of

age groupings threshold was five for our simulations. The use of proportionate mixing formulation may therefore be ok as long as the number of age groupings is small. Also, these results, some coinciding with the CDC group arrangements recommendations for a fast spreading disease like influenza, do tend to vary in the influenza literature publications [55]. Here, we noticed that changing age boundaries does not seem to have (in our limited simulations) a drastic impact in our results as long as the number of age groupings is small. More variations on the age groupings cut off points should be considered.

On the other hand, if we are somewhat certain that our population is well modeled by proportionate mixing then increasing the number of age groups may be ok [25]. Several researchers have used models to evaluate interventions (i.e. social distancing, vaccination) albeit they have not compared the proportionate mixing formulations to the data driven formulations [55]. The final chapter focusses on assessing the impact of mixing formulations and assumptions on the evaluation of vaccine interventions. The analysis in Chapter 4 will assist us in addressing questions such as, what age groupings should receive more vaccine interventions in order to minimize the disease burden [25, 90]? The impact of control measures on the transmission of influenza is analyzed as we vary the level of aggregation. We have included variability on the mean estimates of the average contacts of age-groups under proportionate mixing as well in our analysis.

Chapter 4

ARE THE PROPOSED VACCINATION CONTROL STRATEGIES ROBUST TO CHANGES IN AGE-STRUCTURED MIXING PATTERNS

4.1 Introduction

Mathematical formulations, like proportionate mixing, are unlikely to fully describe how individuals actually mix. These formulations are used by modelers because they capture heterogeneity in a "simple" way and when simulated via dynamical system models it may allow us to assess how disease transmission occurs and how resources should be allocated.

Formulating and understanding human behavior has been a difficult task for researchers, modelers and policy makers. The importance of understanding individual and group behavior is critical when the goal is to minimize the transmission of a fast spreading disease like influenza, which may demand the use of the maximum level of possible interventions.

Recent studies have compared the value of the models when subpopulations are assumed to mix under data-driven or proportionate mixing also or alternative mixing formulations in the study of the dynamics of a fast spreading disease like influenza [25]. These researchers have concluded that the use of certain mathematical formulations (including proportionate mixing) are useful because there is very limited contact data available. However, the robustness of these conclusions have not been addressed. There is, in fact no study in the literature, that has looked at the role of uncertainty in the population contact behavior on the dynamics of disease transmission. Since conducting a comprehensive population contact survey takes time, is costly, and the

results may become obsolete rather quickly, it has become important to study the impact of variation on estimates of contact on simulated outbreaks. We learned that comparing proportionate mixing against data-driven mixing, on age aggregated models, using parameters from an influenza, manage to produce comparable results when the number of age-groupings used is small. Our results are in agreement with those found in recent influenza publications [25].

We reviewed studies aimed at identifying best mixing models and levels of aggregation under different mitigation strategies for the spread of diseases like influenza [23, 24, 54, 56, 55]. Most of these studies failed to identify the best mitigation strategies when resources are scarce or under heterogeneous mixing [23, 55, 86, 60, 89]. Some studies have focused on the study of optimal vaccination strategies under parameter uncertainty [24, 60, 89] but have failed to consider the effect of uncertainty arising from the nature of the age-structured contacts.

In short, inaccurate applications of control policies and a lack of understanding on the role of aggregation, that is, the number of age groupings used, may lead to highly uncertain conclusions (e.g. erroneous decisions on what is the best target age group). The lack of systematic uncertainty assessment may lead to undesirable disease management outcomes that include: (1) an overestimation of disease spread, which results in greater intervention expenditures or (2) an underestimation of disease impact, which results in greater transmission and, in some cases, greater disease mortality.

The goals of Chapter 4 include: How well does proportionate mixing do against standard data in identifying the "best" disease vaccination control interventions policies? and, would we arrive at the same control policy recommendations when the age aggregation scheme is changed?

4.2 Methods

Two contact matrices, based on data driven mixing and proportionate mixing assumptions [13, 43, 9], were generated using survey data [67]. They were used within the SIR age-structured model below, that includes the vaccination of a fraction of the susceptible individuals. Specifically, the model used is given by the following nonlinear system,

$$\begin{aligned} \dot{S}_i &= -u_i S_i - q S_i \sum_{j=1}^{m} f_{ij} \frac{I_j}{N_j} \\ \dot{I}_i &= q S_i \sum_{j=1}^{m} f_{ij} \frac{I_j}{N_j} - \gamma_i I_i - \sigma_i I_i \\ \dot{R}_i &= u_i S_i + \gamma_i I_i \end{aligned} \tag{4.1}$$

where i and j $\in \{1, 2, ..., m\}$ and m are the number of age groupings; f_{ij} models the mixing formulations depending on its expression as given below:

$$f_{ij} = \begin{cases} C_{ij}, \text{ directly from data driven mixing} \\ C_i \bar{p}_j, \text{ under proportional mixing} \end{cases}$$

Here, we are imputing our estimates for C_{ij} and C_i. The susceptible individuals that get vaccinated move to the recovered class at the per capita rate u_i (referred as the control throughout this chapter). The transmission potential of the pathogen is a function of the population contact matrix (f_{ij}) and the probability of infection per "approximate" contact, q (kept constant).

The comparison between proportionate and the data driven mixing formulations is made using the total cost of disease burden (infection and vaccination) as a proxy. Total cost is computed using direct (implementation of vaccination) and indirect cost (number of cases generated during the outbreak). The total cost is used on the objective function that defines the optimal control problem. It is used to identify optimal interventions (control) rates.

4.2.1 Optimal Control Framework

The aim of this work is to contrast the differences between data-driven and proportionate mixing models via their role in minimizing the total cost. The framework identifies optimal intervention policies that minimizes the number of infected individuals over a finite time interval $[0, T]$. The results were obtained for unconstrained (unlimited vaccines) optimal control problems [32, 56, 55, 57, 78, 86]. The objective functional of this framework, to be minimized is given by,

$$\min_{u_i} \int_0^T \sum_{i=1}^{m} (I_i + u_i^2 \frac{W}{2}) dt \tag{4.2}$$

The optimal control framework is (4.2) subject to (4.1). The first step is to show analytically that for a given initial condition there exists a unique solution to the optimal control problem.

The goal of an optimal vaccination strategy is to minimize the objective functional given constraint equations and the system of equations (the integrand of (4.2) is a convex function of u and the state system satisfies the *Lipshitz* property with respect to the state variables). The existence of optimal controls is guaranteed by standard control theory results [32, 56, 55, 57, 78, 86]. The necessary conditions, that there exists an optimal solution, are derived using Pontryagin's Maximum Principle [79]. This principle converts the optimal control framework into minimizing the Hamiltonian (H) given by:

$$
\begin{aligned}
H &= \sum_{i=1}^{m}[I_i(t) + \frac{W_i}{2}u_i^2(t)] \\
&+ \sum_{i=1}^{m}\lambda_{S_i}[-u_iS_i - qS_i\sum_{j=1}^{m}f_{ij}\frac{I_j}{N_j}] \\
&+ \sum_{i=1}^{m}\lambda_{I_i}[qS_i\sum_{j=1}^{m}f_{ij}\frac{I_j}{N_j} - \gamma_iI_i - \sigma_iI_i] \\
&+ \sum_{i=1}^{m}\lambda_{R_i}[u_iS_i + \gamma_iI_i]
\end{aligned} \tag{4.3}
$$

where m is the number of age groupings.

From this Hamiltonian and Pontryagin's Maximum Principle, we obtain

Theorem: There exist the optimal control $u^*(t)$ and corresponding state solutions, $X^*= (S_i^*, I_i^*, R_i^*)$ on 4.2 over Ω. In order for the previous statement to be true, it is necessary that there exist adjoint variables $\lambda_i(t)$ such that Equations 4.4 and 4.5 are satisfied:

$$
\begin{aligned}
-\dot{\lambda}_{S_i(t)} = -\frac{\partial H}{\partial S_i(t)} &= -[(-u_i(t) - q_iS_i(t), \sum_{j=1}^{m}f_{ij}\frac{I_j(t)}{N_j})\lambda_{S_i(t)} \\
&\quad + (q_iS_i(t)\sum_{j=1}^{m}f_{ij}\frac{I_j(t)}{N_j})\lambda_{I_i(t)} + u_i(t)\lambda_{R_i(t)}] \\
-\dot{\lambda}_{I_i(t)} = -\frac{\partial H}{\partial I_i(t)} &= -[1 - (q_iS_i(t)\frac{f_{ii}}{N_i})\lambda_{S_i(t)} + (q_iS_i(t)\frac{f_{ii}}{N_i} - \gamma_i - \delta_i)\lambda_{I_i(t)} \\
&\quad + \gamma_i\lambda_{R_i(t)} + \delta_i\lambda_{D_i(t)}] \\
-\dot{\lambda}_{R_i(t)} = -\frac{\partial H}{\partial R_i(t)} &= 0
\end{aligned} \tag{4.4}
$$

H is minimized with respect to the control (at the optimal control). We differentiate H with respect to u on the set Ω and arrive at the following optimality condition:

$$
\begin{aligned}
\frac{\partial H}{\partial u_i} &= u_iW - \lambda_{S_i}S_i = 0 \\
u_i* &= \lambda_{S_i}\frac{S_i}{W}
\end{aligned} \tag{4.5}
$$

4.2.2 Cost of Disease

Given that we want to test the effectiveness of proportionate mixing under vaccination control, we make use of the minimal cost of the disease (infection and vaccination) output used in the optimal control formulation. Specifically , we use: $\int_0^T \sum_{i=1}^m (I_i^*(t) + (u_i^*)^2(t)\frac{W}{2})dt$, which will give us the total cost of infection for a given data set. In order to assess the effect of proportionate mixing formulation under vaccination (control) we follow these steps:

1. Compute the minimum cost under the for the Data driven mixing assumption (DD)
   $\int_0^T \sum_{i=1}^m (I_i^{DD,*}(t) + (u_i^{DD,*})^2(t)\frac{W}{2})dt$

2. Compute the minimum cost under the proportionate mixing assumptions (PM)
   $\int_0^T \sum_{i=1}^m (I_i^{PM,*}(t) + (u_i^{PM,*})^2(t)\frac{W}{2})dt$

3. Compare how well the model with proportionate mixing assumption did against the model with the data driven mixing assumptions by comparing their respective minimum cost.

In order to follow through the comparison outlined in the previous step, we do the following:

1. Use the control solution ($\hat{u}_i^{PM,*}(t)$) obtained under the proportionate mixing assumptions and apply it to the Data Driven dynamic epidemic model to compute $I_i^{DD,*}(t)$

2. Calculate the Cost of an imperfect control $\int_0^T \sum_{i=1}^m (I_i^{DD,*}(t) + (u_i^{PM,*})^2(t)\frac{W}{2})dt$

- $u_i^{PM,*}(t)$ is the control under Proportionate Mixing assumptions
- $I_i^{DD,*}(t)$ lnfected individuals using Data Driven Mixing State equations with Proportionate Mixing control

3. Calculate Relative Cost difference

- $\frac{\int_0^T \sum_{i=1}^m (I_i^{DD,*}(t)+(u_i^{DD,*})^2(t)\frac{W}{2})dt - \int_0^T \sum_{i=1}^m (I_i^{PM,*}(t)+(u_i^{PM,*})^2(t)\frac{W}{2})dt}{\int_0^T \sum_{i=1}^m (I_i^{DD,*}(t)+(u_i^{DD,*})^2(t)\frac{W}{2})dt}$

We analyze the relative cost difference of the data driven model against the one generated, via the proportionate mixing model, as we increase the number of age stratification (groupings) from 2 to 6 age groupings. We use the recommended age group structured from the Center of Disease and Control for vaccination for influenza like disease. For comparison purposes, we also incorporate age group division (grouping) found in the literature [55].

4.3 Results

We want to find out how well the proportionate mixing model captures the dynamics of an outbreak as we increase the level of age structure(groupings) under vaccination. We evaluate our results using relative cost differences between the data-driven mixing and the proportionate mixing models [56, 55].Seeing wether or not variation of age boundaries (cut off points) in the age structure grouping changes our results, is important. The age structure grouping recommendations from the CDC for influenza vaccination and the age structure grouping from Lee et al. (2012) were used to apply the age aggregation cut offs.

The results below are collected using data driven mixing (Mossong or M;creating subgroups of contacts within the contact matrix under the data driven mixing formulation), proportionate mixing (SDPM; mean contacts with mixing rates proportional

to both contacts in the group and population of the group), and variation in average contacts under proportionate mixing (TNPM; mean contacts with variability and mixing rates proportion to both group-contacts and group population).

The results below are collected in two subsections. The first collects the output values generated using the CDC framework; the second set of results collects the results generated using Lee et al. (2012) age group framework. It is worth nothing that within each framework, we present the relative cost difference of using the proportionate mixing under standard data (SDPM) and including variability or error in the proportionate mixing formulation (TNPM).

4.3.1 CDC Age Group Framework

The graph below shows the relative cost difference when using Mossong vs. SDPM and mean TNPM contact matrix set up. We notice that as we increase the number of age groups the relative difference increases in both cases. The relative difference on the Mossong vs SDPM output increases significantly after 4 age groups. In the case of Mossong vs mean TNPM we notice there is a higher relative cost difference in every age group division. The difference also increases as we increase the number of age groups (Figure 4.1). We also include a trend line with the relative difference plots.

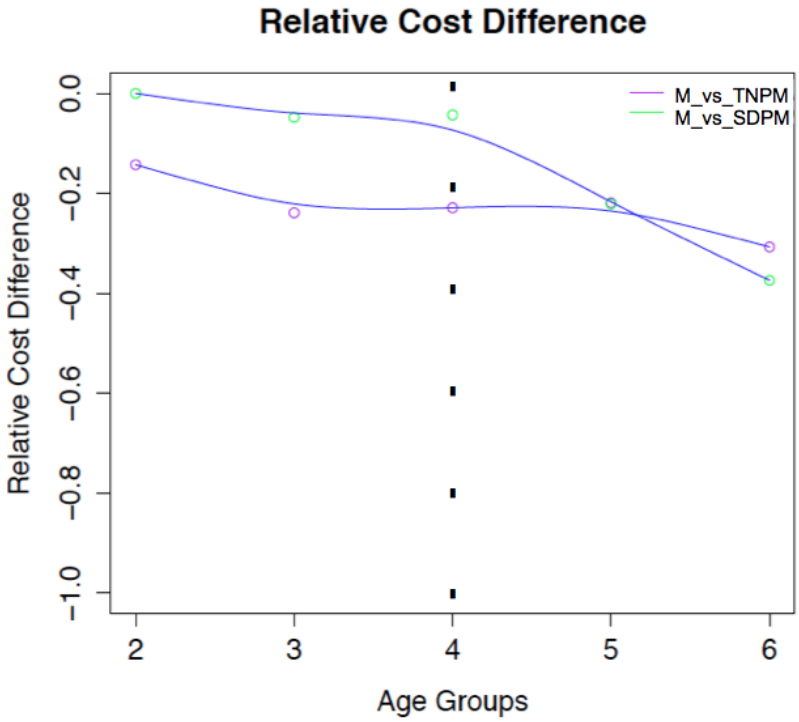

Figure 4.1: Relative Cost Difference using the CDC Framework.

We observe the vaccination policies must be implemented at a maximum rate from the very beginning of the transmission of the epidemic (Figure 4.2) as expected. We observed similar results regardless of the number of age groupings, 2 to 6. Here, we present the output of the control when implementing different type of mixing assumptions (data driven mixing = Mossong, Mean estimates = SDPM, and variability on the mean estimates = TNPM)

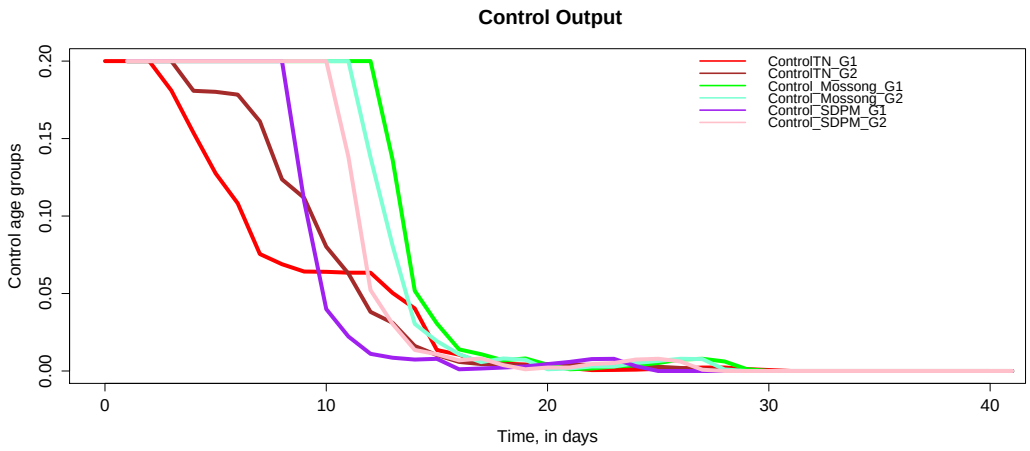

Figure 4.2: Output of the Control Over Time used to Minimize the Infected Individuals: 2 Age Groups Example.

The output of the incidence curve is shown in Figure 4.3. We notice that the incidence output is qualitatively similar when using proportionate vs. using the data driven mixing (Figure 4.3).

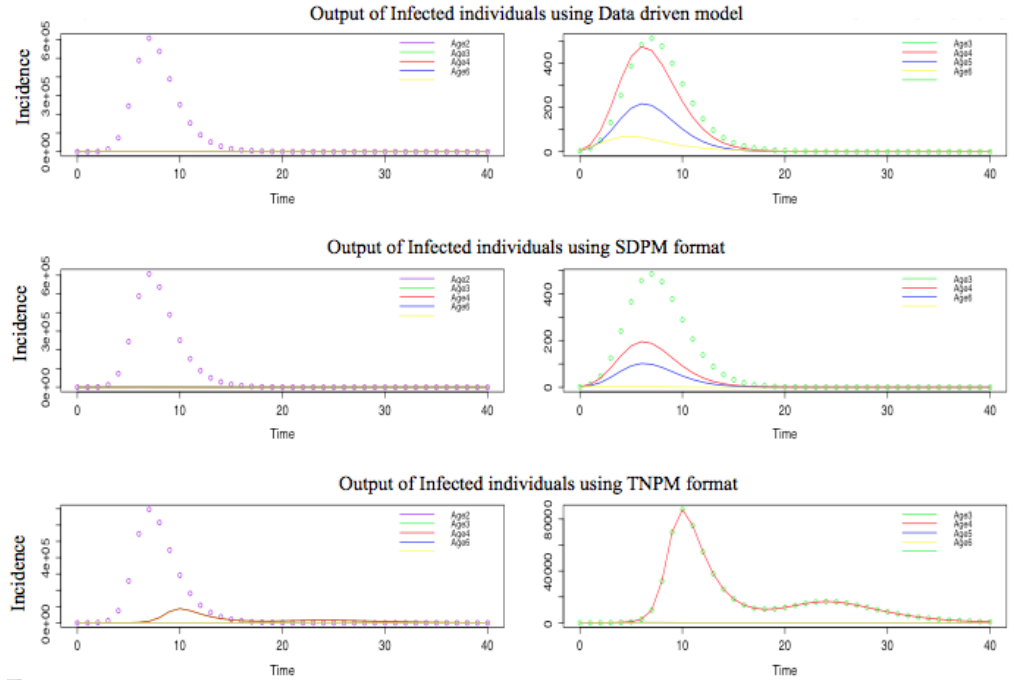

Figure 4.3: Output of the Incidence over Time of Infected Individuals from Age Group 1 to 6 using 3 Different Approaches.

4.3.2 Lee et al. Age Group Framework

Figure 4.4 shows the relative cost difference when we use the age grouping arrangement as per Lee et al. (2012). The relative cost difference in both data driven and proportionate mixing cases increases as we increase the number of age groups. The difference using the SDPM vs Mossong regroup data is smaller in the first 4 age groups divisions.

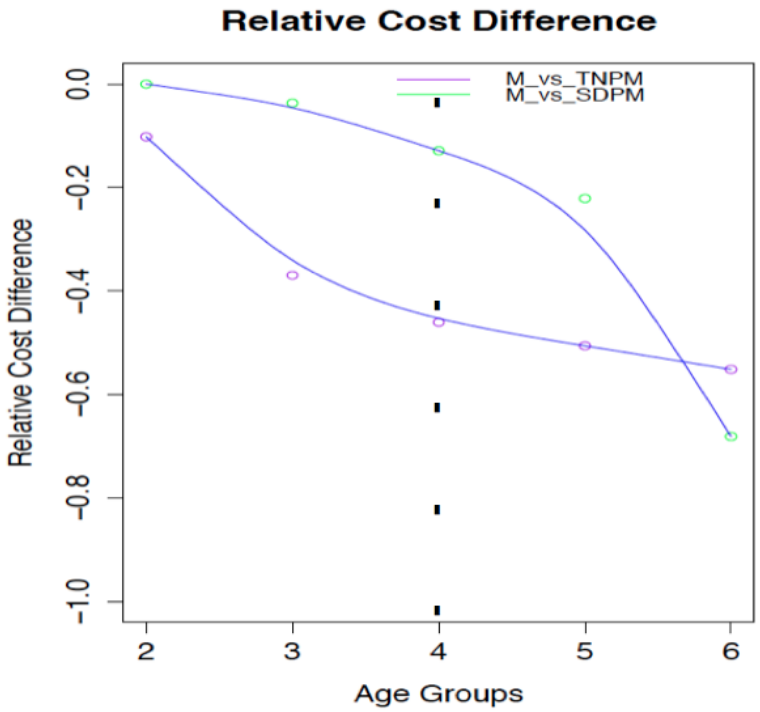

Figure 4.4: Relative Cost difference .

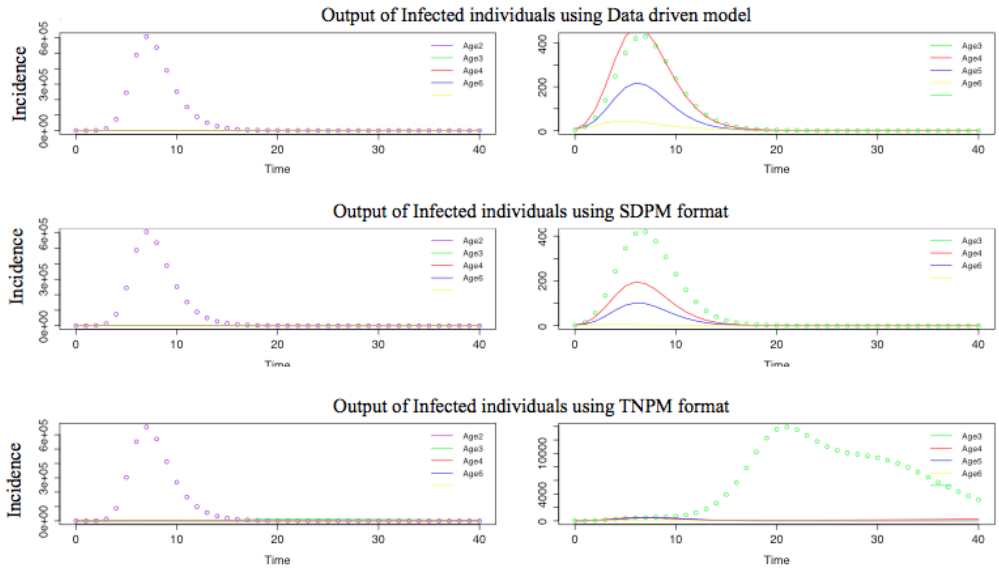

Figure 4.5: Output of the Incidence over Time of Infected Individuals from Age Group 1 to 6 using 3 Different Approaches.

We compared both outcomes and noticed that regardless of the difference in age group cut off points they both have the relative cost difference consistent until 3 age groupings. After this threshold value the trends changes significantly. This could suggest a cut off point of 3 age groupings when implementing vaccine control measures for a fast spreading disease like influenza.

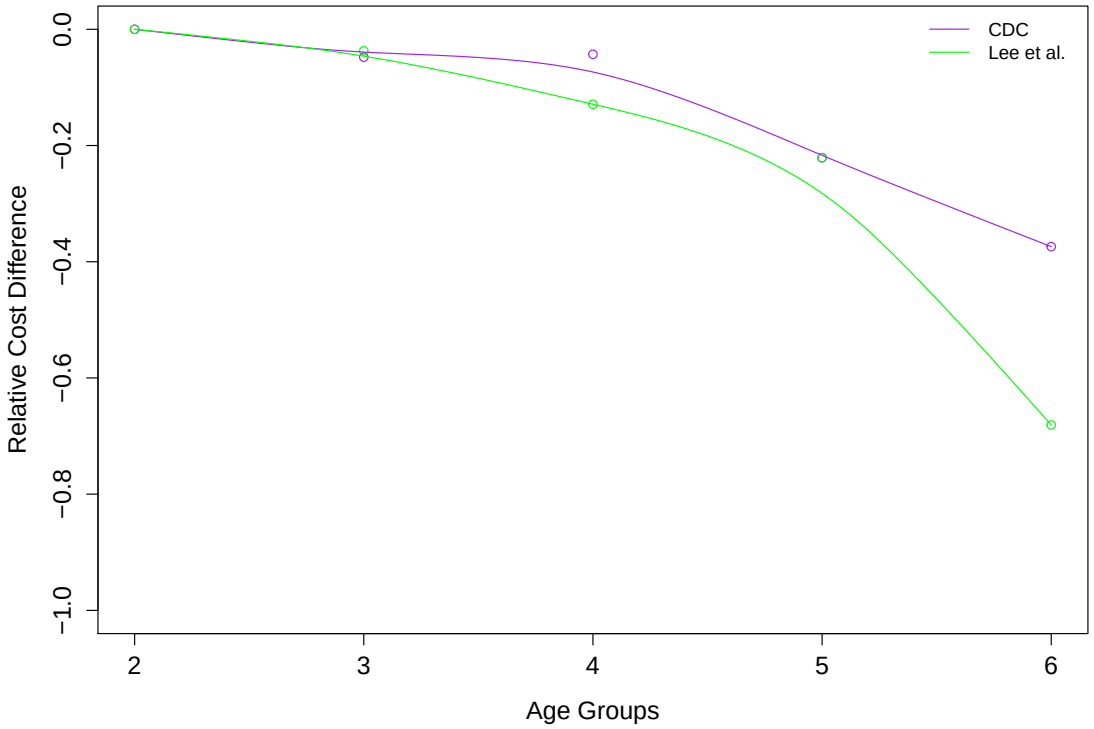

Figure 4.6: Relative Difference Data Driven Contact Mixing vs Mean Estimates in Proportionate Mixing using CDC and Lee et al. (2012) Framework.

4.4 Conclusion

This study has focused on finding how robust the mathematical formulation is to misspecification in the contact matrix; that is, how well models under the proportionate mixing do when compared to models under data driven mixing in the absence or presence of controls (vaccination). The comparisons where carried out when optimal intervention rates that guaranteed minimal total cost of intervention. Finding out wether or not a cut off on the number of age-groupings played a role on the relative cost of the implementation of a vaccine has been part of our analysis. We also study wether or not policy recommendations for age group structured model differed when compared to a non age group structured models. We noticed that age structure and

non age-structured models led to similar vaccination policy recommendation, namely that the majority of resources must be implemented at the beginning of the transmission of a fast disease, like influenza. The use of different age groupings; one following CDC age-grouping recommendations and based on the literature, did not have a significant impact on the total cost of infection up to a particular threshold. The use of 3 age groupings still provided acceptable relative cost difference independently of the age group framework used. Hence, we see that proportionate mixing works well under different age grouping as long as we keep the age groupings below a threshold. Naturally, the threshold may change if we use a different data set. We recommend checking several age grouping cut off points to see when the use of proportionate mixing leads to possibly unacceptable results.

The results on this chapter when assumed unlimited vaccination resources. As we noted in the second chapter (Vaccination strategies when supply is limited) we are planning to incorporate the limited constraint (isoperimetric optimal control problem) in this setting. In our second chapter we identified the range of convergence, however when dealing with age structure models, increasing the number of age groups, made it impossible to identify the convergence range. This challenge will be addressed in future work.

Overall it can be argued that using a population that follows proportionate mixing will be robust to mis- specifications up to a certain critical number of age group divisions, what would this critical value be will depend on the data, model framework, and age-boundaries.

Chapter 5

IMPLICATIONS FROM DATA-DRIVEN MECHANISTIC MODELING APPROACH AND MANAGEMENT RESOURCE ALLOCATIONS TO REDUCE IMPAT OF FAST SPREADING DISEASE

5.0.1 Goal of Thesis

Dynamical models are often used to evaluate interventions for control of infectious diseases in a population. Population's mixing/contact data are of great importance if the goal is to parametrize age structured models or if the aim is to look at age-specific control policies within some realistic scenarios. Contact data is difficult to obtain and most often not available. Hence, modelers have constructed age-dependent mathematical formulations like proportionate mixing that allows them to estimate mixing patterns with limited contact data and, consequently, studies that quantify disease dynamics under these mixing models often the underestimation or overestimation disease burden. In this thesis, we investigate the robustness of misspecification in age-dependent contact matrices (via a mathematical formulation like proportionate mixing) on the transmission dynamics and control of a communicable disease. We compare proportionate and data-driven mixing in the context of influenza outbreaks and vaccination in order to identify optimal vaccination strategies when supply is limited. We also look at the impact of contact estimates in mixing patterns and variations on disease disease dynamics within age structure populations. We also study the role of age-related contact structures and cost within an outbreak when vaccination is available.

5.0.2 Review of Results

The second chapter addresses the optimal distribution of vaccination resources needed to minimize the spread of a disease, particularly, when resources are limited. We find that vaccination resources must be implemented at the beginning of an outbreak regardless of stockpile size. We further conclude that whenever transmission of the disease is low, vaccinating may not have a significant effect in reducing the number of infected individuals. Results are sensitive to mixing patterns in the population i.e., the way individuals interact with each other and their rate of contacts. Hence, we also explore the role of age-structured mixing on disease transmission dynamics. There are rather few published papers that attempt to collect comprehensive data from surveys on the number of contacts that an individual has on a particular day. Several papers introduce mathematical formulations on ways that individuals interact and their impact on disease dynamics. Proportionate mixing on the dynamic results are compared to those obtained when the survey data was used directly. The relative difference in final epidemic size when the number of age groups (from 2 to 8 age-groups) are varied are also compared.

Our results suggest that the relative difference in final outbreak size for models is low (with increases in number of age groups in the models) when the number of age groups is small.

We develop an optimal control framework (expansion from the second chapter) to include age structure and find that the maximum amount of vaccines available must be distributed at the beginning of an outbreak regardless of the number of the age groupings considered; the same outcome coming from non age-structured models. Of course, how the vaccine must be distributed within need not be uniform.

We compare the relative cost of between the output of vaccination data-driven

model and proportionate mixing. The results show that the relative cost difference will increase as we increase the number of age groupings. Further, significant difference are observed when the number of age-groupings increases. Here, using a four age group (or lower) under proportionate mixing produces accurate results on to the amount of control needed to restrain the spread of an influenza like disease spread.

Overall proportionate mixing is a useful formulation, when there is a lack of information or mis-specification on contact patterns (contact matrix) as long as the age-grouping used is low.

5.0.3 Caveats and Future Work

This work should be expanded to consider the implications associated with the limitations of our current study, for example, verifying if similar results are generated when we consider a limited vaccine stockpile or other forms of control, like social distancing. Also, the results must be validated and compared using various data sets (ie. Episims [25]). Other potential expansion may be needed to study the role of the dominant eigenvalue of the contact matrix on the disease dynamics as we increase the number of age groups. We should also consider alternative mathematical formulations in constructing contact matrices like, preferential mixing and reduced mixing.
We plan to use bootstrapping techniques to examine model outcomes when contact patterns are varied. Finally, we plan to consider testing this technique on various fast spreading diseases.

REFERENCES

[1] Defense Health Agency. Questions and answers influenza - pandemic.

[2] Linda JS Allen and Amy M Burgin. Comparison of deterministic and stochastic sis and sir models in discrete time. *Mathematical biosciences*, 163(1):1–33, 2000.

[3] Linda JS Allen and P Van Den Driessche. The basic reproduction number in some discrete-time epidemic models. *Journal of Difference Equations and Applications*, 14(10-11):1127–1147, 2008.

[4] Susan J Baigent and John W McCauley. Influenza type a in humans, mammals and birds: determinants of virus virulence, host-range and interspecies transmission. *Bioessays*, 25(7):657–671, 2003.

[5] Michael G Baker, Craig N Thornley, Clair Mills, Sally Roberts, Shanika Perera, Julia Peters, Anne Kelso, Ian Barr, Nick Wilson, et al. Transmission of pandemic a/h1n1 2009 influenza on passenger aircraft: retrospective cohort study. *Bmj*, 340, 2010.

[6] Frank G Ball and Owen D Lyne. Optimal vaccination policies for stochastic epidemics among a population of households. *Math Biosci*, 177-178:333–54, 2002.

[7] Nicole E Basta, Dennis L Chao, M Elizabeth Halloran, Laura Matrajt, and Ira M Longini, Jr. Strategies for pandemic and seasonal influenza vaccination of schoolchildren in the united states. *Am J Epidemiol*, 170(6):679–86, Sep 2009.

[8] Faina Berezovsky, Georgy Karev, Baojun Song, and Carlos Castillo-Chavez. A simple epidemic model with surprising dynamics. *Math Biosci Eng*, 2(1):133–52, Jan 2005.

[9] SP Blythe, C Castillo-Chavez, JS Palmer, and M Cheng. Toward a unified theory of sexual mixing and pair formation. *Mathematical biosciences*, 107(2):379–405, 1991.

[10] Fred Brauer. Mathematical epidemiology is not an oxymoron. *BMC Public Health*, 9(Suppl 1):S2, 2009.

[11] Fred Brauer, Zhilan Feng, and Carlos Castillo-Chavez. Discrete epidemic models. *Math Biosci Eng*, 7(1):1–15, Jan 2010.

[12] S Busenberg and C Castillo-Chavez. A general solution of the problem of mixing of subpopulations and its application to risk- and age-structured epidemic models for the spread of aids. *IMA J Math Appl Med Biol*, 8(1):1–29, 1991.

[13] C Castillo-Chavez, H W Hethcote, V Andreasen, S A Levin, and W M Liu. Epidemiological models with age structure, proportionate mixing, and cross-immunity. *J Math Biol*, 27(3):233–58, 1989.

[14] Carlos Castillo-Chavez. *Cross-immunity in the dynamics of homogenous and heterogeneous populations*, volume 87. Mathematical Sciences Institute, Cornell University, 1987.

[15] Carlos Castillo-Chavez and Shu-Fang Hsu Schmitz. The evolution of age-structured marriage functions: It takes two to tango. In *Structured-Population Models in Marine, Terrestrial, and Freshwater Systems*, pages 533–553. Springer, 1997.

[16] Carlos Castillo-Chavez and Abdul-Aziz Yakubu. Discrete-time sis models with complex dynamics. *Nonlinear Analysis: Theory, Methods & Applications*, 47(7):4753–4762, 2001.

[17] Carlos Castillo-Chávez and Abdul-Aziz Yakubu. Discrete-time sis models with simple and complex population dynamics. *IMA VOLUMES IN MATHEMATICS AND ITS APPLICATIONS*, 125:153–164, 2002.

[18] CDC. The 2009 h1n1 pandemic: Summary highlights, april 2009-april 2010.

[19] Margaret Chan. World now at the start of 2009 influenza pandemic. 2009.

[20] Jiezhong Chen, Yi-Mo Deng, et al. Influenza virus antigenic variation, host antibody production and new approach to control epidemics. *Virol J*, 6(30):3, 2009.

[21] G Chowell, C E Ammon, N W Hengartner, and J M Hyman. Transmission dynamics of the great influenza pandemic of 1918 in geneva, switzerland: Assessing the effects of hypothetical interventions. *J Theor Biol*, 241(2):193–204, Jul 2006.

[22] G Chowell, M A Miller, and C Viboud. Seasonal influenza in the united states, france, and australia: transmission and prospects for control. *Epidemiol Infect*, 136(6):852–64, Jun 2008.

[23] Gerardo Chowell, Cécile Viboud, Xiaohong Wang, Stefano M Bertozzi, and Mark A Miller. Adaptive vaccination strategies to mitigate pandemic influenza: Mexico as a case study. *PLoS One*, 4(12):e8164, 2009.

[24] Damian Clancy and Nathan Green. Optimal intervention for an epidemic model under parameter uncertainty. *Math Biosci*, 205(2):297–314, Feb 2007.

[25] Sara Y Del Valle, J M Hyman, and Nakul Chitnis. Mathematical models of contact patterns between age groups for predicting the spread of infectious diseases. *Math Biosci Eng*, 10(5-6):1475–97, 2013.

[26] Sara Y Del Valle, James M Hyman, Herbert W Hethcote, and Stephen G Eubank. Mixing patterns between age groups in social networks. *Social Networks*, 29(4):539–554, 2007.

[27] O Diekmann and JAP Heesterbeek. *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation. 2000.* Wiley, 2000.

[28] W J Edmunds, G Kafatos, J Wallinga, and J R Mossong. Mixing patterns and the spread of close-contact infectious diseases. *Emerg Themes Epidemiol*, 3:10, 2006.

[29] Martin Enserink. The challenge of getting swine flu vaccine to poor nations.

[30] Eli P Fenichel, Carlos Castillo-Chavez, MG Ceddia, Gerardo Chowell, Paula A Gonzalez Parra, Graham J Hickling, Garth Holloway, Richard Horan, Benjamin Morin, Charles Perrings, et al. Adaptive human behavior in epidemiological models. *Proceedings of the National Academy of Sciences*, 108(15):6306–6311, 2011.

[31] Neil M Ferguson, Matt J Keeling, W John Edmunds, Raymond Gani, Bryan T Grenfell, Roy M Anderson, and Steve Leach. Planning for smallpox outbreaks. *Nature*, 425(6959):681–685, 2003.

[32] WH Fleming and RW Rishel. Deterministic and stochastic optimal control. 1975.

[33] Staff: Centers for Disease Control and Prevention. Cdc resources for pandemic flu.

[34] Carlos Franco-Paredes, Peter Carrasco, and Jose IS Preciado. The first influenza pandemic in the new millennium: lessons learned hitherto for current control efforts and overall pandemic preparedness. *Journal of immune based therapies and vaccines*, 7(1):2, 2009.

[35] Christophe Fraser, Christl A Donnelly, Simon Cauchemez, William P Hanage, Maria D Van Kerkhove, T Déirdre Hollingsworth, Jamie Griffin, Rebecca F Baggaley, Helen E Jenkins, Emily J Lyons, et al. Pandemic potential of a strain of influenza a (h1n1): early findings. *science*, 324(5934):1557–1561, 2009.

[36] Martin Friede, Laszlo Palkonyay, Claudia Alfonso, Yuri Pervikov, Guido Torelli, David Wood, and Marie Paule Kieny. Who initiative to increase global and equitable access to influenza vaccine in the event of a pandemic: supporting developing country production capacity through technology transfer. *Vaccine*, 29:A2–A7, 2011.

[37] Raymond Gani, Helen Hughes, Douglas Fleming, Thomas Griffin, Jolyon Medlock, and Steve Leach. Potential impact of antiviral drug use during influenza pandemic. *Emerg Infect Dis*, 11(9):1355–62, 2005.

[38] G P Garnett and R M Anderson. Sexually transmitted diseases and sexual behavior: insights from mathematical models. *J Infect Dis*, 174 Suppl 2:S150–61, Oct 1996.

[39] John Glasser, Zhilan Feng, Andrew Moylan, Sara Del Valle, and Carlos Castillo-Chavez. Mixing in age-structured population models of infectious diseases. *Math Biosci*, 235(1):1–7, Jan 2012.

[40] Paula A González-Parra, Sunmi Lee, Leticia Velázquez, and Carlos Castillo-Chavez. A note on the use of optimal control on a discrete time model of influenza dynamics. *Math Biosci Eng*, 8(1):183–97, Jan 2011.

[41] Kathy Hancock, Vic Veguilla, Xiuhua Lu, Weimin Zhong, Eboneé N Butler, Hong Sun, Feng Liu, Libo Dong, Joshua R DeVos, Paul M Gargiullo, et al. Cross-reactive antibody responses to the 2009 pandemic h1n1 influenza virus. *New England Journal of Medicine*, 361(20):1945–1952, 2009.

[42] Elsa Hansen and Troy Day. Optimal control of epidemics with limited resources. *J Math Biol*, 62(3):423–51, Mar 2011.

[43] Herbert W Hethcote. An age-structured model for pertussis transmission. *Mathematical biosciences*, 145(2):89–136, 1997.

[44] Kyung-Wook Hong, Hee Jin Cheong, Won Suk Choi, Jacob Lee, Seong-Heon Wie, Ji Hyeon Baek, Hyo Youl Kim, Hye Won Jeong, and Woo Joo Kim. Clinical courses and outcomes of hospitalized adult patients with seasonal influenza in korea, 2011–2012: Hospital-based influenza morbidity & mortality (himm) surveillance. *Journal of Infection and Chemotherapy*, 20(1):9–14, 2014.

[45] Premier Safety Institute. Epidemic and pandemic influenza and respiratory illness.

[46] Fabrizio Iozzi, Francesco Trusiano, Matteo Chinazzi, Francesco C Billari, Emilio Zagheni, Stefano Merler, Marco Ajelli, Emanuele Del Fava, and Piero Manfredi. Little italy: an agent-based approach to the estimation of contact patterns-fitting predicted matrices to serological data. *PLoS Comput Biol*, 6(12):e1001021, 2010.

[47] Marika K Iwane, Kathryn M Edwards, Peter G Szilagyi, Frances J Walker, Marie R Griffin, Geoffrey A Weinberg, Charmaine Coulen, Katherine A Poehling, Laura P Shone, Sharon Balter, et al. Population-based surveillance for hospitalizations associated with respiratory syncytial virus, influenza virus, and parainfluenza viruses among young children. *Pediatrics*, 113(6):1758–1764, 2004.

[48] John A Jacquez, Carl P Simon, James Koopman, Lisa Sattenspiel, and Timothy Perry. Modeling and analyzing hiv transmission: the effect of contact patterns. *Mathematical Biosciences*, 92(2):119–199, 1988.

[49] Denise J Jamieson, Margaret A Honein, Sonja A Rasmussen, Jennifer L Williams, David L Swerdlow, Matthew S Biggerstaff, Stephen Lindstrom, Janice K Louie, Cara M Christ, Susan R Bohm, et al. H1n1 2009 influenza virus infection during pregnancy in the usa. *The Lancet*, 374(9688):451–458, 2009.

[50] E Jung, Suzanne Lenhart, and Z Feng. Optimal control of treatments in a two-strain tuberculosis model. *Discrete and Continuous Dynamical Systems Series B*, 2(4):473–482, 2002.

[51] William O Kermack and Anderson G McKendrick. Contributions to the mathematical theory of epidemics. ii. the problem of endemicity. *Proceedings of the Royal society of London. Series A*, 138(834):55–83, 1932.

[52] Fan-kun Kong. Pilot clinical study on a proprietary elderberry extract: efficacy in addressing influenza symptoms. *Online Journal of Pharmacology and Pharmacokinetics*, 5:32–43, 2009.

[53] Yang Kuang. Special issue on mathematical models, challenges, and lessons learned from the 2009 a/h1n1 influenza pandemic. *Math Biosci Eng*, 8(1):1–238, 2011.

[54] Sunmi Lee, Gerardo Chowell, and Carlos Castillo-Chávez. Optimal control of influenza pandemics: the role of antiviral treatment and isolation. *Journal of Theoretical Biology*, 265:136–150, 2010.

[55] Sunmi Lee, Michael Golinski, and Gerardo Chowell. Modeling optimal age-specific vaccination strategies against pandemic influenza. *Bull Math Biol*, 74(4):958–80, Apr 2012.

[56] Sunmi Lee, Romarie Morales, and Carlos Castillo-Chavez. A note on the use of influenza vaccination strategies when supply is limited. *Math Biosci Eng*, 8(1):171–82, Jan 2011.

[57] Suzanne Lenhart and John T Workman. *Optimal control applied to biological models*. CRC Press, 2007.

[58] Ira M Longini, M Elizabeth Halloran, Azhar Nizam, and Yang Yang. Containing pandemic influenza with antiviral agents. *American journal of epidemiology*, 159(7):623–633, 2004.

[59] Michael Ludkovski and Jarad Niemi. Optimal dynamic policies for influenza management. *Statistical Communications in Infectious Diseases*, 2(1), 2010.

[60] Laura Matrajt, Jr and Ira M Longini. Optimizing vaccine allocation at different points in time during an epidemic. *PLoS One*, 5(11):e13767, 2010.

[61] Jan Medlock, Lauren Ancel Meyers, and Alison Galvani. Optimizing allocation for a delayed influenza vaccination campaign. *PLoS Curr*, 1:RRN1134, 2009.

[62] Daniel Merl, Leah R Johnson, Robert B Gramacy, and Marc Mangel. A statistical framework for the adaptive management of epidemiological interventions. *PLoS One*, 4(6):e5807, 2009.

[63] Christina E Mills, James M Robins, and Marc Lipsitch. Transmissibility of 1918 pandemic influenza. *Nature*, 432(7019):904–906, 2004.

[64] Arnold S Monto, A Webster, and Oliver Keene. Randomized, placebo-controlled studies of inhaled zanamivir in the treatment of influenza a and b: pooled efficacy analysis. *Journal of Antimicrobial Chemotherapy*, 44(suppl 2):23–29, 1999.

[65] Benjamin R Morin, Carlos Castillo-Chavez, Shu-Fang Hsu Schmitz, Anuj Mubayi, and Xiahong Wang. Notes from the heterogeneous: a few observations on the implications and necessity of affinity. *Journal of biological dynamics*, 4(5):456–477, 2010.

[66] Michael R Moser, Thomas R Bender, Harold S Margolis, Gary R Noble, Alan P Kendal, and Donald G Ritter. An outbreak of influenza aboard a commercial airliner. *American journal of epidemiology*, 110(1):1–6, 1979.

[67] Joël Mossong, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine*, 5(3):e74, 2008.

[68] Sandra Mounier-Jack and Richard J Coker. How prepared is europe for pandemic influenza? analysis of national plans. *The Lancet*, 367(9520):1405–1411, 2006.

[69] Sido D Mylius, Thomas J Hagenaars, Anna K Lugnér, and Jacco Wallinga. Optimal allocation of pandemic influenza vaccine depends on age, risk and timing. *Vaccine*, 26(29-30):3742–9, Jul 2008.

[70] KG Nicholson, FY Aoki, ADME Osterhaus, S Trottier, O Carewicz, CH Mercier, A Rode, N Kinnersley, P Ward, et al. Efficacy and safety of oseltamivir in treatment of acute influenza: a randomised controlled trial. *The Lancet*, 355(9218):1845–1850, 2000.

[71] M Nuno, C Castillo-Chavez, Z Feng, and M Martcheva. Mathematical models of influenza: the role of cross-immunity, quarantine and age-structure. In *Mathematical Epidemiology*, pages 349–364. Springer, 2008.

[72] M Nuno, Zhilan Feng, Maia Martcheva, and Carlos Castillo-Chavez. Dynamics of two-strain influenza with isolation and partial cross-immunity. *SIAM Journal on Applied Mathematics*, 65(3):964–982, 2005.

[73] Canadian Institute of Health. The impact of the h1n1 pandemic on canadian hospitals.

[74] World Health Organization. Update on a(h1n1) pandemic and seasonal vaccine availability, 2009.

[75] Paula Andreá Gonzalez Parra. Constrained optimal control for a multi-group discrete time influenza model. 2012.

[76] Patrick H Peters, Stefan Gravenstein, Paul Norwood, Veerle De Bock, Anthony Van Couter, Michael Gibbens, Tony-Andrea Von Planta, and Penelope Ward. Long-term use of oseltamivir for the prophylaxis of influenza in a vaccinated frail older population. *Journal of the American Geriatrics Society*, 49(8):1025–1031, 2001.

[77] Gregory A Poland. Vaccines against avian influenzaa race against time. *New England Journal of Medicine*, 354(13):1411–1413, 2006.

[78] Lev Semenovich Pontryagin. *Mathematical theory of optimal processes*. CRC Press, 1987.

[79] LS Pontryagin, VG Boltyanskii, RV Gamkrelidze, and E Mishchenko. The mathematical theory of optimal processes (international series of monographs in pure and applied mathematics). *Interscience, New York*, 1962.

[80] Anne M Presanis, Daniela De Angelis, Angela Hagy, Carrie Reed, Steven Riley, Ben S Cooper, Lyn Finelli, Paul Biedrzycki, Marc Lipsitch, et al. The severity of pandemic h1n1 influenza in the united states, from april to july 2009: a bayesian analysis. *PLoS medicine*, 6(12):e1000207, 2009.

[81] Associated Press. 40 million doses of h1n1 vaccine will be burned.

[82] World Health Organization. Global Influenza Programme. *Pandemic influenza preparedness and response: a WHO guidance document*. World Health Organization, 2009.

[83] Rino Rappuoli and Philip R Dormitzer. Influenza: options to improve pandemic preparation. *Science*, 336(6088):1531–1533, 2012.

[84] S-F Hsu Schmitz and C Castillo-Chavez. Parameter estimation in non-closed social networks related to dynamics of sexually transmitted diseases, 1994.

[85] Shu-Fang Hsu Schmitz and Carlos Castillo-Chavez. Completion of mixing matrices for non-closed social networks. In *World Congress of Nonlinear Analysts' 92: Proceedings of the First World Congress of Nonlinear Analysts, Tampa, Florida, August 19-26, 1992*, volume 1, page 3163. Walter de Gruyter, 1996.

[86] Eunha Shim. Optimal strategies of social distancing and vaccination against seasonal influenza. *Math Biosci Eng*, 10(5-6):1615–34, 2013.

[87] Jessica L Silvaggio, A Bearden, P Marquez, D Terashita, B Schwartz, and L Mascola. The influence of a mandate for influenza vaccination or masking of healthcare personnel: Experience from a large urban area. In *2015 CSTE Annual Conference*. Cste, 2015.

[88] Christina Spencer. Canada lends mexico 5m doses of h1n1 vaccine.

[89] Matthew W Tanner, Lisa Sattenspiel, and Lewis Ntaimo. Finding optimal vaccination strategies under parameter uncertainty using stochastic programming. *Mathematical biosciences*, 215(2):144–151, 2008.

[90] S Towers and Z Feng. Social contact patterns and control strategies for influenza in the elderly. *Mathematical biosciences*, 240(2):241–249, 2012.

[91] Ashleigh Tuite, David N Fisman, Jeffrey C Kwong, and Amy Greer. Optimal pandemic influenza vaccine allocation strategies for the canadian population. *PLoS Curr*, 2:RRN1144, 2010.

[92] Zoltan Vajo, Ferenc Tamas, Laszlo Sinka, and Istvan Jankovics. Safety and immunogenicity of a 2009 pandemic influenza a h1n1 vaccine when administered alone or simultaneously with the seasonal influenza vaccine for the 2009–10 influenza season: a multicentre, randomised controlled trial. *The Lancet*, 375(9708):49–55, 2010.

[93] P van den Driessche and James Watmough. Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission. *Math Biosci*, 180:29–48, 2002.

[94] Pauline Van den Driessche and James Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180(1):29–48, 2002.

[95] Josep Vaqué Rafart, Julita Gil Cuesta, and María Brotons Agulló. Principales características de la pandemia por el nuevo virus influenza a (h1n1). *Medicina clínica*, 133(13):513–521, 2009.

[96] Jacco Wallinga, Peter Teunis, and Mirjam Kretzschmar. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am J Epidemiol*, 164(10):936–44, Nov 2006.

[97] Laura Forsberg White, Jacco Wallinga, Lyn Finelli, Carrie Reed, Steven Riley, Marc Lipsitch, and Marcello Pagano. Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza a/h1n1 pandemic in the usa. *Influenza and Other Respiratory Viruses*, 3(6):267–276, 2009.

APPENDIX A

OPTIMAL CONTROL

A.1 Optimal Control Formulation

The goal is to find an optimal vaccination strategy that minimizes Objective Functional (3) given Criterion (4) and the regularity of System of equations (1) (the integrand of $\mathcal{F}$ is a convex function of u and the the state system satisfies the *Lipshitz* property with respect to the state variables). The existence of optimal controls is guaranteed by standard results in control theory [32]. The necessary conditions that optimal solutions must satisfy are derived using Pontryagin's Maximum Principle [79]. This principle converts Systems (1) and (4) into minimizing the Hamiltonian H given by

$$\begin{aligned} H &= I(t) + \frac{W}{2}u^2(t) \\ &+ A_1(t)\{-u(t)S(t) - \frac{\beta}{N(t)}(I(t)+J(t))S(t)\} \\ &+ A_2(t)\{\epsilon u(t)S(t) - \eta V(t) - \frac{\beta}{N(t)}(I(t)+J(t))V(t)\} \\ &+ A_3(t)\{(1-\epsilon)u(t)S(t) - \frac{\beta}{N(t)}(I(t)+J(t))F(t)\} \\ &+ A_4(t)\{\frac{\beta}{N(t)}(I(t)+J(t))(S(t)+V(t)+F(t)) - kE(t)\} \\ &+ A_5(t)\{kE(t) - (\alpha+\gamma_1)I(t))\} \\ &+ A_6(t)\{\alpha I(t) - (\gamma_2+\delta)J(t)\} \\ &+ A_7(t)\{u(t)S(t)\} \end{aligned} \tag{A.1}$$

From this Hamiltonian and Pontryagins Maximum Principle, we obtain
There exist the optimal control $u^*(t)$ and corresponding state solutions, X^*= (S^*, V^*, F^*, P^*, E^*, I^*, J^*, R^*, D^*) that minimize $\mathcal{F}(u(t))$ over Ω. In order for the above

statement to be true, it is necessary that there exist adjoint variables $A_i(t)$ such that

$$\begin{aligned}
\dot{A}_1 &= -[A_1(u(t) - A_1\frac{\beta}{N(t)}(I(t)+J(t)) + A_2(\epsilon u(t)) \qquad (A.2)\\
&+ A_3((1-\epsilon)u(t) + A_4\frac{\beta}{N(t)}(I(t)+J(t)) + A_7 u(t)]\\
\dot{A}_2 &= -[A_2 - \eta + A_2(-\frac{\beta}{N(t)}(I(t)+J(t))) + A_4\frac{\beta}{N(t)}(I(t)+J(t))]\\
\dot{A}_3 &= -[-A_3\frac{\beta}{N(t)}(I(t)+J(t)) + A_4\frac{\beta}{N(t)}(I(t)+J(t))]\\
\dot{A}_4 &= -[A_4(-k) + A_5 k]\\
\dot{A}_5 &= -[1 - A_1\frac{\beta}{N(t)}S(t) - A_2\frac{\beta}{N}V(t) - A_3\frac{\beta}{N(t)}F(t)\\
&+ A_4\frac{\beta}{N(t)}(S(t)+V(t)+F(t)) - A_5(\alpha+\gamma_1) + A_6\alpha]\\
\dot{A}_6 &= -[-A_1\frac{\beta}{N(t)}S(t) - A_2\frac{\beta}{N(t)}V(t) + A_3\frac{\beta}{N(t)}(F(t))\\
&- A_4\frac{\beta}{N(t)}(S(t)+V(t)+F(t)) - A_6(\gamma_2+\delta)]\\
\dot{A}_7 &= 0
\end{aligned}$$

satisfying the transversality conditions,

$$\begin{aligned}
A_i(T) &= 0,\ i = 1, \cdots, 6 \qquad (A.3)\\
A_7(T) &= \theta. \qquad (A.4)
\end{aligned}$$

The Hamiltonian H is minimized with respect to the control (at the optimal control). We differentiate H with respect to u on the set Ω and arrive at the following optimality condition:

$$\frac{\partial H}{\partial u} = Wu(t) - A_1(t)S(t) + \epsilon S(t)A_2(t) + (1-\epsilon)A_3(t)S(t) + A_7(t)S(t). \qquad (A.5)$$

Solving for u^* (by evaluating $\frac{\partial H}{\partial u}$ at u^*), the optimality condition

$$u(t) = \frac{S(t)}{W}(A_1(t) - \epsilon A_2(t) - (1-\epsilon)A_3(t) + A_7(t)) \qquad (A.6)$$

is obtained. Furthermore, using the standard argument for control bounds, we arrive at the following expression for the optimal control function

$$u^*(t) = min\big\{max\big\{0, \frac{S(t)}{W}(A_1(t) - \epsilon A_2(t) - (1-\epsilon)A_3(t) + A_7(t))\big\}, b\big\}. \qquad (A.7)$$

Table A.1: Parameter Definitions and Baseline Values (and their Corresponding Sources) Used in Numerical Simulations.

Parameter	Description	Values	Reference
β	Transmission rate (days^{-1})	$0.75 - 1.68$	[21]
k	Rate of progression to infectious (days^{-1})	0.53	[63]
δ	Mortality rate (days^{-1})	0.01	[37]
γ_1	Recovery rate (days^{-1}) for infectious class (days^{-1})	0.34	[21]
γ_2	Recovery rate for hospitalized class (days^{-1})	1.10	[21]
α	Diagnostic rate (days^{-1})	0.51	[21]
ϵ	Efficacy of vaccination	0.5	[58]
$S(0)$	Initial number of susceptible individuals	174673	[21]
$E(0)$	Initial number of exposed individuals	207	[21]
$I(0)$	Initial number of infectious individuals	132	[21]
T	The simulated time (days)	200	-
b	The upper bound of control	$.05 - 0.2$	-
W	Weight constant on control	100	-

The unconstrained solution can be computed by solving the optimality system which excludes the $\dot{Y}(t)$ equation in (1) and the $A_7(t)$ equation in (7). The standard two point boundary method is used to solve the unconstrained problem: first, the state system is solved using a forward method with given initial conditions; secondly, the corresponding adjoint system is solved using a backward scheme with the transversality conditions; thirdly, a convex combination of previously and currently computed controls are used to generate updated controls using the optimality equations; lastly, the process is repeated until a convergence criterion is satisfied.

For, the constrained optimization problem, a new state variable, $Y(t)$ is introduced in (1) from the isoperimetric constraint (5), which requires boundary conditions at $t = 0$ and $t = T$. From the requirements on $Y(t)$, the corresponding adjoint variable to $Y(t)$ must meet a non-zero transversality condition at the final time T, namely that $A_7(T) \equiv \theta$. Note that θ is unknown therefore, an iteration process is needed to find the right transversality condition required to satisfy the isoperimetric constraint ($Y(T) = B$). This additional iteration process uses Newton's method and the procedure used to implement it numerically identifies convergence issues that were never generated in the search of solutions for the unconstrained problem.

The figure A.1 shows the optimal control functions over time when $R_0 = 1.3$ and $R_0 = 2.0$. In this case the corresponding daily incidence rate in the infected class is compared to the situation in the absence of vaccines.

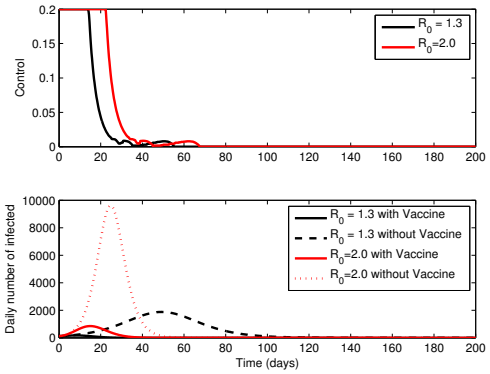

Figure A.1: The Optimal Control Functions over Time are Plotted When $R_0 = 1.3$ and $R_0 = 2.0$.

Figure A.2 shows the comparison of results identifying optimal vaccination policies using different weight constants on the control ($W = 1, 10^2, 10^4, 10^6$) when $R_0 = 1.3$. The optimal controls and the corresponding incidence of infected individuals are illustrated (left). The fraction of the cumulative number of infected and vaccinated individuals are plotted under different weight constants as a function of R_0 (right).

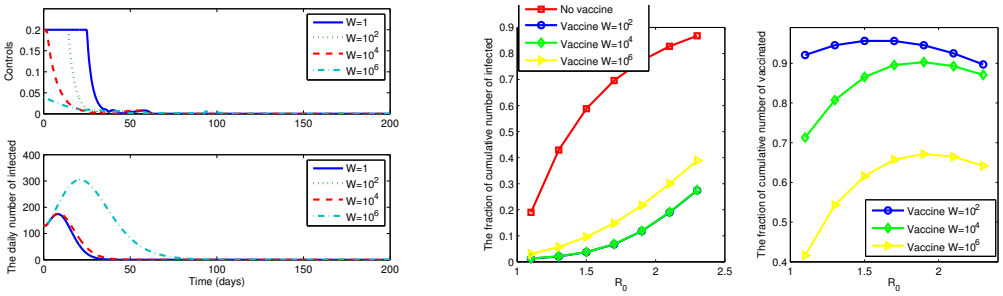

Figure A.2: Optimal Vaccination Policies Results using Different Weight Constants on the Control ($W = 1, 10^2, 10^4, 10^6$) when $R_0 = 1.3$.

Figure A.3 shows the results of implementing optimal vaccination strategies when different upper control bounds ($b = 0.05, 0.1, 0.2, 0.5$) are used with $R_0 = 1.3$. The optimal controls and the corresponding incidence of infected individuals are illustrated (left). The fraction of the cumulative number of infected and vaccinated individuals are plotted for distinct upper bounds as a function of R_0. (right).

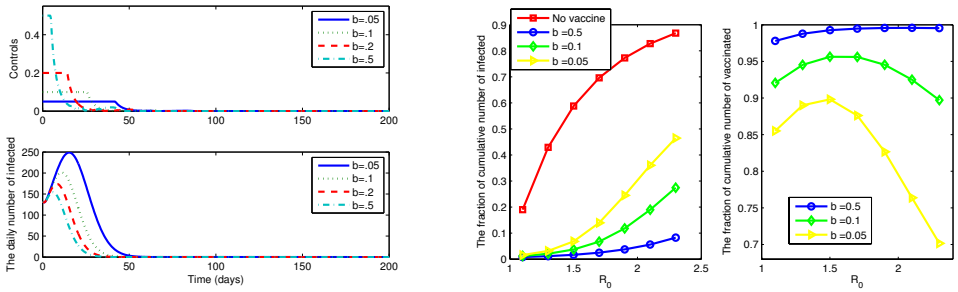

Figure A.3: Optimal Vaccination Strategies Results when Different Upper Control Bounds ($b = 0.05, 0.1, 0.2, 0.5$) are used with $R_0 = 1.3$.

Figure A.4 shows the results of implementing optimal vaccination policies under distinct vaccine efficacy constants on the control ($\epsilon = 0.4, 0.6, 0.8, 1$) when $R_0 = 1.3$. The optimal controls and the corresponding incidence of infected individuals are plotted (left). The fraction of the cumulative number of infected and vaccinated individuals are shown for different efficacy levels as a function of R_0 (right).

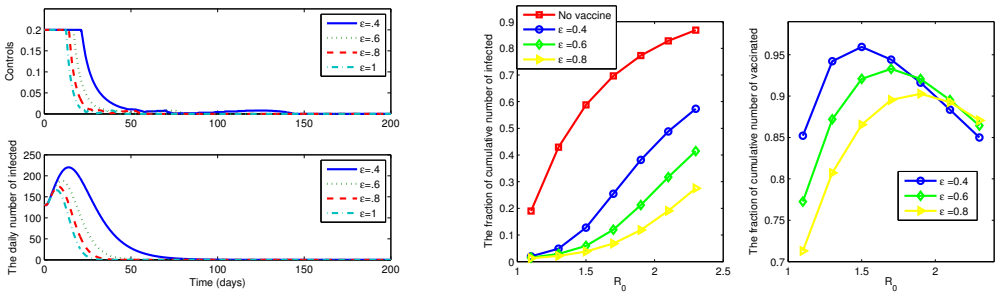

Figure A.4: Optimal Vaccination Policies under Distinct Vaccine Efficacy Constants on the Control ($\epsilon = 0.4, 0.6, 0.8, 1$) when $R_0 = 1.3$.

Figure A.5 shows the optimal control functions plotted when b=0.05 (left) and b=0.2 (right) under three different vaccination coverage (15%, 30%, and 50%). The corresponding daily incidence in the infected class is compared with the one in the absence of vaccines.

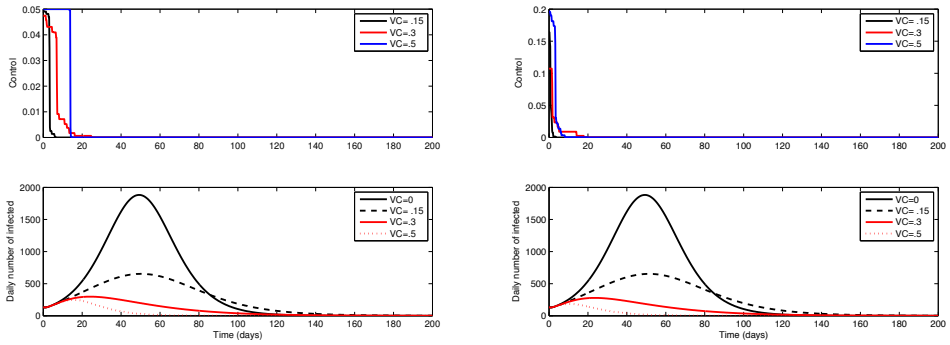

Figure A.5: Optimal Control Functions Plotted When b=0.05 (left) and b=0.2 (right) under Three Different Vaccination Coverage (15%, 30%, and 50%).

Figure A.6 shoes the raction of the cumulative number of infected individuals as a function of R_0 is compared with the one under no vaccines using two different vaccination coverage levels (30% and 50%).

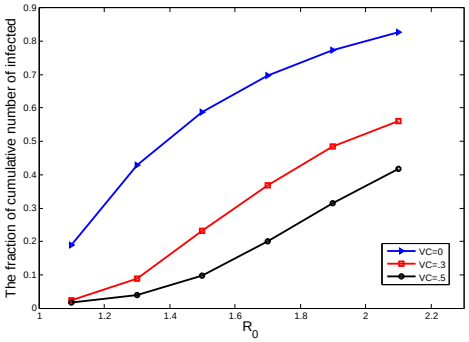

Figure A.6: Fraction of Cumulative Number of Infected Individuals as a Function of R_0 Compared with and without Vaccines.

APPENDIX B

MATHEMATICAL COMPUTATIONS

The contact matrix used in the age-structured dynamics model (Equation 3.1) is commuted from the survey data as follows:

1. For each individual the number of daily contacts were computed.
2. An average number of contacts for a specific age is calculated using data from step 1 and age of individuals
3. Age-groups are constructed and an average contacts per age-group using step 2 are computed

Figure B.1 shows the diagram to formulate the contact matrix. To the left is the participant id and the average age of the individuals he contacted in a given day, the column to the right is the participant key id and age.

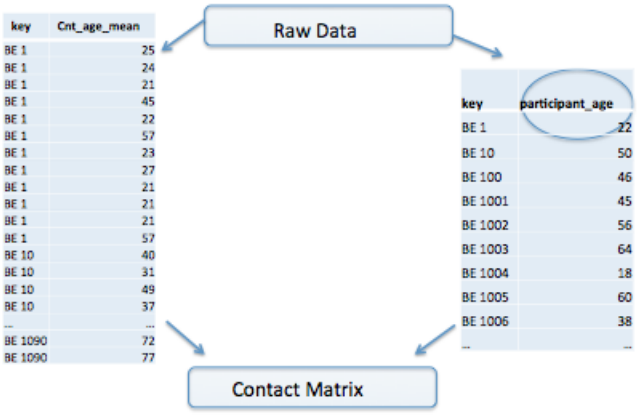

Figure B.1: Diagram to Formulate Contact Matrix .

B.1 R_0 Computation

Using the Next generation Operator we obtain:

$$\mathcal{F} = [qS_i \sum_{j=1}^{m} p_{ij} \frac{I_j}{N_j}] \tag{B.1}$$

$$F = \begin{bmatrix} qS_1\frac{p_1}{N_1} & qS_1\frac{p_2}{N_2} & qS_1\frac{p_3}{N_3} & \dots & qS_1\frac{p_m}{N_m} \\ qS_2\frac{p_1}{N_1} & qS_2\frac{p_2}{N_2} & qS_2\frac{p_3}{N_3} & \dots & qS_2\frac{p_m}{N_m} \\ \dots & \dots & \dots & \dots & \dots \\ qS_m\frac{p_1}{N_1} & qS_m\frac{p_2}{N_2} & qS_m\frac{p_3}{N_3} & \dots & qS_m\frac{p_m}{N_m} \end{bmatrix}$$

$$\mathcal{V} = \begin{bmatrix} \gamma_1 I_1 \\ \gamma_2 I_2 \\ \dots \\ \gamma_m I_m \end{bmatrix}$$

$$V = \begin{bmatrix} \gamma_1 & 0 & 0 & \dots & 0 \\ 0 & \gamma_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \gamma_m \end{bmatrix}$$

$$FV^{-1} = \begin{bmatrix} \frac{qS_1\frac{p_1}{N_1}}{\gamma_1} & \frac{qS_1\frac{f_2}{N_2}}{\gamma_2} & \frac{qS_1\frac{f_3}{N_3}}{\gamma_3} & \dots & \frac{qS_1\frac{p_m}{N_m}}{\gamma_m} \\ \frac{qS_2\frac{p_1}{N_1}}{\gamma_1} & \frac{qS_2\frac{p_2}{N_2}}{\gamma_2} & \frac{qS_2\frac{p_3}{N_3}}{\gamma_3} & \dots & \frac{qS_2\frac{p_m}{N_m}}{\gamma_m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{qS_m\frac{p_1}{N_1}}{\gamma_1} & \frac{qS_m\frac{p_2}{N_2}}{\gamma_2} & \frac{qS_m\frac{p_3}{N_3}}{\gamma_3} & \dots & \frac{qS_m\frac{p_m}{N_m}}{\gamma_m} \end{bmatrix}$$

Taking the maximum eigenvalue of FV^{-1} we obtain:

For 2 Age groups:

$$R_0 = \frac{q(p_2\gamma_1 + p_1\gamma_2)}{\gamma_1\gamma_2} \tag{B.2}$$

For 3 Age groups:

$$R_0 = \frac{q(p_2\gamma_1\gamma_3 + p_1\gamma_2\gamma_3 + p_3\gamma_1\gamma_2)}{\gamma_1\gamma_2\gamma_3} \tag{B.3}$$

For 4 Age groups:

$$R_0 = q\frac{\sum_{i\neq j\neq k\neq l}^{4} p_i\gamma_j\gamma_k\gamma_l}{\prod_{n=1}^{4}\gamma_n} \tag{B.4}$$

B.2 Final Epidemic Size

The expression for final size of an epidemic was computed using equations of the system and the steps are as follows:

$$\begin{aligned} \dot{S_i}(t) &= -qS_i(t)\sum_{j=1}^{m} p_{ij}\frac{I_j(t)}{N_j} \\ \int_0^\infty \dot{S_i}(t)dt &= -\int_0^\infty qS_i(t)\sum_{j=1}^{m} p_{ij}\frac{I_j(t)}{N_j}dt \\ \int_0^\infty \frac{\dot{S}_i(t)}{S_i(t)}dt &= -\int_0^\infty q\sum_{j=1}^{m} p_{ij}\frac{I_j(t)}{N_j}dt \end{aligned} \tag{B.5}$$

In this case we observe that $R_i(\infty) - R_i(0) = Z_i = S_i(0) - S_i(\infty)$ so we obtain

$$\begin{aligned} \ln \frac{S_i(0)}{S_i(\infty)} &= -q \int_0^\infty \sum_{j=1}^{m} p_{ij} \frac{I_j(t)}{N_j} dt \\ \ln(-Z_i) &= -q \int_0^\infty \sum_{j=1}^{m} p_{ij} \frac{I_j(t)}{N_j} dt \end{aligned} \tag{B.6}$$

Given that we have the integral of I we proceed to solve for $\dot{R_i}(t)$ equation and we obtain:

$$\begin{aligned} \int_0^\infty \dot{R_i}(t) &= \int_0^\infty \gamma I_i(t) dt \\ R_i(\infty) - R_i(0) &= \int_0^\infty \gamma I_i(t) dt \\ \frac{Z_i}{\gamma} &= \int_0^\infty I_i(t) dt \end{aligned} \tag{B.7}$$

Putting all back to the Susceptible equation we obtain:

$$\begin{aligned} \ln(-Z_i) &= -q \int_0^\infty \sum_{j=1}^{m} \frac{p_{ij}}{N_j} \frac{Z_j}{\gamma} dt \\ \ln \frac{S_i(0)}{S_i(\infty)} &= -q \sum_{j=1}^{m} \frac{p_{ij}}{N_j} \frac{S_j(0) - S_j(\infty)}{\gamma} \end{aligned} \tag{B.8}$$

B.3 Algorithm to group Participants by their Age and Contacts

1. Construct age groups (say,. 0-25 26-50 51-100; 3 age groups)
2. Use the survey data provided and match the age of the participant with the group division it belongs to (ie. an individual of age 22 can belong on the age group from 0-25)
3. Use survey data and match the age of the individuals with whom the participant had contact with
4. Store the information in a matrix that shows the number of contacts of individuals in one age-group with individuals that belong to a different age-group.

B.4 Algorithm to Construct Mixing Matrix

The pseudocode for the extraction of the data from the survey data: Steps:

1. Initialize the average number of contact parameter for each of the age groups
2. Initialize the proportion of the population for each of the age groups.

3. Write in the population proportion or weight for each age from the data
4. Initialize an age vector from 0-99
5. Initialize an age vector from 1-100
6. Construct a 0 matrix
7. Populate the matrix with the data from the survey times the proportion (the weight) of the population of each age

$$Matrix_{recip} = \begin{bmatrix} C_{11}w_1 & C_{12}w_1 & C_{13}w_1 & \dots & C_{1m}w_1 \\ C_{21}w_2 & C_{22}w_2 & C_{23}w_2 & \dots & C_{2m}w_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{m1}w_m & C_{m2}w_m & C_{m3}w_m & \dots & C_{mm}w_m \end{bmatrix}$$

8. Initialize the end points for each of the age groups
9. Sum all the weights of the population that is contained in the same age group (ie. for a 2 age group example $\sum_{j=0}^{20} w_j$ for the age group [0-20] , and $\sum_{j=10}^{100} w_j$ for age group [21-100])
10. Make sure that reciprocity is taken into accounts (ie. contacts that kids have with adults should be the same as the adults have with kids)
    (a) Using the information stored in the $Matrix_{recip}$ which contains all the 100 by 100 bins
        i. Do a for loop i where you go over the starting point of age group i and end at the starting point of the next age group(i.e. from vector 0-99 we want the group from 0-20 and 21-99 and from vector 1-100 we will divide 1-21 and 22-100)
        ii. Do a for loop j where you go over the starting point of age group j and end at the starting point of the next age group
        iii. Form a smaller matrix with each of the corresponding groups(From our two age group example we obtain)

$$Matrix_{recip} = \begin{bmatrix} \begin{bmatrix} C_{1,1} & C_{1,2} & \dots & C_{1,20} \\ \vdots & \vdots & \ddots & \vdots \\ C_{20,1} & C_{20,2} & \dots & C_{20,20} \end{bmatrix} & \begin{bmatrix} C_{1,21} & C_{2,22} & \dots & C_{1,100} \\ \vdots & \vdots & \ddots & \vdots \\ C_{20,21} & C_{20,22} & \dots & C_{20,100} \end{bmatrix} \\ \begin{bmatrix} C_{21,1} & C_{21,2} & \dots & C_{21,20} \\ \vdots & \vdots & \ddots & \vdots \\ C_{100,1} & C_{100,2} & \dots & C_{100,20} \end{bmatrix} & \begin{bmatrix} C_{21,21} & C_{21,22} & \dots & C_{21,100} \\ \vdots & \vdots & \ddots & \vdots \\ C_{100,21} & C_{100,22} & \dots & C_{100,100} \end{bmatrix} \end{bmatrix}$$

        iv. Sum all the values in each of the smaller matrices and store $\begin{bmatrix} \left[\sum_{i=1}^{20}\sum_{j=1}^{20} C_{i,j}\right] & \left[\sum_{i=1}^{20}\sum_{j=21}^{100} C_{i,j}\right] \\ \left[\sum_{i=21}^{100}\sum_{j=1}^{20} C_{i,j}\right] & \left[\sum_{i=21}^{100}\sum_{j=21}^{100} C_{1,2}C_{i,j}\right] \end{bmatrix}$
        v. Repeat until there are no more age groups

(b) Divide each element of the matrix by its corresponding age weight: this will be New_{Matrix}

$$\text{New}_{Matrix} = \begin{bmatrix} \left[\frac{\sum_{i=1}^{20}\sum_{j=1}^{20} C_{i,j}}{\sum_{i=1}^{20} w_i}\right] & \left[\frac{\sum_{i=1}^{20}\sum_{j=21}^{100} C_{i,j}}{\sum_{i=1}^{20} w_i}\right] \\ \left[\frac{\sum_{i=21}^{100}\sum_{j=1}^{20} C_{i,j}}{\sum_{i=21}^{100} w_i}\right] & \left[\frac{\sum_{i=21}^{100}\sum_{j=21}^{100} C_{1,2}C_{i,j}}{\sum_{i=21}^{100} w_i}\right] \end{bmatrix}$$

(c) Divide the proportion of the population of age group i by the weight of age group i and label as fraction_{Pop}

(d) multiply the New_{Matrix} by the fraction_{Pop}

11. Add all the new information on the matrix output

APPENDIX C

CONTROL FRAMEWORK

C.1 CDC Age Group Framework

The optimal control obtained from the model analysis are shown in the figures below. the goal is to show the difference between models results when mixing matrices was directly estimated from data (Mossong et al. 2008). The notation TN stands for mean value of average number of contacts rates per age-groups where its distribution is assumed to be Normal.

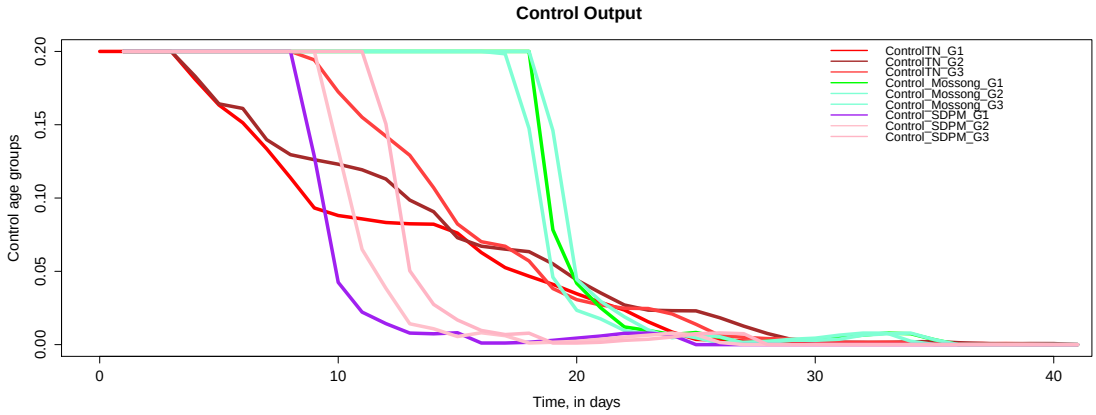

Figure C.1: Control Output to Minimize Infection with 3 Age Groups and with each Mixing Assumption.

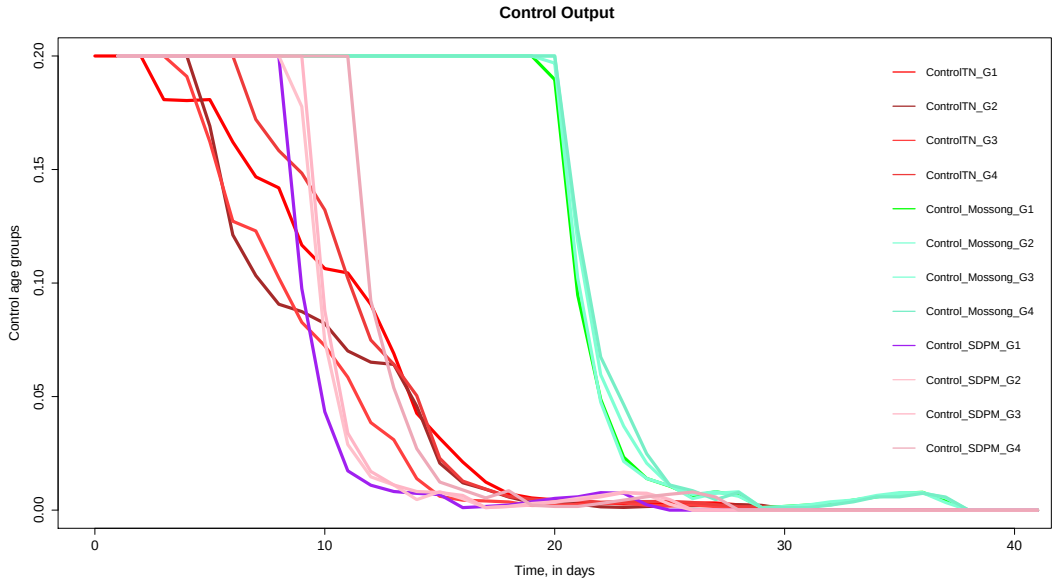

Figure C.2: Control Output to Minimize Infection with 4 Age Groups and with each Mixing Assumption.

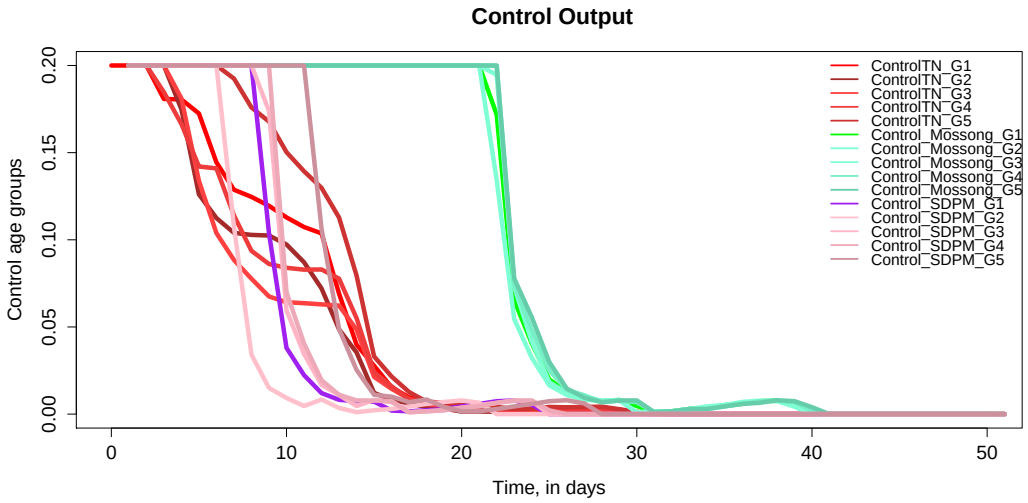

Figure C.3: Control Output to Minimize Infection with 5 Age Groups and with each Mixing Assumption.

The graph below shows the incidence curve using each one of the approaches. For the case of the TNPM we took the average of all the replicates. We observe that the incidence over time is very similar when we have two age groups divisions. However as we increase the number of age groups we notice the incidence output similarities between the Mossong grouping and the TNPM approach. We can infer that the difference might decrease as we increase the number of age groups given the reduction in the density of contacts.

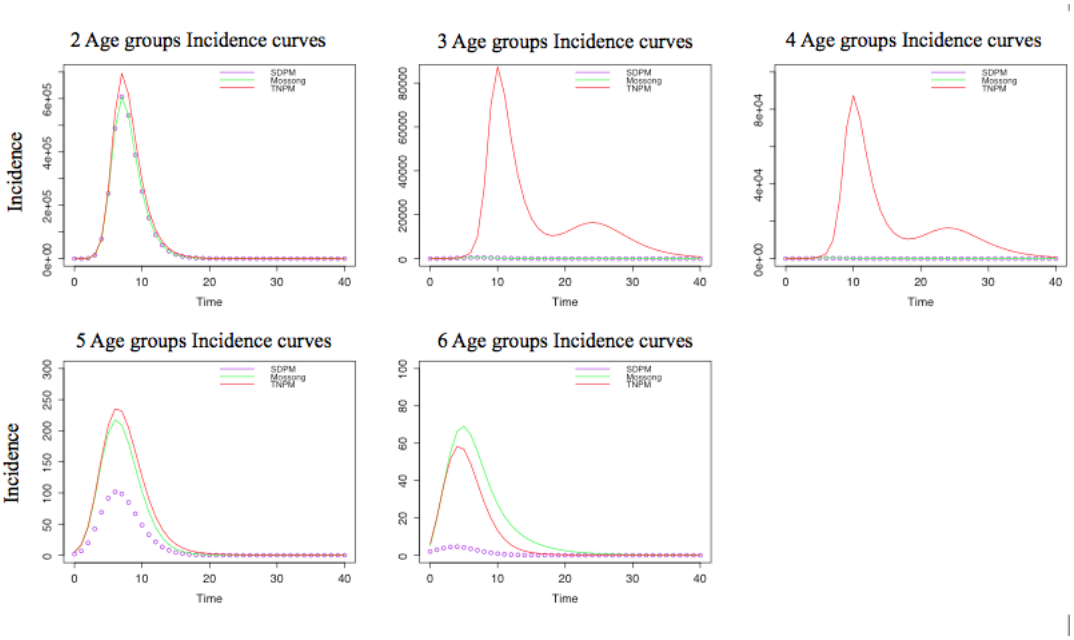

Figure C.4: Output of The Incidence Over Time of Infected Individuals from Age Group 1 to 6 using the 3 Different Approaches.

The graph below shows the incidence over time for each particular approach. We plotted all the age groups simulations. Given that age group 2 was very high, we plotted the rest of the incidence by age excluding age group 2. With the Mossong approach and the SDPM approach we observed that we have the same qualitative behavior among all age groups.

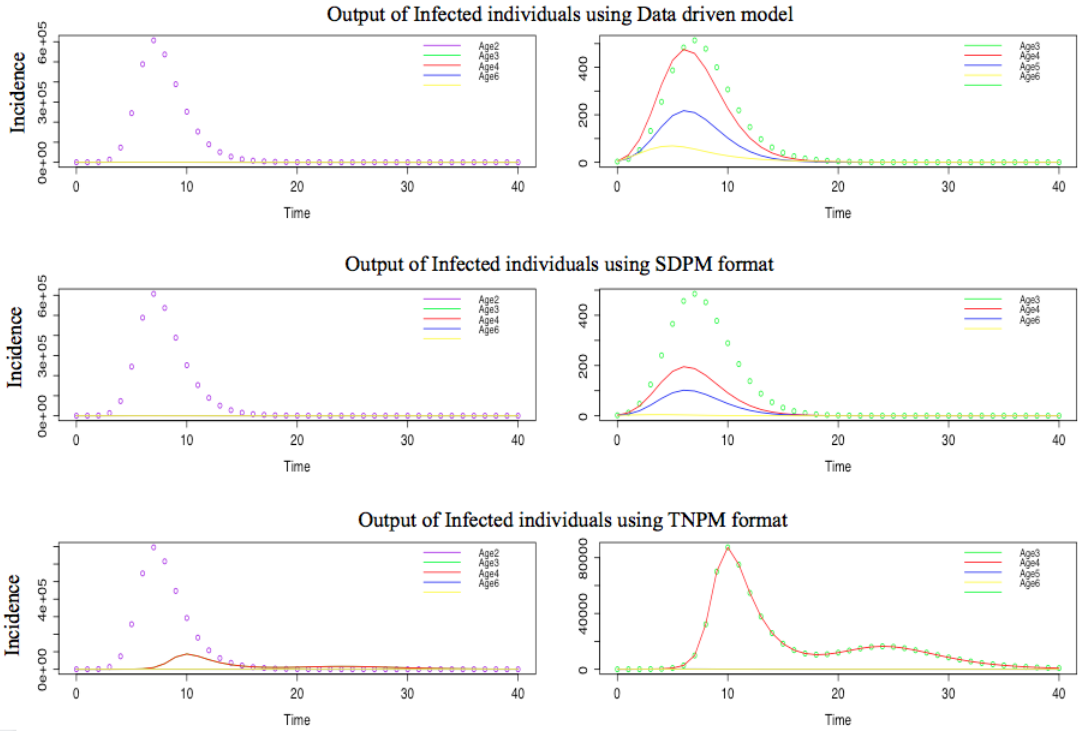

Figure C.5: Output of Infected Individuals Separated by Case.

The graphs below show the replicates of the incidence plots over time using the TNPM approach when two age-groups are considered in the model. We present graphs as we increase the age groups from 2 to 6.

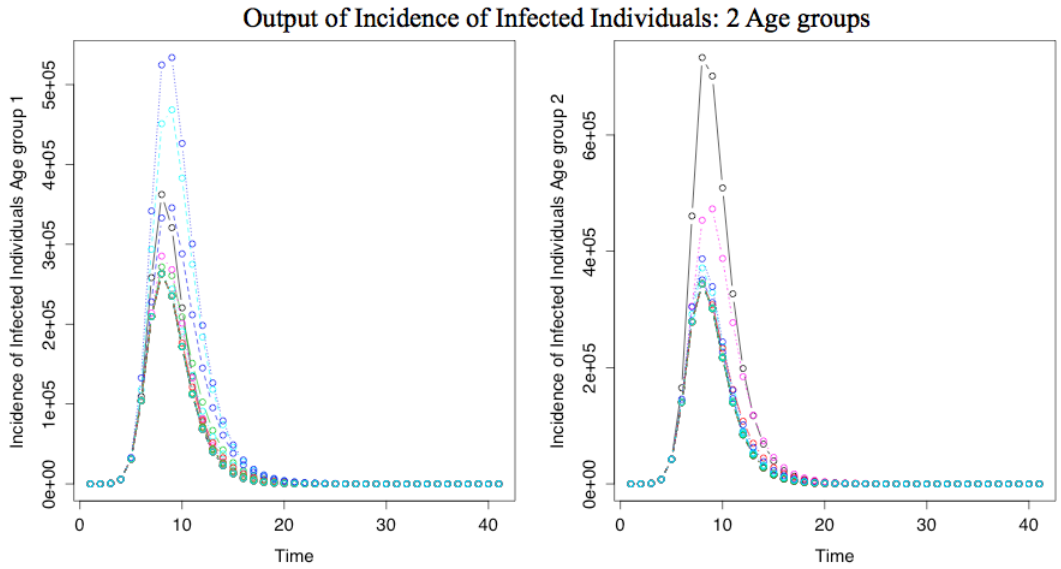

Figure C.6: Output of Infected Replicates for Age Group 2 .

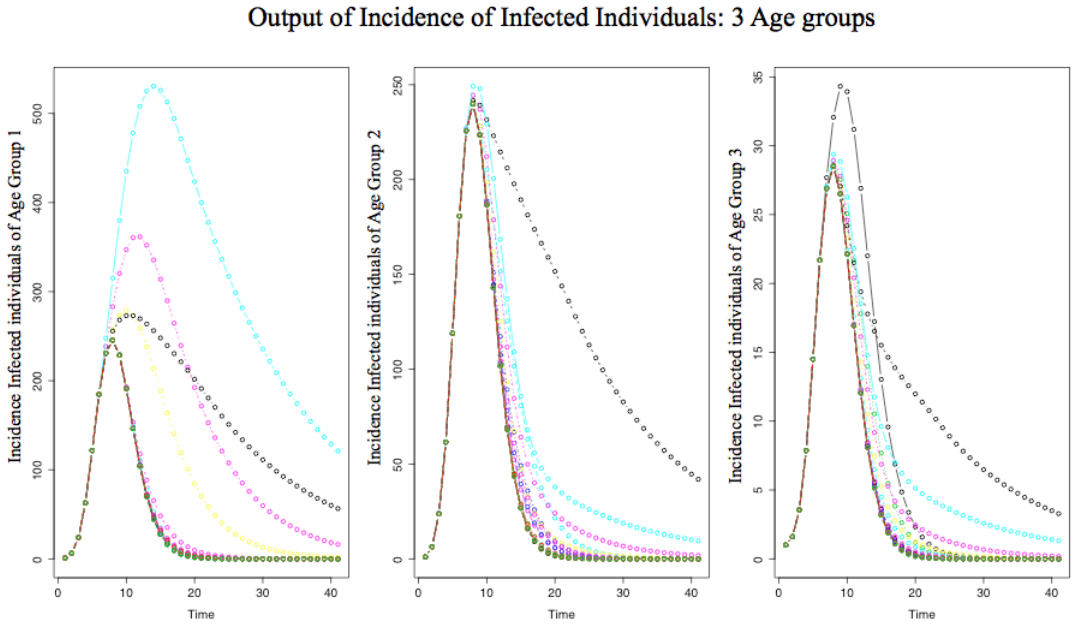

Figure C.7: Output of Infected Replicates for Age Group 3 .

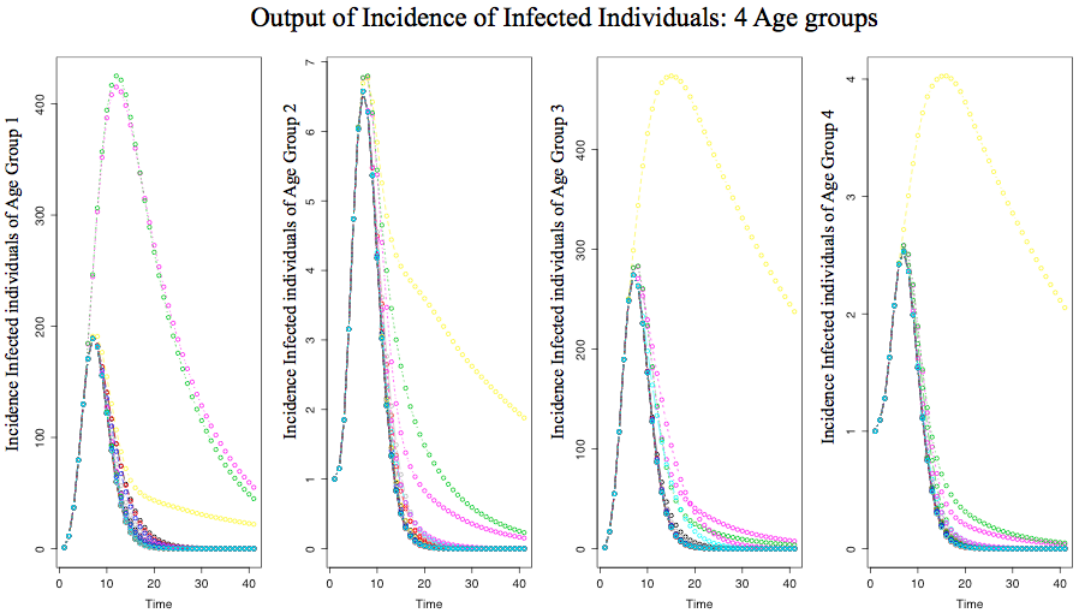

Figure C.8: Output of Infected Replicates for Age Group 4 .

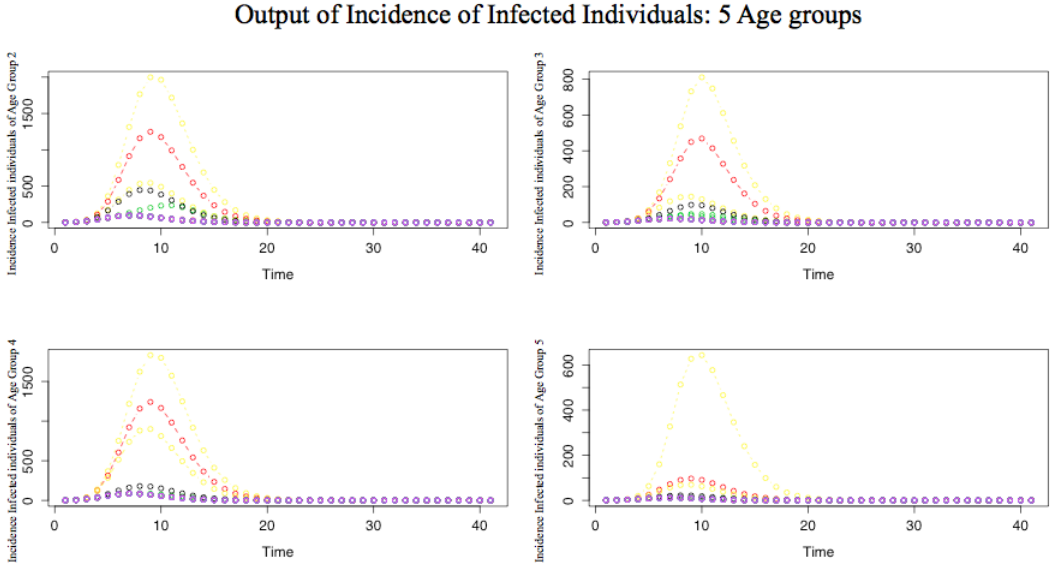

Figure C.9: Output of Infected Replicates for Age Group 5 .

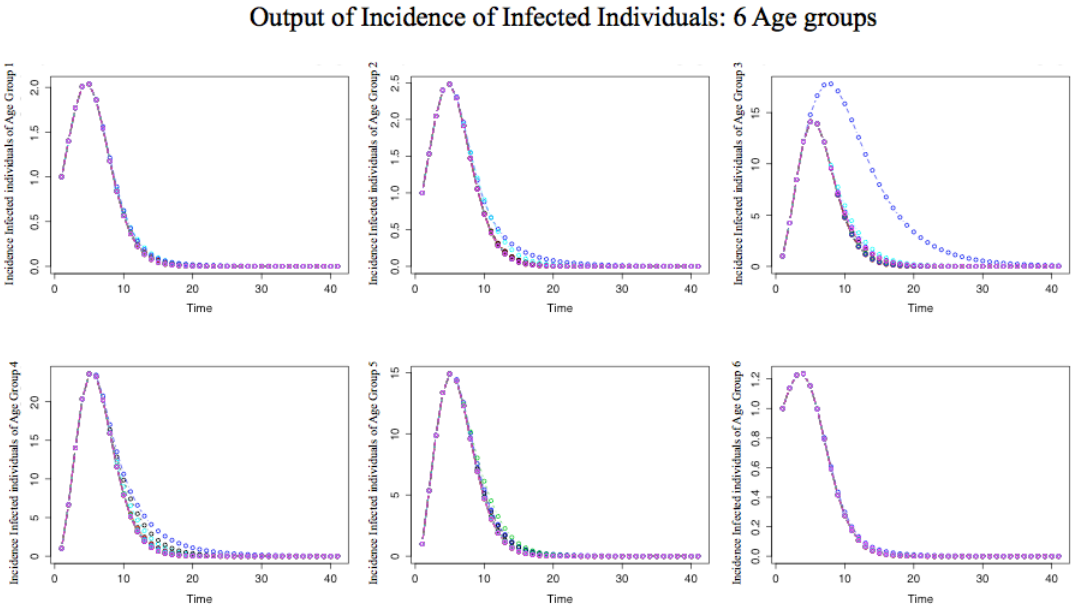

Figure C.10: Output of Infected Replicates for Age Group 6 .

C.2 Lee et al. Age Group Framework

The graph below we show the relative cost difference when we follow the age group arrangement for Lee et al. The relative cost difference in both cases increases as we increase the number of age groups. The trend can be appreciated from the graph to the right. The difference using the SDPM vs Mossong regroup data is smaller in the first 4 age groups divisions.

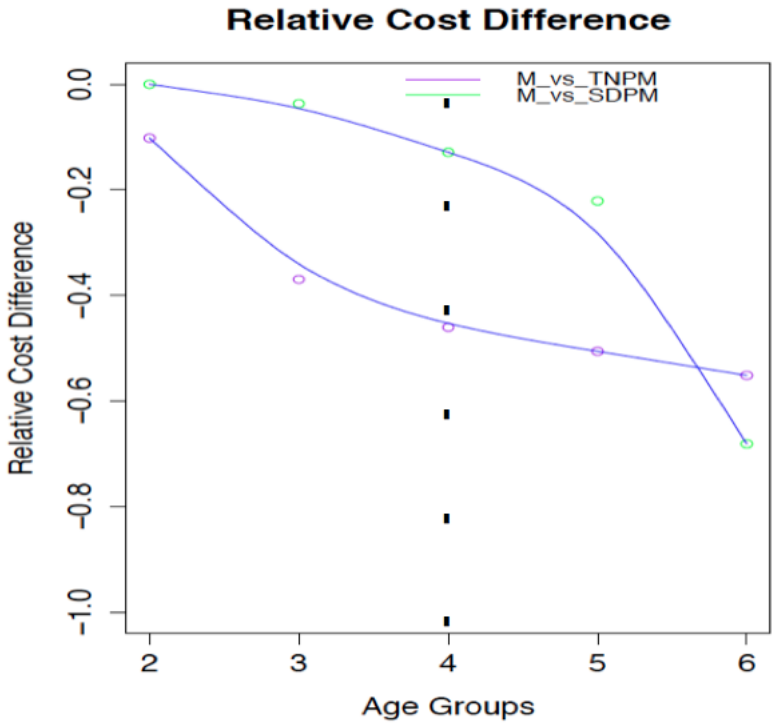

Figure C.11: Relative Cost Difference Between Models using Contact Matrix from Mossong et al. and Contact Matrix that uses Proportionate Mixing Formulations.

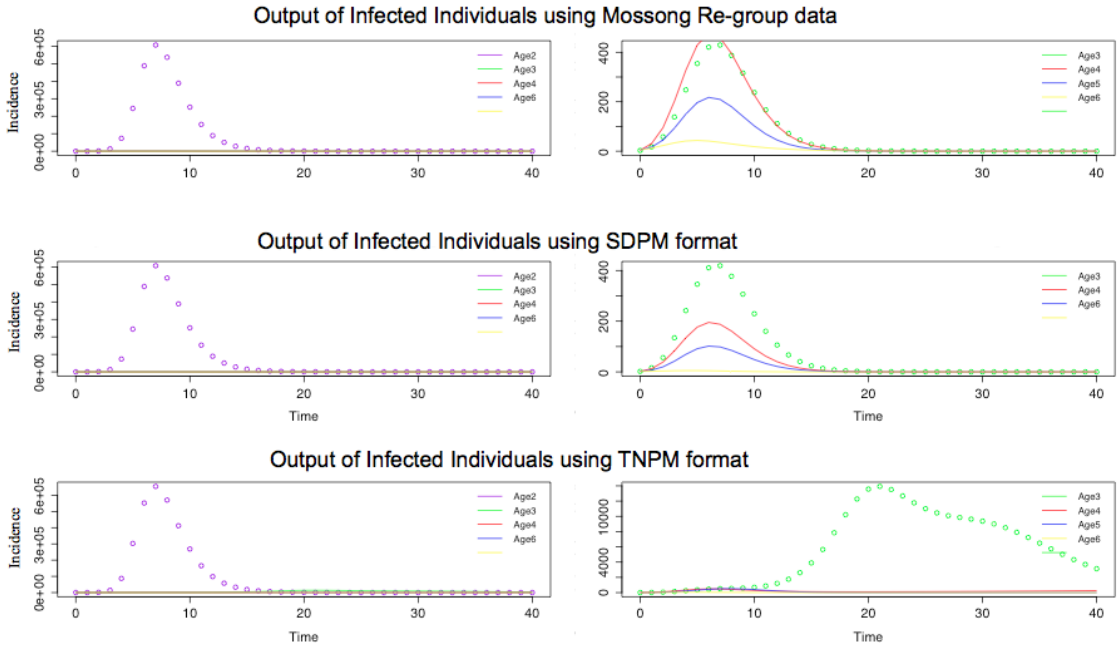

Figure C.12: Output of Infected Individuals Separated by Case.

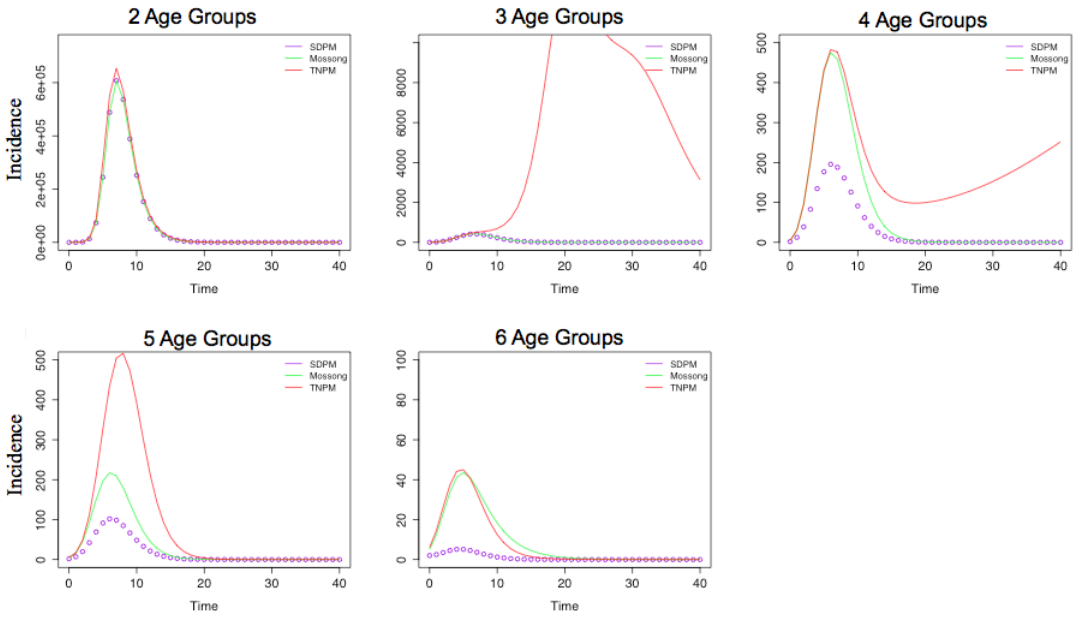

Figure C.13: Output of Infected Individuals from Age Group 1 to 7 using the 3 Different Approaches.

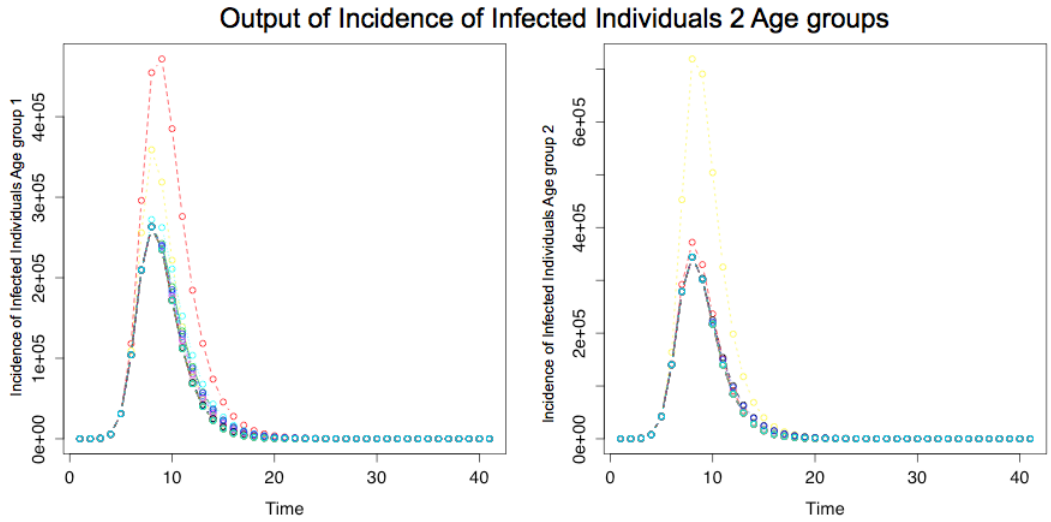

Figure C.14: Output of Infected Replicates for Age Group 2 .

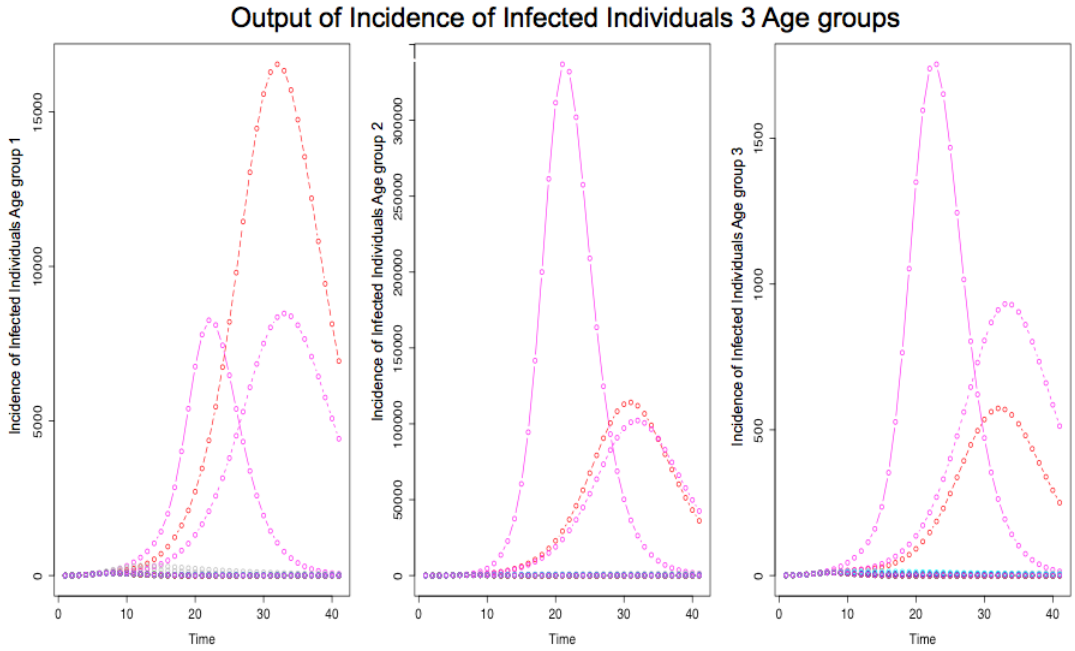

Figure C.15: Output of Infected Replicates for Age Group 3 .

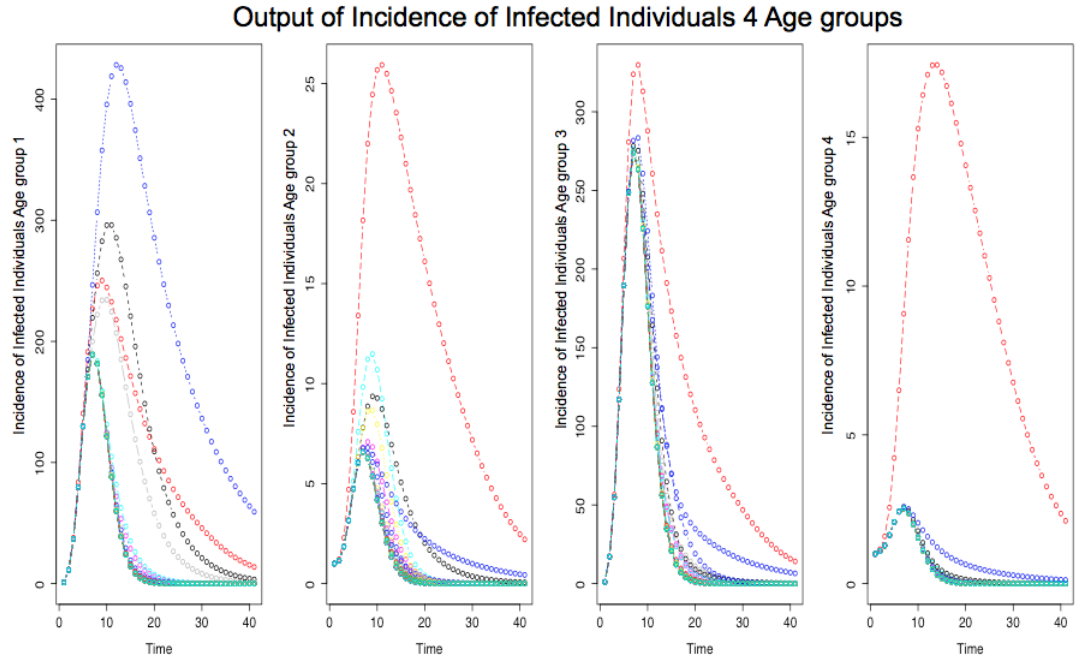

Figure C.16: Output of Infected Replicates for Age Group 4 .

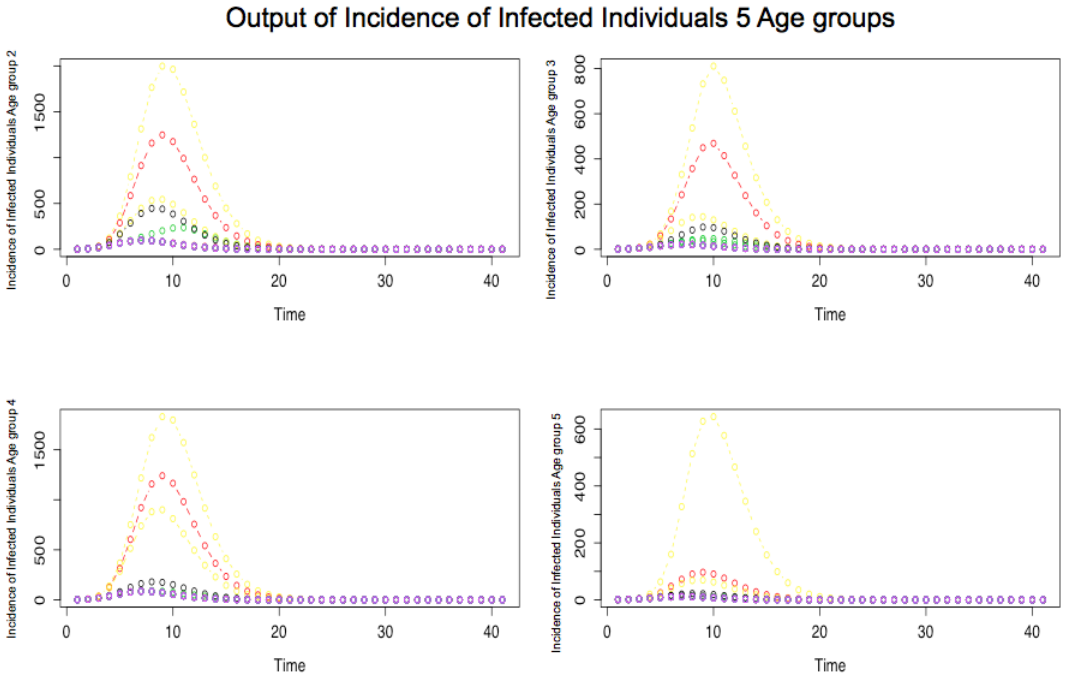

Figure C.17: Output of Infected Replicates for Age Group 5 .

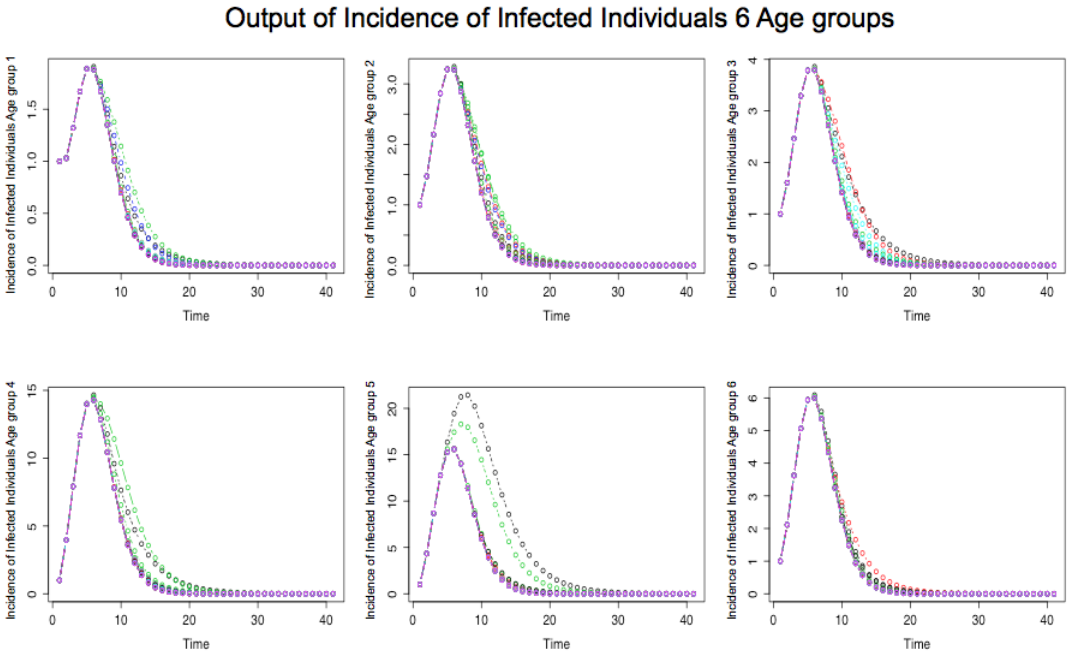

Figure C.18: Output of Infected Replicates for Age Group 6 .

C.3 Isoperimetric Constraint

In this case we will evaluate the case when we have limited resources for example, the Objective Function is:

$$\min_{u_i} \int_0^T \sum_{i=1}^{m} (I_i + u_i^2 \frac{W}{2}) dt \tag{C.1}$$

Subject to the state equations

$$\begin{aligned} \dot{S}_i &= -u_i S_i - q S_i \sum_{j=1}^{m} f_{ij} \frac{I_j}{N_j} \\ \dot{I}_i &= q S_i \sum_{j=1}^{m} f_{ij} \frac{I_j}{N_j} - \gamma_i I_i - \sigma_i I_i \\ \dot{R}_i &= u_i S_i + \gamma_i I_i \end{aligned} \tag{C.2}$$

and the isoperimetric constraint (limited vaccines)

$$\int_0^T \sum_{i=1}^{m} (S_i u_i^2 \frac{W}{2}) dt = B \tag{C.3}$$

In this case B is the maximum amount of vaccines available.

In order to be able to analyze the optimal control problem with the constraint we cannot use the Pontryagins Max principle. So we introduce an additional state variable z, where $\dot{z}$ is

$$\begin{aligned} \dot{z} &= \sum_{i=1}^{m} u_i S_i \\ z(t_0) &= 0 \\ z(t_F) &= B \end{aligned} \tag{C.4}$$

Using the Pontryagin's Maximum Principle [79]. This principle converts ODE

Systems and Constraints into minimizing the Hamiltonian H given by

$$\begin{aligned} H &= \sum_{i=1}^{m}[I_i(t) + \frac{W_i}{2}u_i^2(t)] \\ &+ \sum_{i=1}^{m}\lambda_{S_i}[-u_iS_i - qS_i\sum_{j=1}^{m}f_{ij}\frac{I_j}{N_j}] \\ &+ \sum_{i=1}^{m}\lambda_{I_i}[qS_i\sum_{j=1}^{m}f_{ij}\frac{I_j}{N_j} - \gamma_iI_i - \sigma_iI_i] \\ &+ \sum_{i=1}^{m}\lambda_{R_i}[u_iS_i + \gamma_iI_i] \\ &+ \sum_{i=1}^{m}\lambda_z u_iS_i \end{aligned} \tag{C.5}$$

where m is the number of age groups.

$$\begin{aligned} \frac{\partial H}{\partial S_i} &= -[(-u_i - q_iS_i\sum_{j=1}^{m}f_{ij}\frac{I_j}{N_j})\lambda_{S_i} + (q_iS_i\sum_{j=1}^{m}f_{ij}\frac{I_j}{N_j})\lambda_{I_i} + u_i\lambda_{R_i} + u_i\lambda_z] \\ \frac{\partial H}{\partial I_i} &= -[1 - (q_iS_i\frac{f_{ii}}{N_i})\lambda_{S_i} + (q_iS_i\frac{f_{ii}}{N_i} - \gamma_i - \delta_i)\lambda_{I_i} + \gamma_i\lambda_{R_i} + \delta_i\lambda_{D_i}] \\ \frac{\partial H}{\partial R_i} &= 0 \\ \frac{\partial H}{\partial z} &= 0 \end{aligned} \tag{C.6}$$

The Hamiltonian H is minimized with respect to the control (at the optimal control). We differentiate H with respect to u on the set Ω and arrive at the following optimality condition:

$$\begin{aligned} \frac{\partial H}{\partial \mu_i} &= u_iW - \lambda_{S_i}S_i \\ \mu_i* &= \lambda_{S_i}\frac{S_i}{W} \end{aligned} \tag{C.7}$$

BIOGRAPHICAL SKETCH

Romarie Morales Rosado was born in New York and moved to Puerto Rico with her family when she was only 8 months old. Romarie knew first-hand the hardships and trials that an individual has to endure when faced with a life-threatening illness, both of her sisters faced and overcome childhood extreme illness while Romarie faced heart problems when she was in her primary grades of school. With this past experience it became clear to her that she wanted to make a societal contribution by aiding in research to mitigate the effect of diseases. To help reach her goal, she began her academic career as a political science major and, after participating in a summer internship at Princeton University, she realized getting solid background skills in mathematics and statistics will help her better attain the goal of effectively implementing health policies. She became a full-time applied mathematics graduate student at Arizona State University (ASU) under the supervision of Carlos Castillo-Chavez. As a graduate student, Romarie received several research awards and fellowships and published several peer-review journal articles and technical reports. Romarie?s accolades include 1st place awards at the Emerging Researchers National (ERN) Conference and the Infinite Possibilities Conference, an Alfred P. Sloan Foundation fellowship, a Western Alliance for Expanding Student Opportunities (WAESO) fellowship, an Arizona State University College of Liberal Arts and Sciences scholarship, a Graduate Assistantships in Areas of National Need (GAANN) fellowship, and a Finalist Hispanic Fund scholarship. Some of her publications include: A note on the use of influenza vaccination strategies when supply is limited (2011), Ultrasonic, non- invasive classification/discrimination of liquid explosives (LEs) and threat liquids from non-threat liquids in sealed containers (2009), and The impact of the effectiveness of needle exchange programs on addiction-treatment dynamics (2014). Romarie enjoys giving back to the academic and nonacademic community. She was an active member and leader of several student-based organizations, like Society for Industrial and Applied Mathematics @ ASU Chapter, Latin@ Graduate Student Alliance, SHADES mentorship program, and the applied math representative for the Association for All Graduate Students. Romarie was also part of the selection committee for the Mathematical and Theoretical Biology Institute undergraduate summer research program, serves as a Judge for MGE@MSA undergraduate research presentations, volunteers as a grader and tutor for Sonia Kovalevsky day events were the focus is to promote and nurture interest from high school females in mathematics and for Super Saturdays (tutoring kids from Grade 6-12 in mathematics) in downtown phoenix. Recently Romarie accepted a full time teaching position at Arizona State University (Downtown Phoenix Campus), where she is able to nurture future leaders and continue her research on the control of infectious diseases in structured populations.