Mathematical Models and Inverse Algorithms for Childhood Infectious Diseases with Vaccination - Case Studies in Measles

by

Chaochao Jin

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in
Applied Mathematics

Department of Mathematical and Statistical Sciences
University of Alberta

Abstract

Children are at a high risk of infection since they have not yet developed mature immunity. Childhood infectious diseases, such as measles, chicken pox and mumps, remain epidemic and endemic around the world. Yet, their dynamics are still not fully understood. SIR-type models have been proposed and widely applied to understand and control infectious diseases, and the $SEIR$ model has been frequently applied to study childhood infectious diseases. In this thesis, we improve the classic $SEIR$ model by separating the juvenile group and the adult group to better describe the dynamics of childhood infectious diseases. We perform stability analysis to study the asymptotic dynamics of the new model, and perform sensitivity analysis to uncover the relative importance of the parameters on infection.

The transmission rate is a key parameter in controlling the spread of an infectious disease as it directly determines the disease incidence. However, it is essentially impossible to measure the transmission rate due to ethical reasons. We introduce an inverse method for our new model, which can extract the time-dependent transmission rate from either prevalence data or incidence data in existing open databases. Pre- and post-vaccination measles data sets from Liverpool and London are applied to estimate the time-varying transmission rate.

The effectiveness of vaccination has been widely discussed and studied in epidemiology. Outbreaks can still occur if the percentage of susceptible individuals who take the vaccination is low or the vaccination itself is not sufficiently effective. We further extend our model by adding a vaccination term for all children to predict the date and the infection number of a possible measles outbreak peak in the Province of Alberta.

Acknowledgements

First of all, I owe my deepest appreciation to the tremendous support and invaluable guidance of my supervisor, Dr. Hao Wang. It is him who introduced me to the area of mathematical biology and guided me with his extensive knowledge and patience. It is him who met me regularly and can always give me professional suggestions and inspired me on my research. I can never finish my thesis without his help and guidance.

Moreover, I extend my appreciation to the department of mathematical and statistical science who offered me the chance to study in University of Alberta. It's one of the most valuable experiences and had made a profound influence to my life. I also appreciate the financial support from the department which covered most of my life expense during my graduate study.

Furthermore, I want to give my gratitude to the rest of my committee, Dr. Michael Li, Dr. Christoph Frei, Dr. Bin Han, for their support, guidance and valuable suggestions.

I would also like to thank all my friends and colleagues. Thanks for their support and company during my study in Canada. Especially, I would like to thank Jude Kong who always gave me valuable suggestions and help me revise my thesis. on how to organize my thesis.

Last but not the least, thanks my parents and brother who live in the other side of the earth allowed me to left them and encouraged me peruse for my research career.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background information

Infectious diseases are diseases caused by pathogenic microorganisms, such as bacteria, viruses, parasites, and fungi. They can be transmitted directly or indirect. Direct transmission occurs when there is a physical contract between an infective and a susceptible person. Examples of infectious diseases transmitted directly include chickenpox, influenza, measles. Indirectly transmission, on the contrary, occurs when a susceptible individual comes into contact with a contaminated reservoir. Such diseases can be viral in nature, like rotavirus disease or hantavirus pulmonary syndrome; bacterial, such as cholera or legionellosis; or parasitic, such as schistosomiasis, cryptosporidiosis or giardiasis [14]. Most infectious diseases are acute. Acute infectious diseases develop rapidly but only last for a short period of time. Symptoms of an acute infectious disease are usually severe, for example, chicken pox causes red skin rashes, fever, blisters and so on. Chronic infectious diseases on the other hand, develop slowly and last for a long period of time with mild symptoms.

Infectious diseases are a challenge for public health. They have caused an enormously high percentage of deaths. For instance, 75-200 million people were killed by

the Black Death in Europe from 1346 to 1353 [10], more than 20,000 people died in Canada during the typhus epidemic in 1847-48 [8], and more than 700 people were killed during the SARS outbreak in southern China [27]. Uncovering the transmission mechanism of an infectious disease is of significant importance in controlling it and ultimately getting rid of them.

In epidemiology, an infectious disease is said to be endemic when it always prevails. An epidemic occurs when the incidence of a disease increases sharply in a short period of time. An epidemic is restricted to a local region. However, if it spreads to other countries or continents all over the world and affects a substantial number of people simultaneously, it is termed as pandemic. Examples of pandemics in history include smallpox and tuberculosis, as well as HIV and HIN1 pandemic [11].

The study of the transmission and control of infectious diseases is called epidemiology. Epidemiologists track the occurrence of infectious diseases using two measures: incidence and prevalence. Incidence is the number of new cases of a disease in a given area during a given period of time. It is determined by the length of the duration and the dynamics of the disease itself which reflects the increase rate of new patients. Prevalence is the number of total cases of a disease in a given area during a given period of time.

Children are more susceptible to infectious diseases for a number of reasons including immature immunity. Infectious diseases are a leading cause of death among children. One of these diseases that kill many young people yearly is measles. According to the World Health Organization, approximately 122,000 people died in the whole world from measles in 2012 - mostly children under the age of five [28].

Fortunately, most childhood infectious diseases like measles are vaccine-preventable. Accelerated immunization activities have a major impact on reducing measles-induced deaths. According to data from Alberta Health, before 1954, about 5000 measles cases were confirmed every year, and caused 400 cases of brain inflamma-

tion. Since its two-dose vaccine regimen was introduced in 1996, measles cases have reduced dramatically. In 1995, 2362 cases of measles were reported in Canada [30] and in 2005, only 6 cases were reported [29]. Although vaccination has been efficient in controlling most infectious diseases, outbreaks of a disease can still occur if the vaccination rate is too low or the pathogen evolves to escape the vaccine-induced immune. Actually, a high rate of vaccination, not necessarily a hundred percent, can keep the number of patients low (no outbreak). What vaccination rate should the government and the public health organizations strive to reach in order to prevent disease outbreaks? What vaccination policies should the government make under different situations? We introduce different vaccination strategies and explore those questions in chapter 3.

In this thesis, we use mathematical models in terms of ordinary differential equations to study directly transmitted acute childhood infectious diseases.

1.2 Mathematical models

A mathematical model is a description of a real system using mathematical concepts and language. Main roles of mathematical models are *understanding* and *prediction*. Models can be used to understand how an infectious disease spreads and how various factors affect its dynamics [14]. Models can make reasonably accurate predictions and allow early control decisions and the estimation of epidemic features (like peak amplitude and time).

1.2.1 The SIR model

A simple model, the SIR model, was first introduced and studied by Kermack and Mckendrick (1927) [15] to describe infectious diseases with assumptions listed below:

- The population is divided into three groups: susceptible, infective, and recovered. We use S, I, and R to denote the fractions of these three groups, respectively,
- The total population size remains constant, i.e. $S + I + R = 1$.
- The population is mixed homogeneously and every individual travels randomly in the given region.
- People gain permanent immunity after recovery or vaccination.

A mathematical model was build on how individuals move between groups. The SIR model has two key transitions: $S \to I$ and $I \to R$. First, individuals move from group S to group I when susceptible individuals get infected via physical contacting with infectious individuals. The transmission strength depends on many factors such as the level of susceptibles, the prevalence of infectives, the social network, and the success probability of transmission during a contact, etc. The commonly used transmission term is βSI whose derivations can be dated back to Kermack-McKendrick [15]. The transmission term βSI is determined by S, I and the transmission rate β. It is clear that more susceptibles and infectives will lead to more patients. The transmission rate β is a composite result of many factors such as seasonality, the behavior of host (holidays, etc.), prevalence of pathogen, and virulence of pathogen. Second, individuals move from group I to group R only when they recover by their immune response or treatment. We assume that the parameter ν is the recovery rate.

With the above analysis, we obtain a flow diagram with transition terms on top of black arrows which indicate the transition directions.

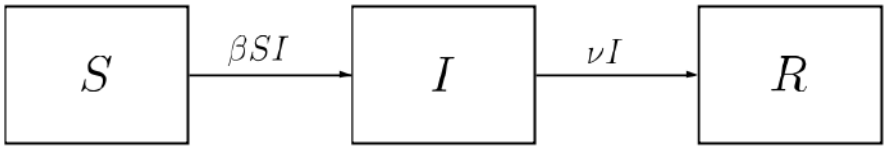


Figure 1.1: A flow diagram demonstrating the relationships among Susceptible (S), Infective (I) and Recovered (R).

According to the flow diagram in Fig 1.1, we obtain the following SIR model:

$$\frac{dS}{dt} = -\beta(t)S(t)I(t), \tag{1.2.1}$$
$$\frac{dI}{dt} = \beta(t)S(t)I(t) - \nu I(t), \tag{1.2.2}$$
$$\frac{dR}{dt} = \nu I(t). \tag{1.2.3}$$

This model is the most basic framework of the SIR-type model. With this model, the disease will ultimately go extinct since there is no susceptible influx. If we are interested in the endemic dynamics of an infectious disease, clearly the demography is important. A common way of introducing vital rates into the SIR model is to assume that there is a average lifespan of every individual, $1/\delta$ years. Thus the natural mortality rate is $\delta \cdot \text{year}^{-1}$. Again, we assume that the total population is constant. With these assumptions, we obtain a generalized SIR model

$$\frac{dS}{dt} = \delta - \beta S(t)I(t) - \delta S(t),$$
$$\frac{dI}{dt} = \beta S(t)I(t) - \nu I(t) - \delta I(t),$$
$$\frac{dR}{dt} = \nu I(t) - \delta R(t).$$

The birth rate is the same as the death rate δ since we assume that the total population is constant.

1.2.2 The $SEIR$ model

For many diseases, individuals can not infect susceptibles immediately after they get infected. It is not until human bodies are colonized by enough pathogens can they infect other susceptibles as shown in Figure 1.2. They are neither susceptibles nor infectives before the pathogen abundance reaches the threshold. Researchers call this exposed period and classify these individuals in a new group E. In fact, most childhood infectious diseases have this exposed period.

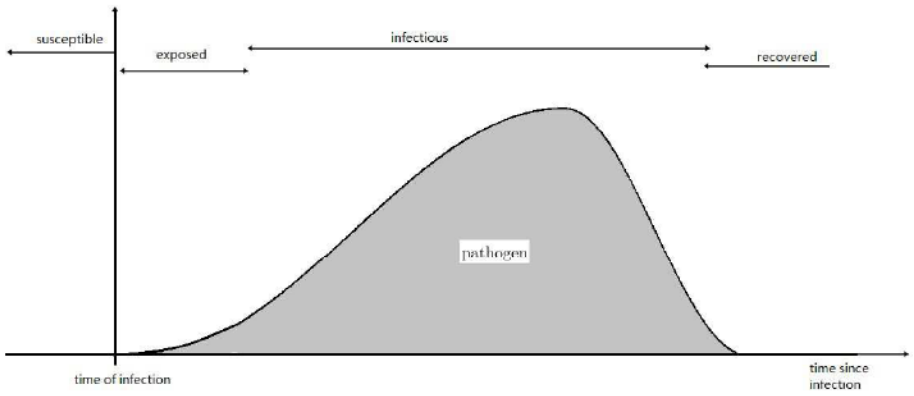


Figure 1.2: The time-line of the infections, showing the dynamics of the pathogen and labeling the infection classes: Susceptible (S), Exposed (E), Infective (I), and Recovered (R).

In the $SEIR$ model, the transition between susceptible and exposed is the fraction of new individuals get infected over a unit time, therefore it is also βSI. We assume that the average duration from exposed to infective is $1/a$, thus the transition term from group E to group I is aE.

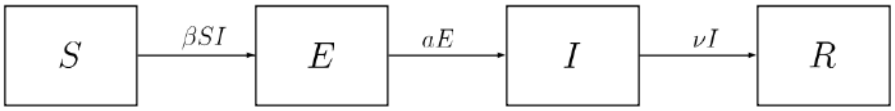


Figure 1.3: Flow diagram of the $SEIR$ model explaining the transitions among Susceptibles (S), Exposed (E), Infectives (I) and Recovered (R).

According to the flow diagram, we obtain the $SEIR$ model as

$$\begin{aligned}
\frac{dS}{dt} &= \delta - \beta(t)S(t)I(t) - \delta S(t),\\
\frac{dE}{dt} &= \beta S(t)I(t) - aE(t) - \delta E(t),\\
\frac{dI}{dt} &= aE(t) - \nu I(t) - \delta I(t),\\
\frac{dR}{dt} &= \nu I(t) - \delta R(t).
\end{aligned}$$

Vaccination is considered as the most effective way of controlling endemic infectious diseases. The most common way of vaccination is pediatric vaccination, which vaccinates newborn babies and young infants. We denote the vaccination rate p. Vaccinated newborn babies will be classified as recovered as they will built immunity after vaccination. The $SEIR$ model with vaccination can be written as

$$\begin{aligned}
\frac{dS}{dt} &= \delta(1-p) - \beta(t)S(t)I(t) - \delta S(t),\\
\frac{dE}{dt} &= \beta S(t)I(t) - aE(t) - \delta E(t),\\
\frac{dI}{dt} &= aE(t) - \nu I(t) - \delta I(t),\\
\frac{dR}{dt} &= \nu I(t) - \delta R(t) + \delta p.
\end{aligned}$$

1.3 Two existing algorithms to extract the transmission rate $\beta(t)$

Studying the transmission rate β can yield critical information on the disease and can help determine which control strategy should be applied. However, the transmission rate is essentially impossible to measure directly due to ethical and economic reasons. Mathematical models are a powerful tool to estimate and study the time-varying transmission rate.

1.3.1 The prevalence algorithm

An algorithm had been proposed recently by Pollicott et al. (2012) to extract the time-dependent transmission rate $\beta(t)$ from the existing prevalence data using a novel inverse method. We call it the prevalence algorithm. We recall the theorem on the inverse problem below [19]:

Theorem 1: Given a smooth function $f(t)$, $\nu > 0$, $\beta_0 > 0$, and $T > 0$, there exists $K > 0$ such that if $\beta_0 < K$, the system (1.2) has a solution $\beta(t)$ with $\beta(0) = \beta_0$ such that $I(t) = f(t)$ for $0 \leq t \leq T$ if and only if $f'(t)/f(t) > -\nu$ for $0 \leq t \leq T$.

This theorem is formatted into the following steps for extracting the transmission rate:

Step 1 Smoothly interpolate the infection data with a spline or trigonometric function to generate a smooth $f(t)$. Check condition 1: $f'(t)/f(t) > -\nu$, where ν is the removal rate .

Step 2 Compute the function $p(t) = (f''(t)f(t) - f'(t)^2)/(f(t)(f'(t) + \nu f(t)))$. Condition 1 prevents a zero denominator.

Step 3 Choose $\beta(0)$ and compute the integral $P(t) = \int_0^t p(\tau)d\tau$. Check condition 2: $\beta(0) < 1/\int_0^T e^{P(s)}f(s)ds$, where T is the time length of the infection data. Alternatively, choose $\beta(0)$ sufficiently small to satisfy condition 2.

Step 4 Apply the formula $\beta(t) = 1/[e^{-P(t)}/\beta(0) - e^{-P(t)}\int_0^t e^{P(s)}f(s)ds]$ to compute

$\beta(t)$ on the given interval $[0, T]$.

This algorithm involves the first and second derivatives of $I(t)$. But, in practice, data is discrete. Hence, as the method used in [19], we first interpolate discrete data with a smooth spline or trigonometric function $f(t)$. We then use $f(t)$ to replace $I(t)$ in the algorithm where we need a smooth function.

For measles in two cities in UK, Liverpool and London, we have weekly measles notification data during 1944-1994 from public databases such as the International Disease Data Archive [13] and Bolker's measles data archive [12]. To aggregate weekly infectious data into prevalence data, we simply sum up the weekly data as previous studies [5,17]. For a week across two months, this weekly infectious data is separated into be two parts. For example, if one week has three days in May and four days in June, then we multiply the number of notification during this week by $3/7$ and incorporate it into the May data, and we multiply the number of notifications during this week by $4/7$ and incorporate it into the June data.

To understand the seasonal dependence of disease transmission, we can calculate the dominant frequency of transmission rate β through the discrete Fourier transform. Figure 1.4 shows the extracted time-dependent transmission rate $\beta(t)$ with initial value $\beta(0) = 140$ and its frequency spectrum.

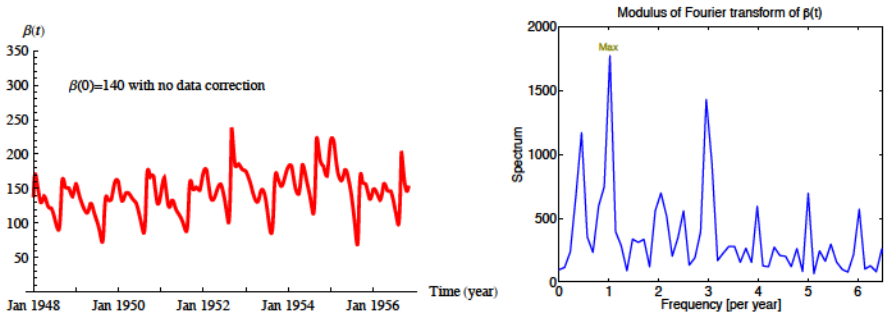


Figure 1.4: The transmission rate β extracted from England&Walse and its modulus of Fourier transform from 1948 to 1958. We use the $SEIR$ model with the parameter values $\nu = 52/12/\text{month}$, $a = 52/12/\text{month}$, $\delta = 1/70/12/\text{month}$, and the initial value $\beta(0) = 140/year$.

The modulus of Fourier transform of a discrete data set will peak at where the data set has its strongest frequency. We can see that there are two dominant frequencies: 1 and 3 per year. The first frequency is obviously from seasonality. The second one reflects the three big holidays in UK: Christmas, Easter, and summer vacation.

1.3.2 The incidence algorithm

Another algorithm was proposed by Haleder in 2011 [9]. The formula uses incidence data instead of prevalence data. Also, no derivatives are needed, while the prevalence algorithm needs both the first and the second derivatives. Thus the incidence algorithm should be more stable numerically. We call it the incidence algorithm. We extend it from the SIR model to the $SEIR$ model.

We denote $\omega(t)$ the number of new cases over a unit time, i.e. $\omega(t) = \beta SI$. We can rewrite the $SEIR$ model as

$$\begin{aligned}
\frac{dS}{dt} &= \delta - \omega(t) - \delta S(t), \\
\frac{dE}{dt} &= \omega(t) - aE(t) - \delta E(t), \\
\frac{dI}{dt} &= aE(t) - \nu I(t) - \delta I(t), \\
\frac{dR}{dt} &= \nu I(t) - \delta R(t).
\end{aligned}$$

According to the $SEIR$ model, we derive a formula to extract the transmission rate $\beta(t)$ from incidence data.

Theorem 2: For the $SEIR$ model, the time-dependent transmission function is

$$\beta(t) = \frac{\omega(t)}{S(t)I(t)}$$

where

$$\begin{aligned}
S(t) &= S(0)e^{-\delta t} + \int_0^t (\delta - \omega(s))e^{-\delta(t-s)}ds, \\
I(t) &= I(0)e^{-(\nu+\delta)t} + a\int_0^t (E(0)e^{-(a+\delta)s} + \int_0^s \omega(\sigma)e^{-(a+\delta)(s-\sigma)}d\sigma)e^{-(\nu+\delta)(t-s)}ds.
\end{aligned}$$

Its proof is similar to the proof of theorem 5 in chapter 2.

1.3.3 Connection between prevalence algorithm and incidence algorithm

We believe that prevalence and incidence formulas are the same formulas, but in different forms. In this subsection, we prove this statement using the SIR model as an example.

With $\omega(t) = \beta(t)S(t)I(t)$, the SIR model can be rewritten as

$$S'(t) = -\omega(t), \tag{1.3.1}$$

$$I'(t) = \omega(t) - \nu I(t), \tag{1.3.2}$$

$$R'(t) = \nu I(t). \tag{1.3.3}$$

From the prevalence algorithm, we know that

$$\beta(t) = \frac{e^{P(t)}}{\frac{1}{\beta_0} - \int_0^t e^{P(s)} f(s) ds}, \tag{1.3.4}$$

where $P(t) = \int_0^t p(t)dt$ and $p(t) = \frac{f''(t)f(t) - f'(t)^2}{f(t)(f'(t) + \nu f(t))}$. Solving (1.3.2) by method of constant variation, we can see that $f(t) = e^{-\nu t} f(0) + \int_0^t \omega(s) e^{-\nu(t-s)} ds$, where $f(t)$ is the smooth function from interpolation of discrete data of $I(t)$. If we plug the above equation into (1.3.4), then $\beta(t)$ can be rewritten in incidence $\omega(t)$. Instead of plug it in directly, we use the relationship between $I(t)$ and $\omega(t)$ as equation (1.3.2) to simplify the process . Take the derivate of (1.3.2) respect to t , we get

$$f''(t) = \omega'(t) - \nu f'(t) = \omega'(t) - \nu(\omega(t) - \nu f(t)) = \omega'(t) - \nu\omega(t) + \nu^2 f(t).$$

Equation (1.3.4) can be rewrite as $f'(t) + \nu f(t) = \omega(t)$.

Therefore,

$$\begin{aligned} p(t) &= \frac{f''(t)f(t) - f'(t)^2}{f(t)(f'(t) + \nu f(t))} \\ &= \frac{(\omega'(t) - \nu\omega(t) + \nu^2 f(t))f(t) - (\omega(t) - \nu f(t))^2}{f(t)\omega(t)} \\ &= \frac{\omega'(t)f(t) + \nu\omega(t)f(t) - \omega^2(t)}{f(t)\omega(t)} \\ &= \frac{\omega'(t)}{\omega(t)} + \nu - \frac{\omega(t)}{f(t)}. \end{aligned}$$

Integrate $p(t)$ from 0 to t:

$$\begin{aligned}
P(t) &= \int_0^t p(s)ds \\
&= \int_0^t \frac{\omega'(s)}{\omega(s)} + \nu - \frac{\omega(s)}{f(s)}ds \\
&= \int_0^t \frac{\omega'(s)}{\omega(s)}ds - \int_0^t \frac{\omega(s) - \nu f(s)}{f(s)}ds \\
&= \int_0^t \frac{\omega'(s)}{\omega(s)}ds - \int_0^t \frac{f'(s)}{f(s)}ds \\
&= \ln\frac{\omega(t)}{\omega(0)} - \ln\frac{f(t)}{f(0)} \\
&= \ln\frac{\omega(t)f(0)}{\omega(0)f(t)},
\end{aligned}$$

i.e. $e^{P(t)} = \frac{\omega(t)f(0)}{\omega(0)f(t)}$. Plug it into (1.3.4):

$$\begin{aligned}
\beta(t) &= \frac{e^{P(t)}}{\frac{1}{\beta_0} - \int_0^t e^{P(s)}f(s)ds} \\
&= \frac{\frac{\omega(t)f(0)}{\omega(0)f(t)}}{\frac{1}{\beta_0} - \int_0^t \frac{\omega(s)f(0)}{\omega(0)f(s)}f(s)ds} \\
&= \frac{\omega(t)}{f(t)(\frac{\omega(0)}{f(0)\beta_0} - \int_0^t \omega(s)ds)} \\
&= \frac{\omega(t)}{f(t)(S(0) - \int_0^t \omega(s)ds)} \\
&= \frac{\omega(t)}{(e^{-\nu t}f(0) + \int_0^t \omega(s)e^{-\nu(t-s)}ds)(S(0) - \int_0^t \omega(s)ds)} \\
&= \frac{\omega(t)}{f(t)S(t)}.
\end{aligned}$$

This is consistent with the incidence algorithm.

1.4 Outline

In chapter 2, we modify the $SEIR$ model to the $SEIRA$ model which better describes childhood infectious diseases. We analyze its positivity, boundedness, and stability. Next, we perform sensitivity analysis of the steady state, the peak value and time of the disease outbreak with respect to all parameters. In the second part of this chapter, we extend both the prevalence and incidence algorithms for the $SEIRA$ model and apply them to real pre-vaccination measles data from Liverpool and London.

In chapter 3, we first consider pediatric vaccination. We perform qualitative and sensitivity analysis again, but the emphasis will be focused on how vaccination rate can affect the dynamics of $I(t)$. We study how different transmission and vaccination rates affect the outbreak peak value, the outbreak peak time, and the endemic steady state of an infectious disease. We also extract $\beta(t)$ from post-vaccination measles data from London and Liverpool during the period 1974-1986. In the last part, we consider a vaccination strategy for all susceptible kids to study the present measles outbreak in the Province of Alberta.

Chapter 2

The $SEIRA$ model

2.1 Derivation of the $SEIRA$ model

Recall that when introducing the $SEIR$ model, we divide the total population into four compartments: susceptible, exposed, infective and recovered. This model can be applied to all infectious diseases satisfying its assumptions. However, it is not suitable to apply this model when considering childhood infectious diseases since, in the $SEIR$ model, adults who have never been infected nor vaccinated are also considered as susceptible.

When studying childhood infectious diseases, we first classify the population as the adult group (group A) and the juvenile group. Then we divide the juvenile group into susceptible (S), exposed (E), infective (I), and recovered (R). We modify the $SEIR$ model to the $SEIRA$ model based on our new classification.

We used to consider natural death rate for every group. But children will not die naturally. They die only because of some specific reasons like accidents or diseases and their death rate is much lower than the natural death rate. Therefore, we ignore nature death rate for the juvenile group. Instead, we consider growth rate as children will grow up and no longer be susceptible. We assume that children will be moved

to group A in rate g, i.e. people under $1/g$ years old will be considered as juvenile. Transition terms between S, E, I and R are the same with it in the $SEIR$ model. Figure 2.1 presents the flow diagram of the $SEIRA$ model.

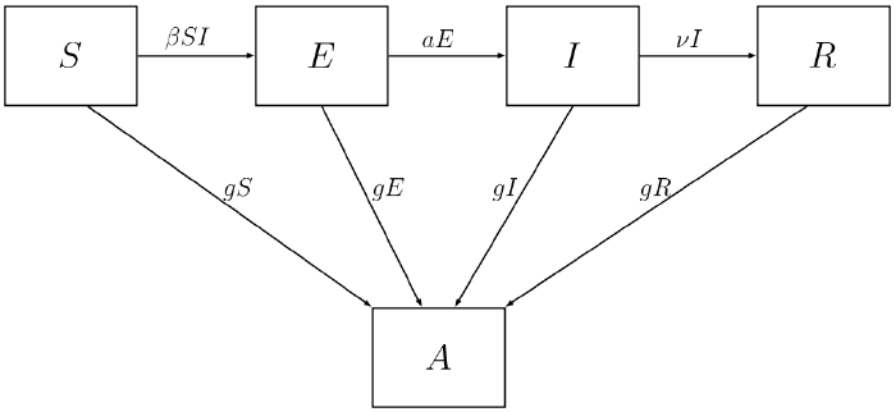


Figure 2.1: The $SEIRA$ flow diagram. A is adult group. Juvenile will be moved to group A in rate g. Flow of birth and death are omitted for clarity.

As a conclusion, we obtain the $SEIRA$ model

$$\frac{dS}{dt} = \delta A(t) - \beta(t)S(t)I(t) - gS(t), \tag{2.1.1}$$
$$\frac{dE}{dt} = \beta(t)S(t)I(t) - aE(t) - gE(t), \tag{2.1.2}$$
$$\frac{dI}{dt} = aE(t) - \nu I - gI, \tag{2.1.3}$$
$$\frac{dR}{dt} = \nu I(t) - gR(t), \tag{2.1.4}$$
$$\frac{dA}{dt} = g(S(t) + E(t) + I(t) + R(t)) - \delta A(t), \tag{2.1.5}$$

with all the parameters and their values for measles listed in table 2.1 (parameters p and q are vaccination rate which we will use later).

Parameter	Value	Description	Units
p	0 -100%	Pediatric vaccination rate	
δ	1/64	Natural death/birth Rate	year $^{-1}$
β	≈ 1000	Transmission rate	year^{-1}
g	1/16	Growth rate	year $^{-1}$
ν	52	Recovery rate	year $^{-1}$
a	52	Rate at which exposed individuals become infective	year $^{-1}$
q	0 -100%	Vaccination rate for all children	year $^{-1}$

Table 2.1: Parameter description and values for measles

Values of rate a and ν for measles are from Anderson and May [2, 23]. Rate $a = 52/year$ and $\nu = 52/year$ mean that the average time duration of exposed and infective are 1 $year/52 = 1week$. We assume that the average life span is 80 years and only kids under 16 are susceptible to measles. Therefore, $g = 1/16 \cdot year^{-1}$ for growth rate and $\delta = 1/64 \cdot year^{-1}$ for the natural death rate for the population between 17 and 80 years old. Values of p and q are mainly determined by vaccination policy, which can be as low as 0 % if there is no vaccination or as high as 100% if every individual is vaccinated. Value of transmission rate β is time dependent and usually considered as a sinusoid function. According to Tidd. et al. [24], we choose $\beta = 1000\ year^{-1}$ on average.

2.2 Qualitative analysis

In this section we analyze positivity, boundedness, and stability of the $SEIRA$ model. Positivity describes the property that a system will stay positive if starts with a positive initial value. An ODE system $\frac{dy_i}{dt} = f_i(y_1, y_2, ..., y_n)$ $(n \in \mathbb{N})$ has positivity if $\forall\ 1 \leq i \leq n,\ f_i \geq 0$ when $y_i = 0$ and $y_j \geq 0,\ j \neq i$. Denote $S,\ E,\ I,\ R,\ A$ as f_i and the functions in the right hand side of the equal signs in (2.1) as $y_i,\ for\ 1 \leq i \leq 5$, respectively. Therefore $f_1 = \delta y_5 \geq 0, f_2 = \beta y_1 y_3 \geq$

$0, f_3 = ay_2 \geq 0, f_4 = \nu y_3 \geq 0, f_5 = g(y_1+y_2+y_3+y_4) \geq 0$. Hence (2.1) has positivity. Also, we notice that $S,\ E,\ I,\ R$, and A sum to be one and their derivative sum to be zero. Combining with positivity, we conclude that the solution of (2.1) will stay in $\{(S, E, I, R, A)\colon S \geq 0, E \geq 0, I \geq 0, R \geq 0, A \geq 0, S + E + I + R + A = 1\}$ if the system starts at a positive initial value.

Stability is a tool to study the endemic behavior of a system. To analyze the stability of the $SEIRA$ system (2.1), we have to figure out all the equilibria first. Assuming that $(S^*, E^*, I^*, R^*, A^*)$ is an equilibrium, then it should satisfy

$$0 = \delta A^* - \beta S^* I^* - gS^*, \tag{2.2.1}$$

$$0 = \beta S^* I^* - (a + g)E^*, \tag{2.2.2}$$

$$0 = aE^* - (\nu + g)I^*, \tag{2.2.3}$$

$$0 = \nu I^* - gR^*, \tag{2.2.4}$$

$$0 = g - (g + \delta)A^*. \tag{2.2.5}$$

From (2.2.5), we know that $A^* = \frac{g}{g+\delta}$. From (2.2.3), we know that $E^* = \frac{\nu+g}{a}I^*$. Plug it into (2.2.2) we get

$$\beta S^* I^* - (a + g)\frac{\nu + g}{a}I^* = 0$$

$$\Rightarrow (\beta S^* - \frac{(a + g)(\nu + g)}{a})I^* = 0$$

$$\Rightarrow S^* = \frac{(a + g)(\nu + g)}{a\beta} \quad or \quad I^* = 0$$

If $I^* = 0$, then $E^* = \frac{\nu+g}{a}I^* = 0$, $R^* = \frac{\nu}{g}I^* = 0$, $S^* = \frac{\delta}{g}A^* = \frac{\delta}{g+\delta}$.

If $S^* = \frac{(a+g)(\nu+g)}{a\beta}$, then

$$
\begin{aligned}
I^* &= \frac{\delta A^* - gS^*}{\beta S^*} = \frac{ag\delta}{(a+g)(g+\delta)(\nu+g)} - \frac{g}{\beta}, \\
E^* &= \frac{\nu+g}{a} I^* = \frac{g\delta}{(a+g)(g+\delta)} - \frac{g(\nu+g)}{a\beta}, \\
R^* &= \frac{\nu}{g} I^* = \frac{a\nu\delta}{(a+g)(g+\delta)(\nu+g)} - \frac{\nu}{\beta}.
\end{aligned}
$$

Hence, there are two equilibria:

the disease-free equilibrium

$$
(S_1^*,\ E_1^*,\ I_1^*,\ R_1^*,\ A_1^*) = (\frac{\delta}{g+\delta},\ 0,\ 0,\ 0,\ \frac{g}{g+\delta})
$$

and the endemic equilibrium

$$
(S_2^*,\ E_2^*,\ I_2^*,\ R_2^*,\ A_2^*) = (\frac{(a+g)(\nu+g)}{a\beta},\ \frac{g\delta}{(a+g)(g+\delta)} - \frac{g(\nu+g)}{a\beta},\ \frac{ag\delta}{(a+g)(g+\delta)(\nu+g)} - \frac{g}{\beta},
$$

$$
\frac{a\nu\delta}{(a+g)(g+\delta)(\nu+g)} - \frac{\nu}{\beta},\ \frac{g}{g+\delta}).
$$

Theorem 3 (Local stability): Assume that all the parameters are positive. We conclude that

- When $a\delta\beta < (a+g)(g+\delta)(\nu+g)$, the disease-free equilibrium is locally asymptotic stable and the endemic equilibrium is not feasible ;
- When $a\delta\beta > (a+g)(g+\delta)(\nu+g)$, the endemic equilibrium is locally asymptotic stable and the disease-free equilibrium is unstable .

Proof. We first calculate the Jacobian matrix of the $SEIRA$ model. Since $S,\ E,\ I,\ R,\ A$ sum to be 1, there are only four free variables. To calculate the Jacobian matrix, we only need to consider any four of them. We ignore

the fourth equation (2.1.4) and obtain the Jacobian matrix as

$$J(S,E,I,A) = \begin{pmatrix} -\beta I - g & 0 & -\beta S & \delta \\ \beta I & -(a+g) & \beta S & 0 \\ 0 & a & -(\nu+g) & 0 \\ 0 & 0 & 0 & -(g+\delta) \end{pmatrix}.$$

- When $a\delta\beta < (a+g)(g+\delta)(\nu+g)$, notice that I_2^* can be rewritten as

$$I_2^* = \frac{g}{\beta(a+g)(g+\delta)(\nu+g)}(a\delta\beta - (a+g)(g+\delta)(\nu+g)),$$

thus $I_2^* < 0$, $R_2^* = \frac{\nu}{g} I_2^* < 0$, and $E_2^* = \frac{\nu+g}{a} I_2^* < 0$, i.e. the endemic equilibrium is not feasible when $a\delta\beta < (a+g)(g+\delta)(\nu+g)$.

For the disease-free equilibrium, the Jacobian matrix is

$$J(S_1^*, E_1^*, I_1^*, A_1^*) = \begin{pmatrix} -g & 0 & -\frac{\delta}{g+\delta}\beta & \delta \\ 0 & -(a+g) & \frac{\delta}{g+\delta}\beta & 0 \\ 0 & a & -(\nu+g) & 0 \\ 0 & 0 & 0 & -(g+\delta) \end{pmatrix}$$

and

$$\lambda I - J(S_1^*, E_1^*, I_1^*, A_1^*) = \begin{pmatrix} \lambda+g & 0 & \frac{\delta}{g+\delta}\beta & -\delta \\ 0 & \lambda+a+g & -\frac{\delta}{g+\delta}\beta & 0 \\ 0 & -a & \lambda+\nu+g & 0 \\ 0 & 0 & 0 & \lambda+g+\delta \end{pmatrix}.$$

The characteristic equation is

$$(\lambda+g)(\lambda+\nu+g)(\lambda^2+(a+2g+\nu)\lambda+(a+g)(\nu+g)-\frac{a\delta}{g+\delta}\beta)=0.$$

Therefore eigenvalues of $J(S_1^*,E_1^*,I_1^*,A_1^*)$ are $\lambda_1=-g$, $\lambda_2=-(g+\delta)$ and λ_3, λ_4 which are two solutions of

$$\lambda^2+(a+2g+\nu)\lambda+(a+g)(\nu+g)-\frac{a\delta}{g+\delta}\beta=0.$$

$$\Rightarrow \lambda_{3,4}=\frac{-(a+2g+\nu)\pm\sqrt{(a+2g+\nu)^2-4((a+g)(\nu+g)-\frac{a\delta}{g+\delta}\beta)}}{2}$$

$$=\frac{-(a+2g+\nu)\pm\sqrt{(a-\nu)^2+4\frac{a\delta}{g+\delta}\beta}}{2}$$

We can see that λ_3, λ_4 are always real and both λ_3, λ_4 are negative if and only if $(a+g)(\nu+g)-\frac{a\delta}{g+\delta}\beta>0$.

According to stability theorem,

the disease-free equilibrium is stable

$\Longleftrightarrow$all eigenvalues of its Jacobian matrix is negative

$\Longleftrightarrow (a+g)(\nu+g)-\frac{a\delta}{g+\delta}\beta>0$

$\Longleftrightarrow a\delta\beta<(a+g)(g+\delta)(\nu+g).$

- When $a\delta\beta>(a+g)(g+\delta)(\nu+g)$, notice that $S_2^*=\frac{(a+g)(\nu+g)}{a\beta}>0$, $A_2^*=\frac{g}{g+\delta}>0$ if all parameters are positive. Also, $I_2^*=\frac{g}{\beta(a+g)(g+\delta)(\nu+g)}(a\delta\beta-(a+g)(g+\delta)(\nu+g))>0$, and $R_2^*=\frac{\nu}{g}I_2^*>0$, $E_2^*=\frac{\nu+g}{a}I_2^*>0$, i.e. the endemic equilibrium is feasible when all the parameters are positive and $a\delta\beta>(a+g)(g+\delta)(\nu+g)$.

  Also, when $a\delta\beta>(a+g)(g+\delta)(\nu+g)$, from the previous proof, we know that the disease-free equilibrium is unstable.

For the endemic equilibrium, once again, we calculate the Jacobian matrix:

$$J(S_2^*, E_2^*, I_2^*, A_2^*) = \begin{pmatrix} -\beta I_2^* - g & 0 & -\beta S_2^* & \delta \\ \beta I_2^* & -(a+g) & \beta S_2^* & 0 \\ 0 & a & -(\nu+g) & 0 \\ 0 & 0 & 0 & -(g+\delta) \end{pmatrix}.$$

$$\Longrightarrow$$

$$\lambda I - J(S_2^*, E_2^*, I_2^*, A_2^*) = \begin{pmatrix} \lambda + \beta I_2^* + g & 0 & \beta S_2^* & -\delta \\ -\beta I_2^* & \lambda + a + g & -\beta S_2^* & 0 \\ 0 & -a & \lambda + \nu + g & 0 \\ 0 & 0 & 0 & \lambda + g + \delta \end{pmatrix}.$$

From the above matrix, we can see that one eigenvalue of the characteristic equation is $\lambda_1 = -(g+\delta)$, and the other three are the eigenvalues of matrix

$$M = \begin{pmatrix} -\beta I_2^* - g & 0 & -\beta S_2^* \\ \beta I_2^* & -(a+g) & \beta S_2^* \\ 0 & a & -(\nu+g) \end{pmatrix}.$$

$$\Longrightarrow$$

$$\lambda I - M = \begin{pmatrix} \lambda + \beta I_2^* + g & 0 & \beta S_2^* \\ -\beta I_2^* & \lambda + a + g & -\beta S_2^* \\ 0 & -a & \lambda + \nu + g \end{pmatrix},$$

and the characteristic equation is

$$\lambda^3 + (-tr(M))\lambda^2 + (\alpha(M))\lambda + (-Det(M)) = 0,$$

where

$$tr(M) = -(\beta I_2^* + a + 3g + \nu),$$
$$\alpha(M) = (\beta I_2^* + g)(a+g) + (\beta I_2^* + g)(\nu + g) + (a+g)(\nu+g) - a\beta S_2^*,$$
$$Det(M) = -(\beta I_2^* + g)(a+g)(\nu+g) + ag\beta S_2^*.$$

To study stability, we apply the third order Routh-Hurwitz stability criterion. We first review the third order Routh-Hurwitz stability criterion.

Proposition 1 (Third-order Routh-Hurwitz stability criterion): Real parts of all solutions of a third-order polynomial $P(s) = a_3 s^3 + a_2 s^2 + a_1 s + a_0 = 0$ are negative if the coefficients satisfy $a_3 > 0, a_2 > 0, a_1 > 0, a_0 > 0$ and $a_2 a_1 > a_3 a_0$.

Compare the characteristic equation with the third order Routh-Hurwitz stability criterion and plug in $S_2^* = \frac{(a+g)(\nu+g)}{a\beta}$:

$$\begin{aligned}
a_3 &= 1 > 0,\\
a_2 &= -tr(M) = \beta I_2^* + 3g + a + \nu > 0,\\
a_1 &= \alpha(M) = (\beta I_2^* + g)(a+g) + (\beta I_2^* + g)(\nu+g) + (a+g)(\nu+g) - a\beta S_2^*\\
&= (\beta I_2^* + g)(a+g) + (\beta I_2^* + g)(\nu+g) + (a+g)(\nu+g) - a\beta\tfrac{(a+g)(\nu+g)}{a\beta}\\
&= (\beta I_2^* + g)(a + 2g + \nu) > 0,\\
a_0 &= -Det(M) = (\beta I_2^* + g)(a+g)(\nu+g) - ag\beta S_2^*\\
&= \beta I_2^*(a+g)(\nu+g) + g(a+g)(\nu+g) - ag\beta S_2^*\\
&= \beta I_2^*(a+g)(\nu+g) + g(a+g)(\nu+g) - ag\beta\tfrac{(a+g)(\nu+g)}{a\beta}\\
&= \beta I_2^*(a+g)(\nu+g) > 0,\\
a_2 a_1 &= \alpha(M) * (-tr(M))\\
&= (\beta I_2^* + g)(a + 2g + \nu)(\beta I_2^* + 3g + a + \nu)\\
&> \beta I_2^*(a+g)(\nu+g) = -Det(M) = a_3 a_0.
\end{aligned}$$

Conditions of Routh-Hurwitz stability criterion are satisfied and the endemic equilibrium is asymptotically stable if all the parameters are positive

and $a\delta\beta > (a+g)(g+\delta)(\nu+g)$.

□

To be short, theorem 3 can be rewritten as:

Assume that all the parameters are positive.

- When $R_0 \leq 1,\ \lim_{t\to+\infty}(S(t),\ E(t),\ I(t),\ R(t),\ A(t)) \longrightarrow$ Disease-free equilibrium (DFE),
- When $R_0 > 1,\ \lim_{t\to+\infty}(S(t),\ E(t),\ I(t),\ R(t),\ A(t)) \longrightarrow$ Endemic equilibrium (EE),

where $R_0 = \frac{a\delta\beta}{(a+g)(g+\delta)(\nu+g)}$ is the basic reproduction number.

Recall that when $R_0 \geq 1$, the disease can spread and when $R_0 < 1$, the disease will finally disappear. Thus, stability conclusion of theorem 3 is consistent with conclusion from R_0. We can rewrite R_0 as

$$R_0 = \frac{\beta}{\nu+g} \cdot \frac{a}{a+g} \cdot \frac{\delta}{\delta+g}.$$

We know that βSI is the number of new cases over a unit time. Hence the average number for one infective individual is $\frac{\beta SI}{I} = \beta S$. Therefore, the number of susceptibles an infective individual can infect is the average number over a unit time multiply by the average length of duration an infective indiidual stay infectious which is $\beta S \cdot \frac{1}{\nu+g}$. The fraction of the infected individuals who can finally become infective is the possibility an exposed individual will become infective in a unit time multiply by the time duration of an exposed individual stay exposed which is $a \cdot \frac{1}{a+g}$. The expected fraction of susceptibles is $\frac{\delta}{\delta+g}$ which is the value of susceptibles at the disease-free steady state. Therefore, $\beta S \cdot \frac{1}{\nu+g} \cdot \frac{a}{a+g} = \beta \cdot \frac{\delta}{\delta+g} \cdot \frac{1}{\nu+g} \cdot \frac{a}{a+g}$ is the average number of susceptibles that one infective individual can infect in the duration of infection.

2.3 Sensitivity analysis

Sensitivity analysis studies the quantity of how uncertainty factors can affect important results. The normalized forward sensitivity index is defined as

$$S.I. = \frac{p}{X^*}\frac{\partial X^*}{\partial p} \tag{2.3.1}$$

where X^* is the quantity being considered, and p is the parameter which X^*depends upon. Sensitivity indices can be positive or negative which indicate the nature of the relationship, and it is the magnitude that ranks the strength of the relationship as compared to the other parameters.

In this section, we calculate, analyze and compare the sensitivity indices of outbreak peak value, time of outbreak peak, and steady state value of $I(t)$, with respect to different parameters. When studying quantities like the peak value or the peak time which do not have explicit formulas, we compute the approximation values of their indices by calculating

$$S.I. = \frac{p}{X^*(p)}\frac{X^*(p+\Delta p) - X^*(p-\Delta p)}{2\Delta p}. \tag{2.3.2}$$

We calculate $S.I.$ with respect to one specific parameter by perturbing this parameter only and keeping the others unchanged. In our calculation in this section, we take $\Delta p = 1\%p$.

2.3.1 Sensitivity analysis of the outbreak peak value

The sensitivity indices of the amplitude of the outbreak peak show how the first epidemic depends on the parameters as seen in Table 2.2.

Parameter	Sensitivity of peak	Description
δ	0.0099	Natural death/birth Rate
β	1.7200	Transmission Rate
g	-0.0195	Growth Rate
ν	-2.8204	Removal Rate
a	1.1102	Rate at which exposed individuals become infective

.

Table 2.2: The sensitivity indices of the value of the outbreak peak respect to the parameters values $\delta = 1/64/12/month,\ \beta = 55/month,\ g = 1/16/12/month,\ \nu = 52/12/month,\ a = 52/12/month$ and initial values $S_0 = 0.2,\ E_0 = 0.002,\ I = 0.002,\ R_0 = 0.006,\ A_0 = 0.79$.

The removal rate ν has the strongest relationship to the magnitude of the outbreak peak. The negative value tells us that a lower removal rate would lead to a more severe epidemic. In contrast to the birth/death rate δ which has among the lowest of sensitivity indices, ν would thus be an important parameter to control in order to reduce the harm of an outbreak.

Both transmission rate β and rate a have strong positivity relationship to the peak outbreak as higher β would result in a higher level of group E and higher a would move more exposed to infections.

The sensitivity index with respect to the human birth/death rate μ is very low in comparison to all the others. This makes sense, because the initial peak of an epidemic occurs relatively quickly after the introduction of sick people, and the birth and death of new susceptibles would take much longer time.

The growth rate g is negative sensitive to the outbreak peak because a larger growth rate will move infections to group A faster, thus reduces the outbreak peak. Similarly with birth/death rate δ, growth rate g has a small influence on the peak value.

In fact, parameters related to demography, such as birth/death rate and growth rate, would have small influence on the outbreak level as the initial peak appears

relatively quickly. Parameters which directly related to the inputs and outputs of infections would have important influence on the initial peak. For instance, rate a determines how fast exposed individuals will become infectious, ν determines how quickly infections will be moved to group R and β determines how many susceptibles will be infected and both of them have strong relationship with the outbreak.

2.3.2 Sensitivity analysis of the outbreak peak time

Sensitivity indices of the outbreak peak time measure how the first epidemics outbreak depends on different parameters as seen in the Table 2.3.

Parameter	Sensitivity of peak time	Description
δ	-0.0011	Natural death/birth Rate
β	-0.7403	Transmission Rate
g	0.0027	Growth Rate
ν	0.3075	Removal Rate
a	-0.2908	Rate at which exposed individuals become infective

Table 2.3: The sensitivity of the outbreak peak time respect to the parameters with values $\delta = 1/64/12/month, \beta = 55/month, g = 1/16/12/month, \nu = 52/12/month, a = 52/12/month.$ and initial values $S_0 = 0.2,\ E_0 = 0.002,\ I = 0.002,\ R_0 = 0.006,\ A_0 = 0.79.$

As outlined previously, we have the same reason that birth/death rate δ and growth rate g have less influence on the outbreak time than the other three parameters.

We can see from Table 2.3 that the transmission rate β has the largest influence on the dynamics of the system. This suggests that β is a more important quantity to control to prevent outbreaks. The negative relationship tells us that a larger transmission rate would lead to a quicker outbreak.

The relationship between rate a and the time of the maximum outbreak is negative, because a higher contact rate (shorter latent period) will cause more new

infections and the timing of the maximum would be attained earlier.

The removal rate still has important effect on the outbreak time. The positive relationship between ν and the outbreak time is because patients will recover faster with larger ν thus postpones the outbreak time.

2.3.3 Sensitivity analysis of the endemic steady state

Endemic steady state determines the levels of groups of an endemic infectious disease. It represents the expectation of the final size of all groups. In the table below, we list sensitivity indices of I_2^* respects to all parameters.

Parameter	Sensitivity of I_2^*	Description
δ	1.3221	Natural death/birth Rate
β	0.6526	Transmission Rate
g	-0.3260	Growth Rate
ν	-1.6508	Removal Rate
a	0.0020	Rate at which exposed individuals become infective

Table 2.4: Sensitivity of the endemic steady state with respect to the parameters with values $\delta = 1/64/12/month, \beta = 55/month, g = 1/16/12/month, \nu = 52/12/month, a = 52/12/month$.

The endemic level of infective individuals is most sensitive to the recovery rate ν and birth/death rate δ. The negative relationship with ν is strong because recovery is the main way that infectives leave the infected component. The relationship with δ is positive as larger δ means more newborn susceptibles will possibly become infective. Rate a has a weak but positive relationship with I_2^* as expected because a larger a leads to the fact that more exposed will become infectives. Transmission rate β is also of great importance in controlling the endemic level of the infectives. The positive relationship is obvious since larger β means more susceptibles will get infected.

2.3.4 Numerical simulations

Sensitivity and stability were discussed previously and numerical simulations are presented in the following diagrams.

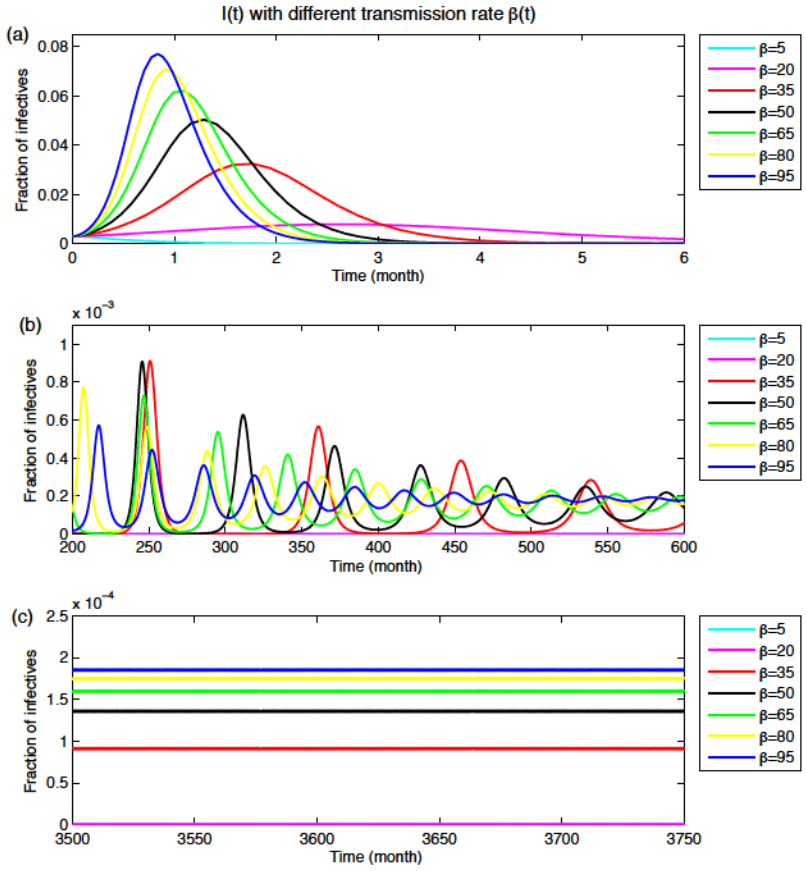


Figure 2.2: Fraction of infectives with different transmission rates β and initial value $S_0 = 0.29, E_0 = 0.003, I_0 = 0.003, R_0 = 0.003, A_0 = 0.7$ and parameters a=52/12/month, ν=52/12/month, g=1/16/12/month, δ=1/64/12/month. (a) The first peak of fraction of infectives from time 0 to 6 the first outbreak of $I(t)$. (b) Fraction of infections from time 200 to 600. We observe many outbreaks whose peak values decay. There is a threshold β_0 that when $\beta(t) < \beta_0$, the disease will go to disease-free, i.e. no more peaks after the first one and when $\beta(t) > \beta_0$, it will have small peaks after the first peak. (c) Fraction of infections from time 3500 to 3750: the disease finally goes to a steady state.

Figure 2.2 shows the dynamics of $I(t)$ with transmission rate β varying from 5 to 95 in different time intervals. The first figure (a) plots infectives $I(t)$ from time 0 to 6. During this time, the biggest peak appears quickly after the disease was first introduced into the population. The larger β is, the more severe the peak is. The second figure (b) plots infectives $I(t)$ from time 200 to 600. We know that the threshold for stability is $R_0 = \frac{a\delta\beta}{(a+g)(g+\delta)(\nu+g)} = 1$. Using the values we choose in the simulation, the threshold is $\beta_0 = 21.7$, i.e. when $\beta_0 \geq 21.7$, the disease will always exit and come back in peaks, and when $\beta_0 < 21.7$, the disease will finally disappear. We can see that when $\beta = 5,\ 20$, infectives $I(t)$ has no more small peaks compare to waning peaks when β is larger than 21.7. The third figure (c) plots $I(t)$ from time 3500 to 3750 when infetives $I(t)$ finally goes to its steady state. From (c) we can see that when condition $R_0 < 1$ or $\beta < 21.7$ ($\beta = 5, 20$ in the figure) is satisfied, the steady state of $I(t)$ is zero. When $R_0 < 1$ or $\beta < 21.7$ is satisfied, $I(t)$ finally becomes a constant, the value of which depends on the transmission rate β. When β is larger, the steady state will have a higher level.

In fact, the behavior of $I(t)$ can be much more complicated since parameters are changing while we consider them as constant in the simulation. For instance, the transmission rate β for measles is usually considered as a periodic function driven by seasonality and school holidays.

2.4 Two algorithms to extract the transmission rate $\beta(t)$ from pre-vaccination data

We extend both algorithms to the $SEIRA$ model. We first present the theoretical results of the formulas of $\beta(t)$ based on prevalence data and incidence data. Then proofs and steps of algorithms will be followed.

2.4.1 The prevalence algorithm

The prevalence algorithm is used to extract the transmission rate $\beta(t)$ from prevalence data set by solving the inverse problem.

Theorem 4: For the $SEIRA$ model, the time-dependent transmission function $\beta(t)$ satisfies the following differential equation:

$$M\beta''\beta^2 + N(\beta')^2\beta + P\beta'\beta^2 - L\beta^4 - Q\beta^3 = 0, \tag{2.4.1}$$

where

$M = D = -Hf^3,$

$N = C = 2Hf^3,$

$P = B - (2g+\delta)Hf^3 = 2Hf'f^2 - 2H'f^3 - (2g+\delta)Hf^3,$

$Q = -(A + (2g+\delta)H'f^3 - (2g+\delta)Hf'f^2 + g(g+\delta)Hf^3)$

$\quad = -(H''f^3 - Hf''f^2 - 2H'f'f^2 + 2H(f')^2f + (2g+\delta)H'f^3 - (2g+\delta)Hf'f^2 + g(g+\delta)Hf^3),$

$L = -(H'f^4 + (g+\delta)Hf^4 - ag\delta f^4).$

Proof. Rewrite equation(2.1.3) as $f'(t) + (\nu+g)f(t) = aE(t)$. Differentiate both sides: $f''(t) + (\nu+g)f'(t) = aE'(t)$. Plugging (2.1.2) and (2.1.3) into the above equation, we get

$$
\begin{aligned}
f''(t)+(\nu+g)f'(t) &= aE'(t) \\
&\overset{(2.1.2)}{=} a(\beta(t)S(t)f(t)-(a+g)E(t)) \\
&= a\beta(t)S(t)f(t)-(a+g)(aE(t)) \\
&\overset{(2.1.3)}{=} a\beta(t)S(t)f(t)-(a+g)(f'(t)+(\nu+g)f(t).)
\end{aligned}
$$

Rewrite the above equation as

$$
aS(t)=\frac{f''(t)+(\nu+2g+a)f'(t)+(a+g)(\nu+g)f(t)}{\beta(t)f(t)}.
$$

Denote $H(t)\triangleq f''(t)+(\nu+2g+a)f'(t)+(a+g)(\nu+g)f(t)$, then $aS(t)=\frac{H(t)}{\beta(t)f(t)}$.
Take first and second derivative of the above equation, we get

$$
\begin{aligned}
aS'(t) =&(\frac{H(t)}{\beta(t)f(t)})'=\frac{H'(t)\beta(t)f(t)-H(t)(\beta'(t)f(t)+\beta(t)f'(t)}{\beta(t)^2f(t)^2}, \\
aS''(t) =&(\frac{H'(t)\beta(t)f(t)-H(t)(\beta'(t)f(t)+\beta(t)f'(t)}{\beta(t)^2f(t)^2})' \\
=&\frac{[H'\beta f-H(\beta' f+\beta f')]'\beta^2f^2-[H'\beta f-H(\beta' f+\beta f')](\beta^2f^2)'}{(\beta^2f^2)^2} \\
=&\frac{H''f^3\beta^3+H'f^3\beta'\beta^2+H'f'f^2\beta^3-H'f^3\beta'\beta^2-H'f'f^2\beta^3-Hf^3\beta''\beta^2-2Hf'f^2\beta'\beta^2-Hf''f^2\beta^3}{\beta^4f^4} \\
&-\frac{2(H'f^3\beta'\beta^2+H'f'f^2\beta^3-Hf^3(\beta')^2\beta-Hf'f^2\beta'\beta^2-Hf'f^2\beta'\beta^2-H(f')^2f\beta^3)}{\beta^4f^4} \\
=&\frac{[H''f^3-Hf''f^2-2H'f'f^2+2H(f')^2f]\beta^3+[2Hf'f^2-2H'f^3]\beta'\beta^2+2Hf^3(\beta')^2\beta-Hf^3\beta''\beta^2}{\beta^4f^4} \\
=&\frac{A\beta^3+B\beta'\beta^2+C(\beta')^2\beta+D\beta''\beta^2}{\beta^4f^4},
\end{aligned}
$$

where $A=H''f^3-Hf''f^2-2H'f'f^2+2H(f')^2f$,

$B=2Hf'f^2-2H'f^3$, $C=2Hf^3$, $D=-Hf^3$.

Since $S+E+I+R+A=1$, i.e. $S+E+I+R=1-A$,

(2.1.5) becomes $A'=g(S+E+I+R)-\delta A=g(1-A)-\delta A=g-(g+\delta)A$

Take derivative of (2.1.1) respect to t: $S'' + (\beta SI)' + gS' = \delta A'$, plug in (2.1.1) and (2.1.5):

$$S'' + (\beta SI)' + gS' = \delta A' = \delta(g - (g+\delta)A) = g\delta - (g+\delta)(\delta A)$$
$$\overset{(2.1.1)}{=} g\delta - (g+\delta)(S' + \beta SI + gS) \qquad \text{Multiply } a \text{ to both sides of}$$

the above equation and simplify:

$$aS'' + (2g+\delta)(aS') + g(g+\delta)(aS) + (a\beta SI)' + (g+\delta)(a\beta SI) = ag\delta$$

Plug in the formula of aS, aS' and aS'':

$$(\frac{A\beta^3 + B\beta'\beta^2 + C(\beta')^2\beta + D\beta''\beta^2}{\beta^4 f^4}) + (2g+\delta)(\frac{H'(t)\beta(t)f(t) - H(t)(\beta'(t)f(t) + \beta(t)f'(t)}{\beta(t)^2 f(t)^2})$$
$$+ g(g+\delta)(\frac{H(t)}{\beta(t)f(t)}) + H' + (g+\delta)H = ag\delta.$$

Multiply by $\beta^4 f^4$ and expand:

$$A\beta^3 + B\beta'\beta^2 + C(\beta')^2\beta + D\beta''\beta^2 + (2g+\delta)H'f^3\beta^3 - (2g+\delta)Hf^3\beta'\beta^2$$
$$- (2g+\delta)Hf'f^2\beta^3 + g(g+\delta)Hf^3\beta^3 + H'f^4\beta^4 + (g+\delta)Hf^4\beta^4 = ag\delta\beta^4.$$

Collect terms with the same common factors of β^4, β^3, $\beta^2\beta'$, $\beta(\beta')^2$, $\beta^2\beta''$:

$$(H'f^4 + (g+\delta)Hf^4 - ag\delta f^4)\beta^4 + (A + (2g+\delta)H'f^3 - (2g+\delta)Hf'f^2 + g(g+\delta)Hf^3)\beta^3$$
$$+ (B - (2g+\delta)Hf^3)\beta'\beta^2 + C(\beta')^2\beta + D\beta''\beta^2 = 0.$$

i.e. $$-L\beta^4 - Q\beta^3 + P\beta'\beta^2 + N(\beta')^2\beta + M\beta''\beta^2 = 0, \qquad (2.4.2)$$

where

$$M = D = -Hf^3,$$

$$N = C = 2Hf^3,$$

$$P = B - (2g+\delta)Hf^3 = 2Hf'f^2 - 2H'f^3 - (2g+\delta)Hf^3,$$

$$\begin{aligned} Q &= -(A + (2g+\delta)H'f^3 - (2g+\delta)Hf'f^2 + g(g+\delta)Hf^3) \\ &= -(H''f^3 - Hf''f^2 - 2H'f'f^2 + 2H(f')^2 f + (2g+\delta)H'f^3 - (2g+\delta)Hf'f^2 + g(g+\delta)Hf^3), \end{aligned}$$

$$L = -(H'f^4 + (g+\delta)Hf^4 - ag\delta f^4).$$

Notice that $N = -2M$ and rewrite equation(2.4.2) as

$$M\beta''\beta^2 - 2M(\beta')^2\beta + P\beta'\beta^2 - Q\beta^3 - L\beta^4 = 0.$$

Divided by β^4:

$$\begin{aligned} & M\frac{\beta''\beta^2 - 2(\beta')^2\beta}{\beta^4} + P\frac{\beta'\beta^2}{\beta^4} - Q\frac{\beta^3}{\beta^4} - L\frac{\beta^4}{\beta^4} \\ =& - M(2\beta^{-3}(\beta')^2 - \beta^{-2}\beta'') - P(-\beta^{-2}\beta') - Q(\beta^{-1}) - L \\ =& 0. \end{aligned}$$

Let $y = \beta^{-1}$, $y' = -\beta^{-2}\beta'$, $y'' = 2\beta^{-3}(\beta')^2 - \beta^{-2}\beta''$.
(2.4.2) can be rewrite as $My'' + Py' + Qy + L = 0$ with $y = \beta^{-1}$.

□

According to theorem 4, we obtain the prevalence algorithm to extract the transmission rate from the $SEIRA$ model.

Step 1 Smoothly interpolate the infection data with a spline or trigonometric function to generate a smooth f(t). Check condition : $f'(t)/f(t) > -(\nu + g)$, where ν is the removal rate and g is the growth rate .

Step 2 Calculate the function $H(t) = f''(t) + (\nu + 2g + a)f'(t) + (a+g)(\nu+g)f(t)$. Calculate M, N, P, Q, and L by plugging $H(t)$ into formulas (2.4.1).

Step 3 Choose $\beta(0)$, $\beta'(0)$, and interval $[0,T]$, use an ODE solver to solve equation $M\beta''\beta^2 + N(\beta')^2\beta + P\beta'\beta^2 - L\beta^4 - Q\beta^3 = 0$ for $\beta(t)$ on interval $[0,T]$.

2.4.2 The incidence algorithm

To extend the incidence algorithm to the $SEIRA$ model, we first rewrite it in terms of incidence $\omega(t)$. With $\omega(t) = \beta SI$ and $S + E + I + R + A = 1$, the $SEIRA$ model can be rewritten as

$$\frac{dS}{dt} = (\delta A(t) - \omega(t)) - gS(t) \tag{2.4.3}$$

$$\frac{dE}{dt} = \omega(t) - (a+g)E(t), \tag{2.4.4}$$

$$\frac{dI}{dt} = aE(t) - (\nu+g)I(t), \tag{2.4.5}$$

$$\frac{dR}{dt} = \nu I(t) - gR(t), \tag{2.4.6}$$

$$\frac{dA}{dt} = g - (g+\delta)A(t). \tag{2.4.7}$$

We can extend the incidence algorithm to the $SEIRA$ model following the theorem below:

Theorem 5: For the $SEIRA$ model, the time-dependent transmission function is

$$\beta(t) = \frac{\omega(t)}{S(t)I(t)},$$

where

$$S(t) = S_0 e^{-gt} + \int_0^t (\delta(A_0 e^{-(g+\delta)s} + \int_0^s g e^{(g+\delta)(\sigma-s)} d\sigma) - \omega(s)) e^{g(s-t)} ds, \quad (2.4.8)$$

$$I(t) = I_0 e^{-(\nu+g)t} + \int_0^t a(E_0 e^{-(a+g)s} + \int_0^s \omega(\sigma) e^{(a+g)(\sigma-s)} d\sigma) e^{(\nu+g)(s-t)} ds. \quad (2.4.9)$$

Proof. First solving (2.4.7) with initial value $A(0) = A_0$ by using method of variation of constant. Assume that

$$\begin{aligned}
&\frac{dA}{dt} = -(g+\delta)A \\
\Rightarrow &\frac{dA}{A} = -(g+\delta)dt \\
\Rightarrow &ln(A) = -(g+\delta)t + C_1 \\
\Rightarrow &A = e^{-(g+\delta)t+C_1} = e^{-(g+\delta)t}C \quad (C = e^{C_1})
\end{aligned}$$

Then we assume that the solution of (2.4.7) is

$$\begin{aligned}
&A = C(t)e^{-(g+\delta)t} \\
\Rightarrow &A' = C'(t)e^{-(g+\delta)t} + C(t)(-(g+\delta)e^{-(g+\delta)t})
\end{aligned}$$

Plug it into (2.4.7):

$$\begin{aligned}
A' &= C'(t)e^{-(g+\delta)t} + C(t)(-(g+\delta)e^{-(g+\delta)t}) \\
&= g - (g+\delta)C(t)e^{-(g+\delta)t} \\
\Rightarrow &C'(t)e^{-(g+\delta)t} = g \\
\Rightarrow &C'(t) = ge^{(g+\delta)t} \\
\Rightarrow &C(t) = C(0) + \int_0^t ge^{(g+\delta)s} ds
\end{aligned}$$

When t=0, $A_0 = A(0) = A(t)|_{t=0} = C(t)e^{-(g+\delta)t}|_{t=0}$
$= (C(0) + \int_0^t ge^{(g+\delta)s})e^{-(g+\delta)t}|_{t=0}$
$= C(0),$

the solution of (2.4.7) with initial value $A(0) = A_0$ is

$$A(t) = C(t)e^{-(g+\delta)t} = (A(0) + \int_0^t ge^{(g+\delta)s}ds)e^{-(g+\delta)t}$$

$$\Rightarrow A(t) = A_0e^{-(g+\delta)t} + \int_0^t ge^{(g+\delta)(s-t)}ds. \tag{2.4.10}$$

By using the same method, we can get the solutions for $S(t)$, $I(t)$ and $E(t)$ with initial values $S(0) = S_0$, $I(0) = I_0$ and $E(0) = E_0$:

$$S(t) = S_0e^{-gt} + \int_0^t (\delta A(s) - \omega(s))e^{g(s-t)}ds \tag{2.4.11}$$

$$I(t) = I_0e^{-(\nu+g)t} + \int_0^t aE(s)e^{(\nu+g)(s-t)}ds \tag{2.4.12}$$

$$E(t) = E_0e^{-(a+g)t} + \int_0^t \omega(s)e^{(a+g)(s-t)}ds \tag{2.4.13}$$

Plug (2.4.13) into (2.4.12) and plug (2.4.10) into (2.4.11):

$$S(t) = S_0e^{-gt} + \int_0^t (\delta(A_0e^{-(g+\delta)s} + \int_0^s ge^{(g+\delta)(\sigma-s)}d\sigma) - \omega(s))e^{g(s-t)}ds \tag{2.4.14}$$

$$I(t) = I_0e^{-(\nu+g)t} + \int_0^t a(E_0e^{-(a+g)s} + \int_0^s \omega(\sigma)e^{(a+g)(\sigma-s)}d\sigma)e^{(\nu+g)(s-t)}ds \tag{2.4.15}$$

Thus $\beta(t) = \frac{\omega(t)}{S(t)I(t)}$ with $S(t)$ and $I(t)$ in (2.4.14) and (2.4.15) □

Now we turn the above theorem into an algorithm to extract time-dependent transmission rate $\beta(t)$ using incidence data:

Step 1 Smoothly interpolate incidence data of infectious with a spline or trigonometric function to generate a smooth $\omega(t)$ (In fact, we only need $\omega(t)$ to be continuous, not necessary smooth). Assume that the time period for $\omega(t)$ is $[0, T]$

Step 2 Divide $[0, T]$ into n small intervals with middle points and end points

$t_i = \frac{T}{n}i, \quad 0 \le i \le n$. Calculate $\beta(t_i) = \frac{\omega(t_i)}{S(t_i)I(t_i)}$ with

$$S(t_i) = S_0 e^{-gt_i} + \int_0^{t_i} (\delta(A_0 e^{-(g+\delta)s} + \int_0^s g e^{(g+\delta)(\sigma-s)} d\sigma) - \omega(s)) e^{g(s-t_i)} ds,$$
$$I(t_i) = I_0 e^{-(\nu+g)t_i} + \int_0^{t_i} a(E_0 e^{-(a+g)s} + \int_0^s \omega(\sigma) e^{(a+g)(\sigma-s)} d\sigma) e^{(\nu+g)(s-t_i)} ds.$$

Step 3 Repeat *step 2* for $0 \le i \le n$.

2.4.3 Numerical simulations

To test our two algorithms derived in the previous sections, we do simulations with both fake and real measles data from England&Wales. Figure 2.3 plots $\beta(t)$ extracted from fake prevalence data $f(t) = 10^{-3}[1.4 + cos(2\pi t/12)]$. From (a) we can see that it works in a short time. However, as shown in (b), it will have singularity finally. Although we did not find an explicit formula for $\beta(t)$, we believe that singularity is caused by a zero denominator, which is the reason in Pollicott et al. [19]. Our formula needs higher order of derivatives which makes it more unstable and has a higher possibility of singularity.

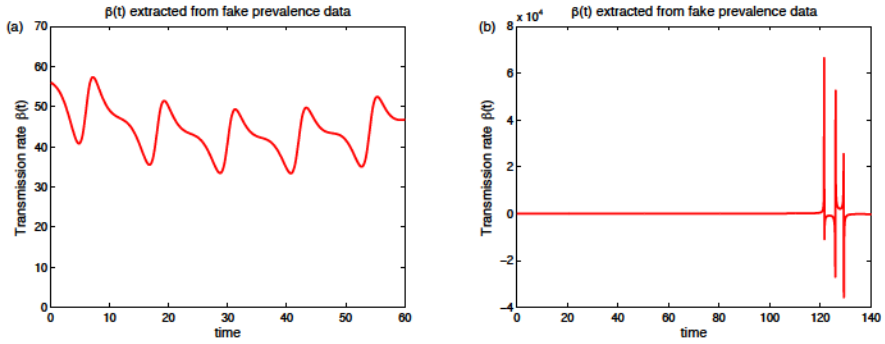


Figure 2.3: The transmission rate $\beta(t)$ extracted from fake prevalence data $f(t) = 10^{-3}[1.4 + cos(2\pi t/12)]$ with initial value $\beta(0) = 56, \beta'(0) = -1$, and parameters $\nu = 52/12, a = 52/12, g = 1/16/12, \delta = 1/64/12$. (a) $\beta(t)$ from time 0 to 60.(b) $\beta(t)$ from time 0 to 140.

Figure 2.4 plots $\beta(t)$ from simulated incidence data set. Since $\beta(t) = \frac{\omega(t)}{S(t)I(t)}$, the algorithm always works except when $S(t)$ or $I(t)$ becomes zero. Equation (2.4.2) indicates that $I(t)$ is always positive. In theory, $S(t)$ might be zero or negative when $\omega(t)$ is too large so that all susceptibles are infected. While in reality, it is rarely possible, thus the incidence algorithm will always work for real data. As long as we pick reasonable simulated notification data, it will also work for a long time.

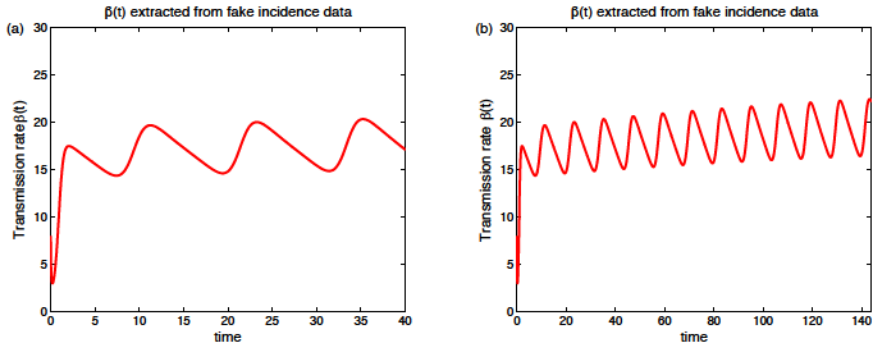


Figure 2.4: $\beta(t)$ extracted from fake incidence data $f(t) = 10^{-4}[2.7+1.5sin(2\pi t/12)]$ with initial value $S_0 = 0.25, E_0 = 0.0009, I_0 = 0.0001, A_0 = 0.7$, and parameters $\nu = 52/12, a = 52/12, g = 1/16/12, \delta = 1/64/12$. (a) $\beta(t)$ from time 0 to 40.(b) $\beta(t)$ from time 0 to 144.

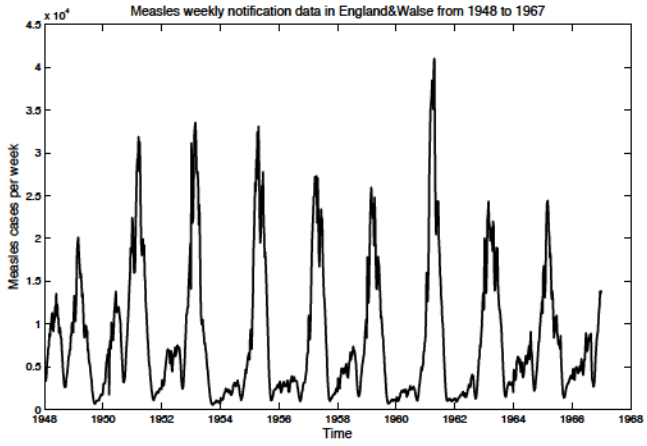


Figure 2.5: Measles weekly notification data in England&Walse from 1948-67.

Figure 2.5 is the weekly notification data of measles in England&Walse. We

convert it to be prevalence data and apply it with the prevalence algorithm. Figure 2.6 (a) plots the part of $\beta(t)$ before singularity happens and (b) plots the whole $\beta(t)$, with singularity. The prevalence algorithm can only work for a short time. The more disappointing result is that the transmission rate in figure 2.6 (a) is not a reasonable estimation for $\beta(t)$. We know that both the data set and interpolation can have errors. We believe that errors accumulated from complex calculation are enough to disturb information of transformation rate.

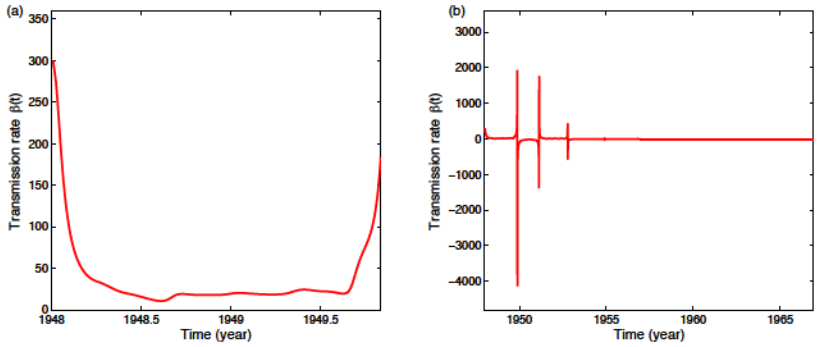


Figure 2.6: The transmission rate $\beta(t)$ extracted from measles data of England&Walse with initial value $\beta(0) = 25/month, \beta'(0) = -10/month$, and parameters $\nu = 52/12, a = 52/12, g = 1/16/12, \delta = 1/64/12$. (a) $\beta(t)$ from time year 1948 to 1949. (b) $\beta(t)$ from year 1948 to 1966.

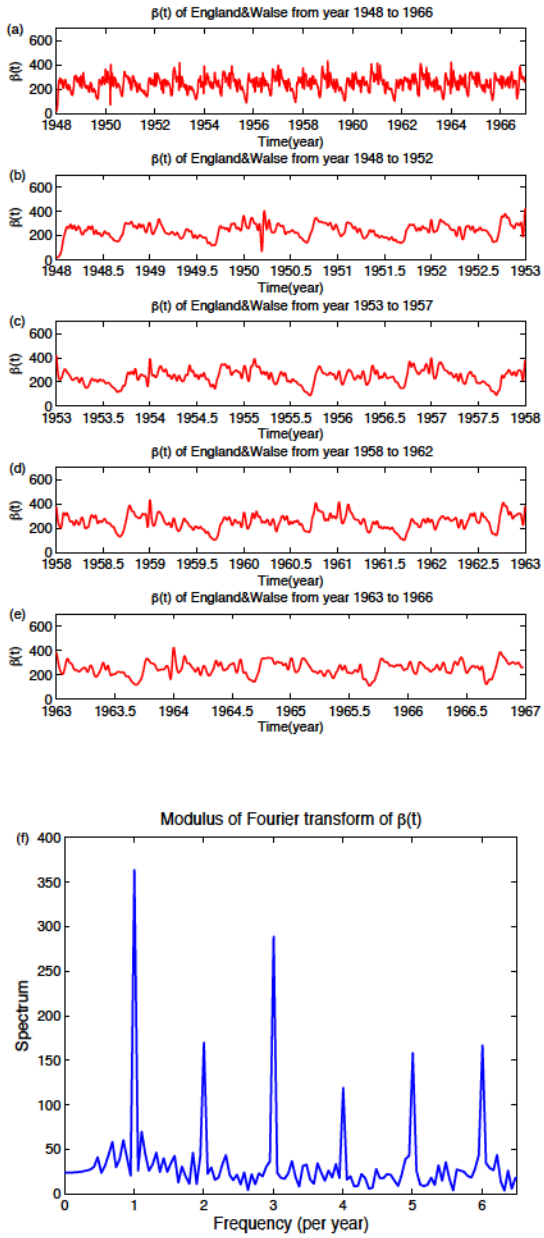


Figure 2.7: (a) Time-dependent transmission rate $\beta(t)$ of England&Walse from year 1948 to 1966. Figure (b), (c), (d) and (e) are time-dependent transmission rate every three years from 1948 to 1952, 1953 to 1957, 1958 to 1962 and 1963 to 1966, respectively. $\beta(t)$ is extracted with parameters $\delta = 1/64/52/week, a = 52/52/week, \nu = 52/52/week, g = 1/16/52/week$ and initial values $S_0 = 0.2,\ E_0 = 0.001,\ I_0 = 0.001,\ A_0 = 0.78$. (f) plots the modulus of Fourier transform of $\beta(t)$ in (a).

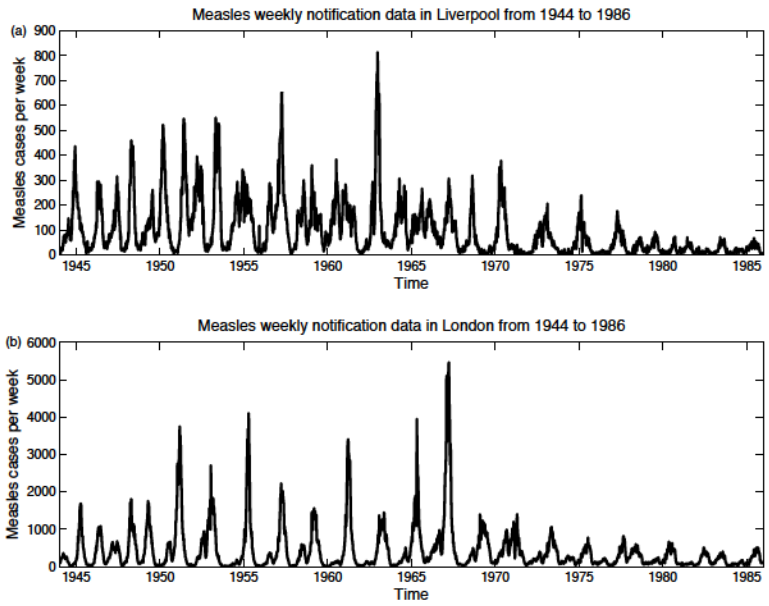


Figure 2.8: Measles weekly notification data in Liverpool and London from 1944-86.

We then apply the incidence algorithm with real weekly measles incidence data set from Liverpool and London. Figure 2.8 presents weekly notification data of those two cities. London is a larger city and has a much higher level of patients.

Discrete Fourier transform (DFT) can convert the domain of equally spaced discrete data to frequency domain. The panel(e) plots the modulus of Fourier transform of corresponding $\beta(t)$ in Figure 2.9 (a). We can see two peaks with frequencies 1/year and 3/year which is consistent with common belief that measles is driven seasonally as well as by school vacations. Also, the dominant frequencies are robust with respect to initial values.

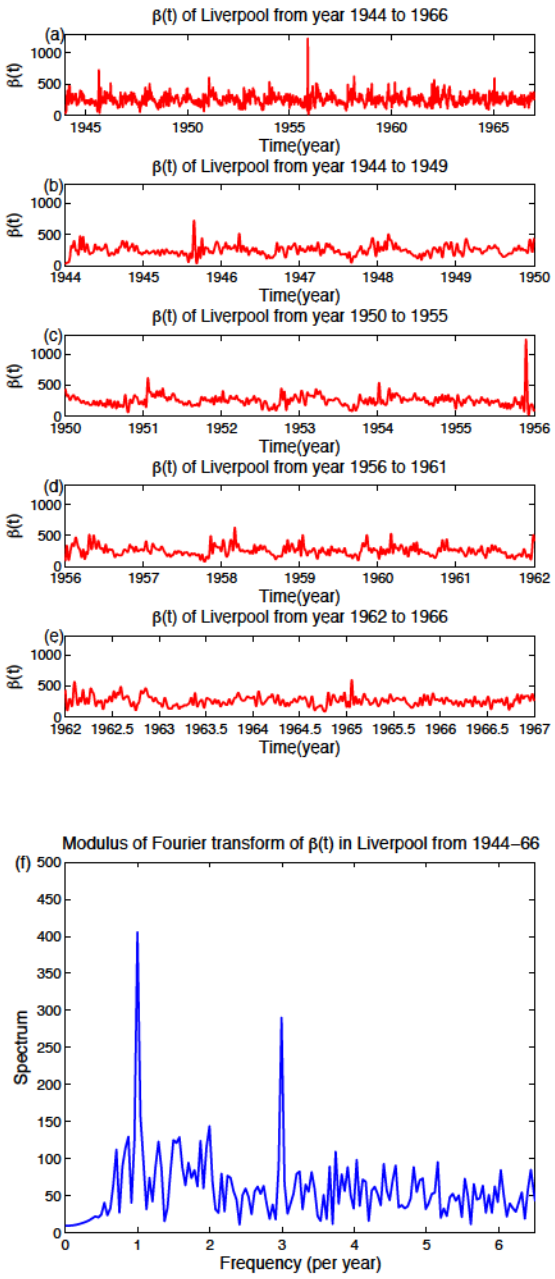


Figure 2.9: (a) Time-dependent transmission rate $\beta(t)$ of Liverpool from year 1944 to 1966. (b), (c), (d) and (e) are time-dependent transmission rate every three years from 1948 to 1952, 1953 to 1957, 1958 to 1962 and 1963 to 1966, respectively. $\beta(t)$ is extracted with parameters $\delta = 1/64/52/week, a = 52/52/week, \nu = 52/52/week, g = 1/16/52/week$ and initial values $S_0 = 0.2, E_0 = 0.001, I_0 = 0.001, A_0 = 0.78$. (f) plots the modulus of Fourier transform of $\beta(t)$ in (a).

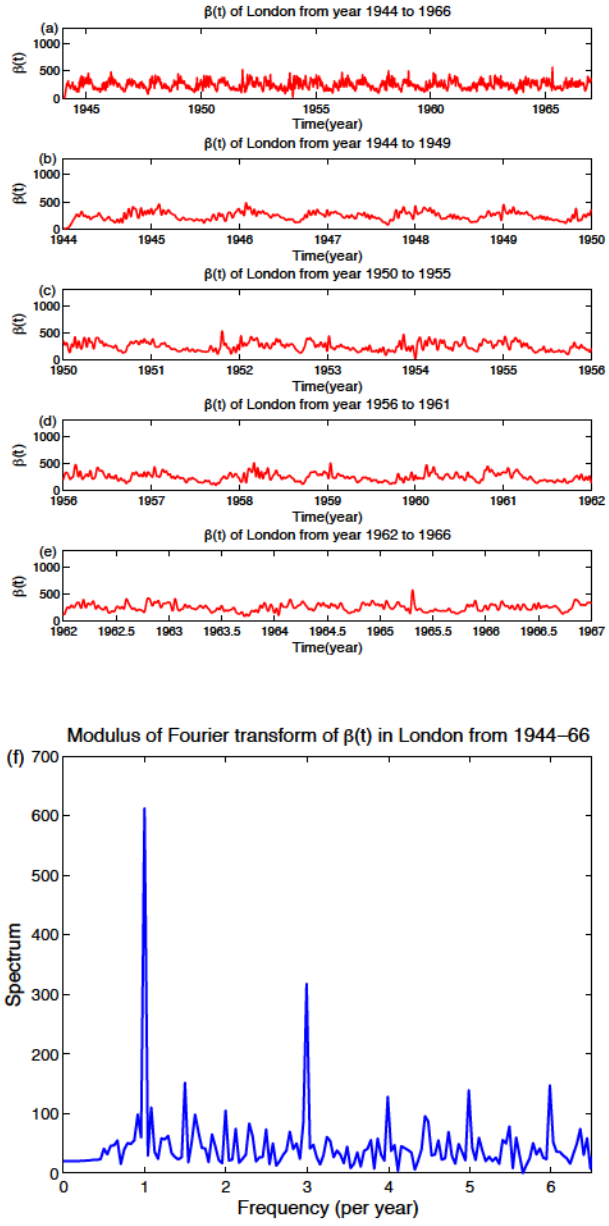


Figure 2.10: (a) Time-dependent transmission rate $\beta(t)$ of London from year 1944 to 1966. (b),(c),(d) and (e) are time-dependent transmission rate every three years from 1948 to 1952, 1953 to 1957, 1958 to 1962 and 1963 to 1966, respectively. $\beta(t)$ is extracted with parameters $\delta = 1/64/52/week, a = 52/52/week, \nu = 52/52/week, g = 1/16/52/week$ and initial values $S_0 = 0.2, E_0 = 0.001, I_0 = 0.001, A_0 = 0.78$. (f) plots the modulus of Fourier transform of $\beta(t)$ in (a).

The incidence algorithm gives the same 1 and 3 per year dominant frequencies as what Pollicott et.al [19] concluded, but we did not find the frequency of every two years which is found in Pollicott et al. [19]

After simulations with data from Liverpool, London and England&Walse, we conclude that

- Measles transmission is driven seasonally and by school holidays.
- The transmission rate is synchronized in different cities.
- Seasonality is more important than school holidays in affecting the transmission rate, especially in large cities.

From the mathematical calculations and simulation results in this chapter, we agree with Hadeler [9] that the incidence algorithm has advantages over the prevalence in many aspects.

First, it is better to use incidence rather than prevalence data. The notification data is the new cases over a unit time. It is exactly what we need in incidence algorithm. For the prevalence algorithm, the prevalence data is converted from notification data. In this thesis, the prevalence data is converted from weekly notification data with the method we mentioned in the first chapter.

Second, the incidence algorithm is more stable. The incidence algorithm will always work (no singularity) as long as $I(t)$ and $S(t)$ are not zero since the formula for β is $\beta(t) = \frac{\omega(t)}{S(t)I(t)}$. However, the prevalence algorithm for the $SEIRA$ model involves until the fourth order derivatives of the smooth function from the interpolation of the prevalence data. It works for a short time and singularity will happen finally.

Third, incidence algorithm has less challenge in dealing with the problem of estimating the initial values. It is difficult to evaluate the initial value of $\beta(t)$, $\beta(t)$. Both of them have great affects on the dynamics of $\beta(t)$. While the only challenge

for the incidence algorithm is to estimate $S(0)$ which can be estimated based on the vaccination policy in the past decade since initial values of the other groups have weak influence on $\beta(t)$.

2.5 Conclusion

In this chapter, we modify the $SEIR$ model to be the $SEIRA$ model to better describe the dynamics of childhood infectious diseases. We analyze the boundedness, positivity and stability of this model. It has positivity and is bounded by $S + E + I + R + A = 1$. Thus, the solution will always be positive if starts with a positive initial value. Also, there are two equilibria, one disease-free equilibrium and one endemic equilibrium. When the basic reproduction number $R_0 < 1$, the disease-free equilibrium is locally asymptotically stable and the endemic equilibrium is not feasible, i.e. the steady state is negative. When $R_0 > 1$, the disease-free equilibrium is unstable and the endemic equilibrium is locally asymptotically stable.

From sensitivity analysis, we know that transmission is the most important in preventing and ameliorating the magnitude of an outbreak while birth and removal rate are most important in controlling the endemic level.

We extend both the prevalence algorithm and the incidence algorithm for the $SEIRA$ model. We verify that the prevalence algorithm is more unstable than the incidence data. From Fourier transform of the transmission rate of Liverpool, London, England&Walse extracted from the incidence data, we conclude that the transmission of measles is driven by seasonality and school holidays, whereas seasonality is more dominant, especially in large cities. Also, cycles of the transmission rate is synchronized in different cities, i.e. frequencies of 1 and 3 per year are the dominant frequencies in different cities.

Chapter 3

The $SEIRA$ model with vaccination

Vaccination is the process that a vaccine stimulates the immune system of an individual to build immunity against a pathogen. Vaccination can ameliorate both mortality and morbidity. The effectiveness of vaccination has been widely studied and verified since the first work of Edward Jenner on smallpox [16].

Different vaccination strategies are used to deal with different situations. Pediatric vaccination is an efficient way in preventing dangerous human infectious diseases. Much work has been focused on the vaccination of newborn babies or infants to reduce the prevalence of diseases like measles, mumps,rubella,etc. Mathematical treatment of vaccination is straight forward and only needs a single addition to the $SEIRA$ model. Note that p is used to denote the fraction of the newborns who are successfully vaccinated. We obtain the following $SEIRA$ model with vaccination:

$$\frac{dS}{dt} = \delta(1-p)A(t) - \beta S(t)I(t) - gS(t), \tag{3.0.1}$$
$$\frac{dE}{dt} = \beta S(t)I(t) - aE(t) - gE(t), \tag{3.0.2}$$
$$\frac{dI}{dt} = aE(t) - \nu I(t) - gI(t), \tag{3.0.3}$$
$$\frac{dR}{dt} = \nu I(t) - gR(t) + \delta pA(t), \tag{3.0.4}$$
$$\frac{dA}{dt} = g(S(t) + E(t) + I(t) + R(t)) - \delta A(t). \tag{3.0.5}$$

However, it is not cost-effective to control rare infectious diseases by pediatric vaccination. Therefore, another vaccination policy, random vaccination are conducted for rare infectious diseases or an potential outbreak which target to vaccinate all unvaccinated individuals, not only the newborns.

It is difficult for a disease to spread as long as the fraction of susceptibels is kept in a low level. Therefore, it is more reasonable that we should vaccinate less if the fraction of susceptibles is lower, and vice-versa. The vaccination policy with vaccination quantity depends on the level of susceptibels is called wildlife vaccination. It was first used to prevent the spread of animal infectious disease since animals are easier to control. But we believe that this model is also reasonable with human populations nowadays, especially in developed countries and cities with high education rate.

3.1 Qualitative analysis

In this section, we analyze positivity, boundedness and stability of the equation system 3. Using the same symbols to analyze positivity, $f_1 = \delta(1-p)y_5 \geq 0,\ f_2 = \beta y_1 y_3 \geq 0,\ f_3 = ay_2 \geq 0,\ f_4 = \nu y_3 + \delta p y_5 \geq 0,\ f_5 = g(y_1+y_2+y_3+y_4) \geq 0$. Therefore

system (3) has positivity. Also, boundedness property is the same as it in chapter 2. Combining positivity and boundedness, we conclude that the solution of (3) will stay in $\{(S, E, I, R, A)$: $S \geq 0, E \geq 0, I \geq 0, R \geq 0, A \geq 0, S + E + I + R + A = 1\}$ if starts at a positive initial value.

The same with the situation in the previous chapter, there are two equilibria for the vaccinated $SEIRA$ model:
the disease-free equilibrium

$$(S_1^*,\ E_1^*,\ I_1^*,\ R_1^*,\ A_1^*) = (\frac{\delta(1-p)}{g+\delta},\ 0,\ 0,\ \frac{\delta p}{g+\delta},\ \frac{g}{g+\delta})$$

and the endemic equilibrium

$$(S_2^*,\ E_2^*,\ I_2^*,\ R_2^*,\ A_2^*) = (\frac{(a+g)(\nu+g)}{a\beta},\ \frac{g\delta(1-p)}{(a+g)(g+\delta)} - \frac{g(\nu+g)}{a\beta},\ \frac{ag\delta(1-p)}{(a+g)(g+\delta)(\nu+g)} - \frac{g}{\beta},$$

$$\frac{a\nu\delta(1-p)}{(a+g)(g+\delta)(\nu+g)} - \frac{\nu}{\beta} + \frac{\delta p}{g+\delta},\ \frac{g}{g+\delta})$$

The calculation of equilibria is almost the same, we will not present the details again. The local stability of these two equilibria are shown below:

Assume that all parameters are positive.

- When $a\delta(1-p)\beta < (a+g)(g+\delta)(\nu+g)$, the disease-free equilibrium is locally asymptotically stable and the endemic equilibrium is not feasible ,
- When $a\delta(1-p)\beta > (a+g)(g+\delta)(\nu+g)$, the endemic equilibrium is locally asymptotically stable and the disease-free equilibrium is unstable .

The basic reproduction ratio is $R_0 = \frac{a\beta\delta(1-p)}{(a+g)(g+\delta)(\nu+g)}$. It can be rewritten as $R_0 = \frac{\beta}{\nu+g} \cdot \frac{a}{a+g} \cdot \frac{\delta}{a+g} \cdot (1-p)$. The first three terms have the same meanings as in chapter 2. When considering vaccination, the level of susceptibles will be *'discounted'* by $p \times 100\%$ because newborns who have vaccination are no longer considered as

susceptible.

3.2 Sensitivity analysis

Sensitivity indices have been discussed in the previous chapter. In this chapter, we focus on the sensitivity indices respect to the vaccination rate p who has weak relationships with both the outbreak peak value and time. But p has a strong relationship with the endemic level of the infected since p directly determines the influx of the system.

3.2.1 Sensitivity analysis of the outbreak peak value

Parameter	Sensitivity of peak	
p	-0.0137	Vaccination Rate
δ	0.0137	Natural death/birth Rate
β	0.8629	Transmission Rate
g	-0.0152	Growth Rate
ν	-1.3635	Removal Rate
a	0.5124	Rate at which exposed individuals become infective

.

Table 3.1: Sensitivity of the value of the outbreak peak to the parameters with the parameter values $p = 0.5, \delta = 1/64/12/month$, $\beta = 150/month$, $g = 1/16/12/month$, $\nu = 52/12/month$, $a = 52/12/month$ and initial values $S_0 = 0.0998$, $E_0 = 0.0001$, $I = 0.0001$, $R_0 = 0.11$, $A_0 = 0.79$.

Comparing sensitivity analysis with it in the second chapter, we can see that the absolute value of sensitivity indices respect to all parameters are smaller, but vaccination does not change ranks of their importance. The vaccination rate p has a negative relationship with the outbreak peak since more pediatric vaccination will result in less infectives.

3.2.2 Sensitivity analysis of the outbreak peak time

Parameter	Sensitivity of the peak time	
p	0.0056	Vaccination Rate
δ	-0.0056	Natural death/birth Rate
β	-1.1359	Transmission Rate
g	0.0071	Growth Rate
ν	0.5821	Removal Rate
a	-0.4484	Rate at which exposed individuals become infective

Table 3.2: Sensitivity of the outbreak peak time to the parameters with parameter values $p = 0.5, \delta = 1/64/12/month,\ \beta = 150/month,\ g = 1/16/12/month,\ \nu = 52/12/month,\ a = 52/12/month$ and initial values $S_0 = 0.0998,\ E_0 = 0.0001,\ I = 0.0001,\ R_0 = 0.11,\ A_0 = 0.79$.

From table 3.2, we can see that the vaccination rate p is the least important to control in preventing outbreaks. The rank of sensitivity indices with respect to different parameters is not changed.

3.2.3 Sensitivity analysis of the endemic steady state

Parameter	Sensitivity of I_2^*	Description
p	-1.4076	Vaccination Rate
δ	1.1261	Natural death/birth Rate
β	0.4077	Transmission Rate
g	-0.1295	Growth Rate
ν	-1.4061	Removal Rate
a	0.0017	Rate at which exposed individuals become infective

Table 3.3: Sensitivity of the endemic steady state to the parameters with parameter values $p = 0.5,\ \delta = 1/64/12/month,\ \beta = 150/month,\ g = 1/16/12/month,\ \nu = 52/12/month,\ a = 52/12/month$.

Table 3.3 shows that the vaccination rate p has the greatest importance in determining the endemic level of infectives. The negative relationship is because

more vaccination will reduce the fraction of susceptibles. Infectives will be less with less susceptibles.

3.2.4 Numerical simulation

Sensitivity and stability properties of the vaccinated $SEIRA$ model were discussed previously and numerical simulation is presented in the following diagrams. When considering a specific disease, parameters $\delta,\ g,\ \nu,\ a$ are constant or vary in a small range while $p(t)$ and $\beta(t)$ is time and locally dependent. The vaccination rate $p(t)$ is mainly determined by the government policy and the transmission rate is determined by many factors like social structure, levels of susceptibles and infectives. Thus, in the following figures, we present the outbreak value, the outbreak time and the steady state of $I(t)$ with $p(t)$ varies from 0 to 1 and $\beta(t)$ varies from 0 to 100 per month.

Figure 3.1 plots the peak value of $I(t)$ with varying $p(t)$ and $\beta(t)$. We can see that the vaccination rate p has little impact on the outbreak maximum value since newborn babies during the outbreak is negligible. In fact, we can ignore demography when considering outbreaks. The transmission rate $\beta(t)$ still has an important effect on the peak value of $I(t)$.

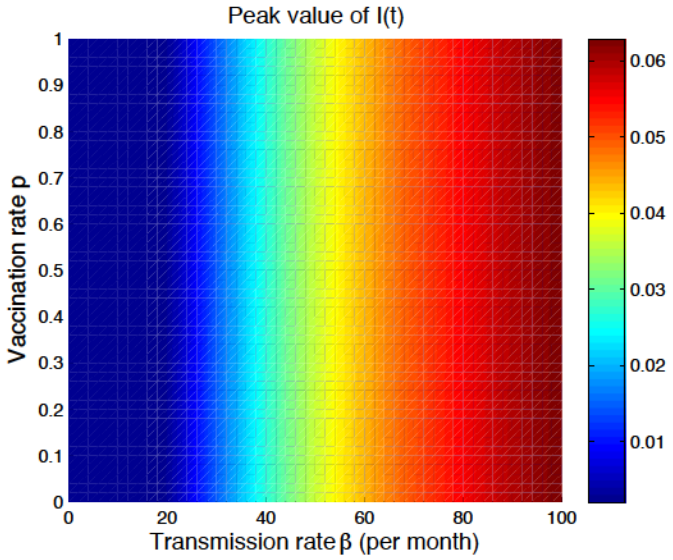


Figure 3.1: The peak value of $I(t)$ with $\beta(t)$ varying form 0 t0 100 per month and the vaccination rate $p(t)$ from 0 to 1. Parameters are $\nu = 52/12/month, \delta = 1/64/12/month$, $a = 52/12/month$, $g = 1/16/12/month$ and initial values are $S_0 = 0.25$, $E_0 = 0.002$, $I_0 = 0.002$, $R_0 = 0.046$, $A_0 = 0.7$.

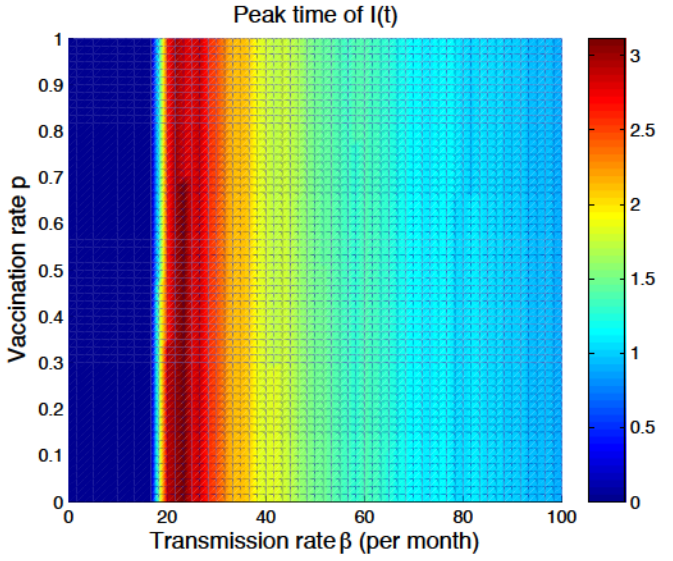


Figure 3.2: The peak time of $I(t)$ with $\beta(t)$ varying form 0 t0 100 per month and the vaccination rate $p(t)$ from 0 to 1 with parameters $\nu = 52/12/month, \delta = 1/64/12/month$, $a = 52/12/month$, $g = 1/16/12/month$ and initial values $S_0 = 0.25, E_0 = 0.002$, $I_0 = 0.002$, $R_0 = 0.046$, $A_0 = 0.7$.

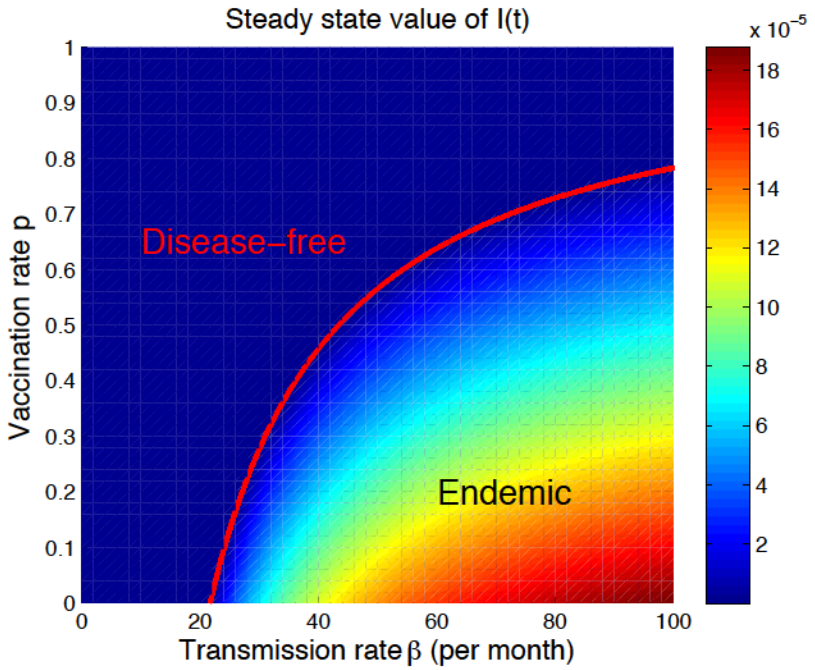


Figure 3.3: Steady state value of $I(t)$ with $\beta(t)$ varying form 0 t0 100 per month and vaccination rate $p(t)$ from 0 to 1. Parameters are $\nu = 52/12/month, \delta = 1/64/12/month,\ a = 52/12/month,\ g = 1/16/12/month.$

Figure 3.2 plots the peak time of $I(t)$ with varying $p(t)$ and $\beta(t)$. The vaccination rate p has little effect on the peak time. The relationship between β and the peak time is complicated. When β is small, $I(t)$ decreases directly and finally goes to zero. Thus peak time is zero. When β is extremely large, outbreaks come quickly. Thus, there is an optimum value for β which can postpone the outbreak most. In the figure, we can see that the optimum value is around 25 per month.

Figure 3.3 plots the endemic steady state of $I(t)$. The red line which is the threshold condition for stability $R_0 = 1$ separates the area of *'Disease-free'* and *'Endemic'* where the solution goes to the disease-free equilibrium and the endemic equilibrium. The disease-free zone is always zero and the endemic zone gradually increases with $\beta(t)$ gets larger and $p(t)$ smaller. This makes sense because a larger transmission rate $\beta(t)$ leads to more infected kids and a smaller vaccination rate

leads to more susceptible kids.

Figure 3.4 plots how sensitivity indices of steady state $I(t)$ change when the vaccination rate $p(t)$ varies. From (a) we can see how sensitivity indices of $I(t)$ change when vaccination rate varies from 0 to 0.87. In figure 3.4, interval for the vaccination rate p is not from 0 to 1 since if p is too large, the endemic equilibrium will no longer be feasible, i.e. it is meaningless to study endemic steady state. Sensitivity of $I(t)$ respect to all parameters become larger as $p(t)$ becomes larger, except for growth rate g. The curve for g stops before where p is 0.3 and continues as a dashed line which is the absolute value sensitivity indices to g. As p increases, g decreases to zero and becomes negative. Growth rate g can affects final size of infectives in two direction. In one hand, larger g will move infectives to the adult group faster thus reduce fraction of infetives. In the other hand, larger g will increase the population of adults, which will result in a larger population of infectives and susceptibles. When g is small, the first effect is dominant and when g is large, the second effect is dominating.

In (b), we zoom in the area where two curves intersects for a better view. Sensitivity respect to $\beta(t)$ increases faster and finally surpasses birth rate δ. This means that in areas where the vaccination rate $p(t)$ is low, birth rate δ and recovery rate ν have a larger influence on the steady state. While in areas where the vaccination rate $p(t)$ is high, $\beta(t)$ becomes more important than δ but still less important than ν in controlling the endemic steady state.

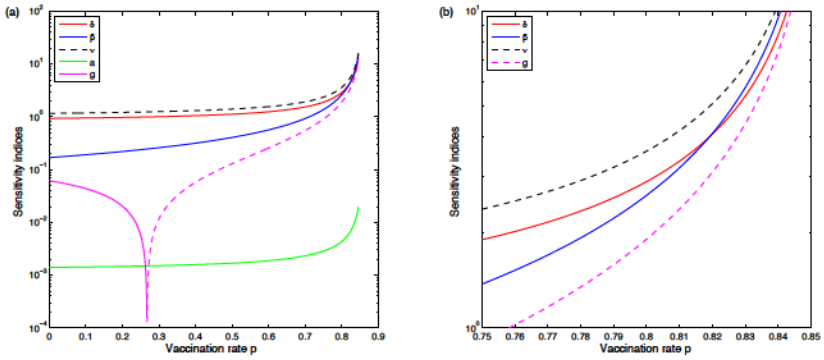


Figure 3.4: Sensitivity of different parameters with varying vaccination rate p. (a) plots figures of parameters δ, β, g, ν, and a with varying vaccination rate p. The dashed line is the S. I of ν and it is dashes means it is negative sensitive. (b) zooms in the part of (a) where S.I. of β surpasses S.I. of δ. Parameters used plotting (a) and (b) are $\delta = 1/64/12/month, \beta = 150/month, g = 1/16/12/month, \nu = 52/12/month, a = 52/12/month$.

3.3 Two algorithms to extract the transmission rate $\beta(t)$ from post-vaccination data

In this chapter, we considered vaccination and extended the $SEIRA$ model to the vaccinated $SEIRA$ model. In this section, we extend our extracting algorithm to both the prevalence and the incidence algorithm. We first states the theoretical results without repeating the details of proofs and the steps of the algorithms.

3.3.1 The prevalence algorithm

For the vaccinated $SEIRA$ model with time dependent vaccination rate $p(t)$, the time-dependent transmission function $\beta(t)$ satisfies the following differential equation :

$$M\beta''\beta^2 + N(\beta')^2\beta + P\beta'\beta^2 - L\beta^4 - Q\beta^3 = 0, \tag{3.3.1}$$

where

$$
\begin{aligned}
H &= f''(t) + (a + 2g + \nu)f'(t) + (a + g)(\nu + g)f(t), \\
M &= -(1 - p)Hf^3, \\
N &= 2(1 - p)Hf^3, \\
P &= (1 - p)(2Hf'f^2 - 2H'f^3 - (2g + \delta)Hf^3) - p'Hf^3, \\
Q &= -(1 - p)(H''f^3 - Hf''f^2 - 2H'f'f^2 + 2H(f')^2 f + (2g + \delta)H'f^3 - (2g + \delta)Hf'f^2 \\
&\quad + g(g + \delta)Hf^3) - p'H'f^3 + p'Hf'f^2 - gp'Hf^3, \\
L &= -(1 - p)(H'f^4 + (g + \delta)Hf^4 - ag\delta(1 - p)f^4) - p'Hf^4.
\end{aligned}
$$

Formulas of coefficients seem quiet different from the previous one. In fact, the formula presented above combines the situation when $0 < p \leq 1$ and when $p = 1$. From (3.0.1) we can see that when $p = 1$, $\frac{dS}{dt}$ is independent of $A(t)$ which degenerate to the case of the $SEIR$ model. We present formula for both cases when $p = 1$ and when $0 \leq p < 1$.

- when p(t)=1, the differential equation of $\beta(t)$ is

$$P\beta' - L\beta^2 - Q\beta = 0$$

This is a Bernoulli equation which is the same as the result of the $SEIR$ model. By letting $y(t) = \frac{1}{\beta(t)}$, the Bernoulli equation can be rewritten as a first order linear differential equation $Py'(t) + Qy(t) + L = 0$ with

$$
\begin{aligned}
P &= -Hf, \\
Q &= -(H'f - Hf' + gHf), \\
L &= -Hf^2.
\end{aligned}
$$

- when $0 \leq p(t) < 1$, the differential equation of the time-dependent transmission rate $\beta(t)$ is $M\beta''\beta^2 + N(\beta')^2\beta + P\beta'\beta^2 - L\beta^4 - Q\beta^3 = 0$ with

$$
\begin{aligned}
M &= -Hf^3, \\
N &= 2Hf^3, \\
P &= 2Hf'f^2 - 2H'f^3 - (\frac{p'}{1-p} + 2g + \delta)Hf^3, \\
Q &= -(H''f^3 - Hf''f^2 - 2H'f'f^2 + 2H(f')^2 f + (\frac{p'}{1-p} + 2g + \delta)H'f^3 \\
&\quad - (\frac{p'}{1-p} + 2g + \delta)Hf'f^2 + g(\frac{p'}{1-p} + g + \delta)Hf^3), \\
L &= -(H'f^4 + (\frac{p'}{1-p} + g + \delta)Hf^4 - ag\delta(1-p)f^4).
\end{aligned}
$$

  Formula (3.3.1) covers both situations. When $0 \leq p(t) < 1$, $1 - p(t) \neq 0$, all the coefficients in the theorem divided by $1 - p$ yields the same result as in the case of $0 \leq p < 1$. When $p = 1$, plugging $1 - p = 0$ into the formulas in the theorem yields $M = 0, N = 0, P = -p'Hf^3, Q = -p'(H'f - Hf' + gHf)f^2, L = -p'Hf^4$. Both situation consistent with formula (3.3.1).

3.3.2 The incidence algorithm

For the vaccinated $SEIRA$ model with time-dependent vaccination rate $p(t)$, the time-dependent transmission rate is $\beta(t) = \frac{\omega(t)}{S(t)I(t)}$ where $S(t)$ and $I(t)$ are

$$
\begin{aligned}
S(t) =& S_0 e^{-gt} + \int_0^t (\delta(\mathbf{1\text{-}p(s)}\ )(A_0 e^{-(g+\delta)s} + \int_0^s g e^{(g+\delta)(\sigma-s)} d\sigma) - \omega(s)) e^{g(s-t)} ds \\
I(t) =& I_0 e^{-(\nu+g)t} + \int_0^t a(E_0 e^{-(a+g)s} + \int_0^s \omega(\sigma) e^{(a+g)(\sigma-s)} d\sigma) e^{(\nu+g)(s-t)} ds
\end{aligned}
$$

The only difference between this formula and the incidence formula in chapter 2 is shown in bold.

3.3.3 Numerical simulations

In the precious chapter, we have presented two algorithms to extract $\beta(t)$ and their numerical simulations on fake and real pre-vaccination data. When we did simulation with post-vaccination data of London and Liverpool, the prevalence algorithm still only worked for a short time. Thus we only present $\beta(t)$ extracted from the incidence algorithm of two cities, Liverpool and London.

Figure 3.5 presents $\beta(t)$ extracted from post-vaccination measles weekly notification data of Liverpool from year 1974 to 1986 using the incidence algorithm. Liverpool is a relatively smaller city compare to London. The population of Liverpool is less then 1/10 of London. Therefore, there are more noises with the weekly data. For example, we find three consecutive data as $8, 0, 43$. This situation is rare when there is no new patients in one week and 43 new patients in the following week. Term $\omega(t)$ is the function from interpolating the notification data. However, with data sequence with zeros as ' 8, 0, 43', $\omega(t)$ will be negative locally. In this case, we replace 0 with a small positive number to make $\omega(t)$ positive. For instance, in the case of $8, 0, 43$, we replace 0 with 4 to make $\omega(t)$ positive. We also observe frequencies pf one and three per year in the modulus of Fourier transform in (f).

Figure 3.6 presents $\beta(t)$ extracted from post-vaccination measles weekly notification data in London from year 1974 to 1986 by the incidence algorithm. We cut $\beta(t)$ into smaller parts, zoom in and plot them in (b), (c), (d), (e). Figure 3.6 (f) plots modulus of Fourier transform of $\beta(t)$ in London. We can see two peaks at 1 and 3 as well which reflects cycles of one and three times per year.

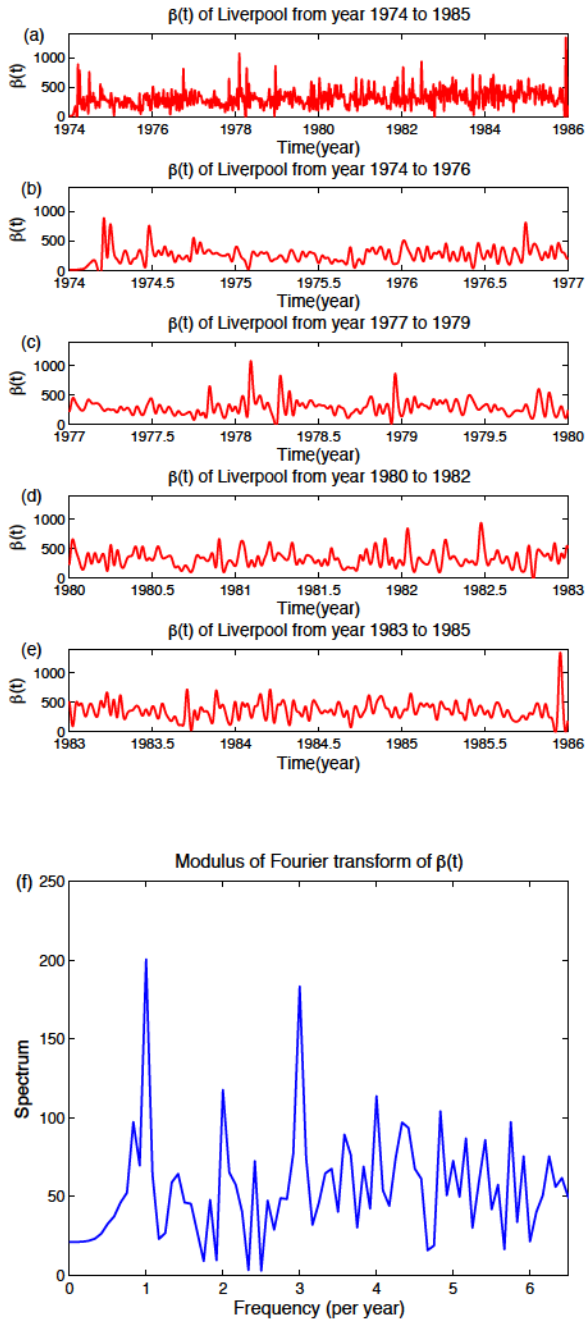


Figure 3.5: (a) Time-dependent transmission rate $\beta(t)$ of Liverpool from year 1974 to 1985. (b),(c),(d) and (e) are time-dependent transmission rate every three years from 1974 to 1976, 1977 to 1979, 1980 to 1982 and 1983 to 1985, respectively. $\beta(t)$ is extracted with parameters $\delta = 1/64/52/week, a = 52/52/week, \nu = 52/52/week, g = 1/16/52/week$ and initial values $S_0 = 0.25, E_0 = 0.001, I_0 = 0.001, A_0 = 0.7$. The panel (f) is the Fourier transform of transmission rate in as shown in (a).

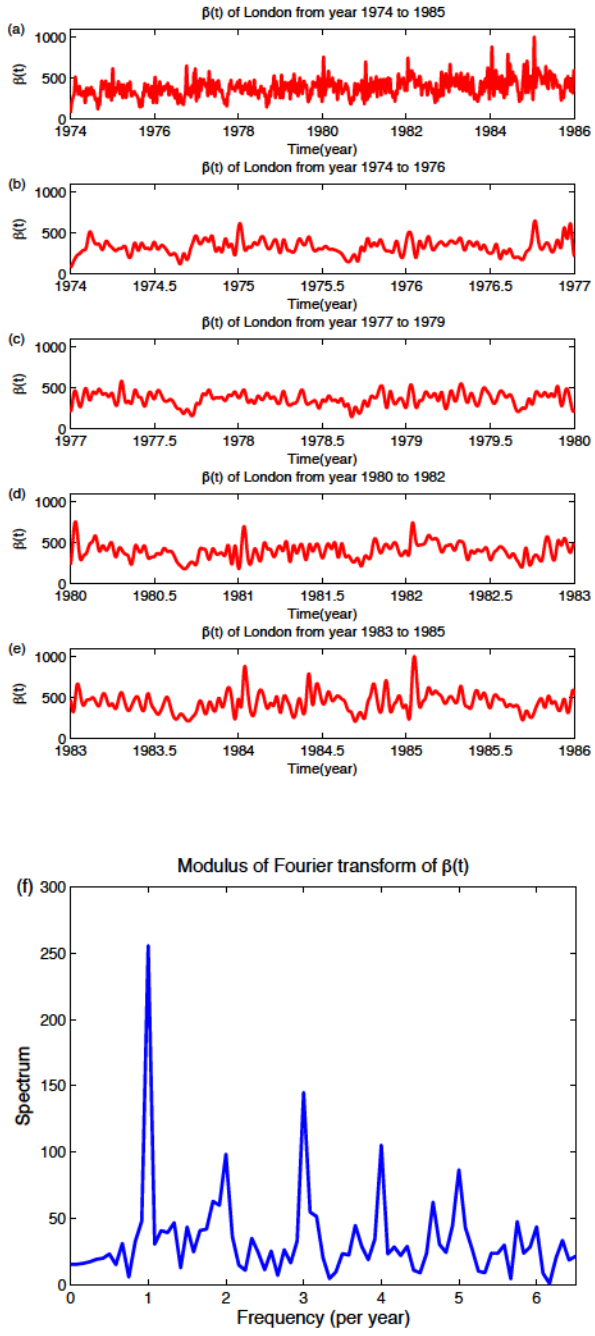


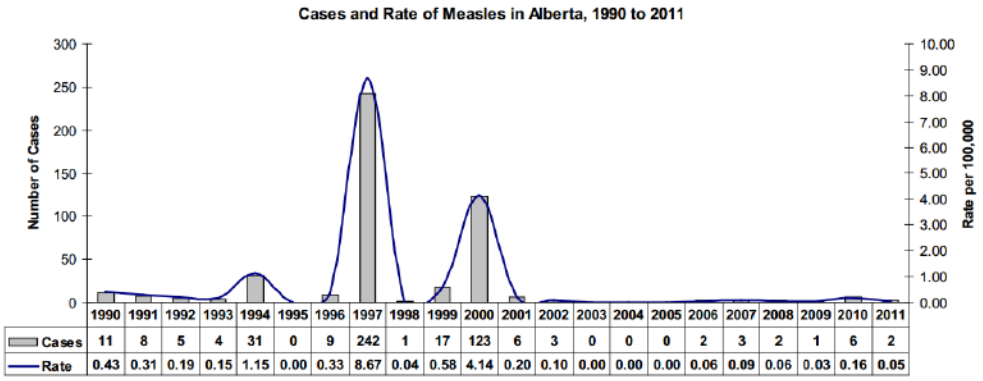
Figure 3.6: (a) Time-dependent transmission rate $\beta(t)$ of London from year 1974 to 1985. (b), (c), (d) and (e) are time-dependent transmission rate every three years from 1974 to 1976, 1977 to 1979, 1980 to 1982 and 1983 to 1985, respectively. $\beta(t)$ is extracted with parameters $\delta = 1/64/52/week, a = 52/52/week, \nu = 52/52/week, g = 1/16/52/week$ and initial values S_0=0.15, E_0=1 e-05, I_0=1 e-04, A_0=0.7.The panel (f) is the Fourier transform of transmission rate as shown in (a).

Although the dominant frequencies for pre-vaccination is the same with post-vaccination, the dominant is stronger in the pre-vaccination situation. This is because we use two data sets, the measles notification data and the vaccination rate data, to extract the transmission rate while only one data set is used in the pre-vaccination situation.

Dominant frequencies in London have less noise than those in Liverpool because London is a much larger cite in population and it is less sensitive to unexpected factors. The fact that vaccination does not change the dominant frequencies in both cities strengthens our belief that transmission of measles is driven by seasonality and school holidays. Again, we prove that transmission cycles are synchronized in different cities.

3.4 Case study: the current Alberta measles outbreak in 2014

Measles is well controlled in Alberta. During the last decade, from 2001 to 2011, only total of 25 measles cases all over Alberta. Figure 3.7 presents number and rate of measles cases of Alberta from 1990 to 2001 [31]. There were two outbreak since its two-dose vaccine regimen was introduced in 1996. An outbreak is unlikely to happen when the pool of un-vaccinated kids are small. But after more than 10 'peaceful' years in Alberta, vaccination rate is lagging and more susceptibles will possibly lead to an measles outbreak. This year, 31 cases were confirmed which is huge compared to the number last decade [32].



	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Cases	11	8	5	4	31	0	9	242	1	17	123	6	3	0	0	0	2	3	2	1	6	2
Rate	0.43	0.31	0.19	0.15	1.15	0.00	0.33	8.67	0.04	0.58	4.14	0.20	0.10	0.00	0.00	0.00	0.06	0.09	0.06	0.03	0.16	0.05

Source: Alberta Health Communicable Disease Reporting System data pulled by onset date on January 12, 2012.

Figure 3.7: Cases and rate of measles in Alberta, 1990 to 2011.

In this section, we use our $SEIRA$ model to predict the possible measles outbreak and expected date in Alberta. Everyone who is not vaccinated is at high risk, thus, when there is a potential outbreak, the government would try to vaccinate all unvaccinated individuals, especially school kids. In this sense, we we use the wildlife model. We add an additional vaccination term for all kids to make our model more realistic. We assume that the vaccination rate for all kids who are not vaccinated nor never get infected (group S) q, i.e. except for newborn babies vaccination, hospitals should vaccinate $q\times$ 100% more kids every year. Thus, we modify our model as

$$\begin{aligned}
\frac{dS}{dt} =& \delta(1-p)A(t) - \beta S(t)I(t) - gS(t) - \mathbf{qS(t)},\\
\frac{dE}{dt} =& \beta S(t)I(t) - aE(t) - gE(t),\\
\frac{dI}{dt} =& aE(t) - \nu I(t) - gI(t),\\
\frac{dR}{dt} =& \nu I(t) - gR(t) + \delta pA(t) + \mathbf{qS(t)},\\
\frac{dA}{dt} =& g(S(t)+E(t)+I(t)+R(t)) - \delta A(t).
\end{aligned}$$

With above model, we test the dynamics of $I(t)$ with different q. We fix pediatric vaccination rate as $p = 0.8$. Other parameters are the default values for measles as shown in table 2.1. The population of Alberta is 4082571, 2014 [6]. Thus the initial value of I is about (30 confirmed cases)/(Alberta population) $\approx$ 0.75 e-05. We choose initial values of both $I(t)$ and $E(t)$ as 0.5 e-05, which is smaller than the fraction of the total confirmed cases. The vaccination rate for newborn babies can be as high as 85% nowadays, but the average rate over the past decade is not as high. Therefore, we assume that the average vaccination rate is 72 %. As we assumed previously, the fraction of susceptibles is 20 % and the fraction of group A is 80 %. Thus, we pick the initial value of $S(t)$ as 20%$\times$(1-72%)=0.056 and $R(t)$ as 0.144.

For the transmission rate $\beta(t)$, we follow Tidd *et al.* (1993) and set

$$\beta(t) = \beta_0(1 + \beta_1\phi(t))$$

where

$$\phi(t) = 1.5(0.68 + cos2\pi t)/(1.5 + cos2\pi t)$$

Here, $\beta_0 = 1000$ year^{-1} or 2.74 day^{-1}, and β_1 can vary between 0.2 and 0.28. This formula is more accurate in modeling school year effect than a simple sinusoid function.

Figure 3.8 predicts the outbreak peak and time of measles cases with different q. From the figure, we can see that when q is small (less than 5%), there are two outbreaks and the second one is more severe. As q becomes larger(greater than 5%), the first outbreak is more severe. Recall that when we simulate the behavior of $I(t)$ in chapter 2 as shown in figure 2.2, peaks are decaying. However, the second peak is more severe when $q < 5\%$. This is because the value of the transmission rate β is a function instead of a constant. Severe outbreak happens when vaccination can not offset the high level of the transmission rate during school days.

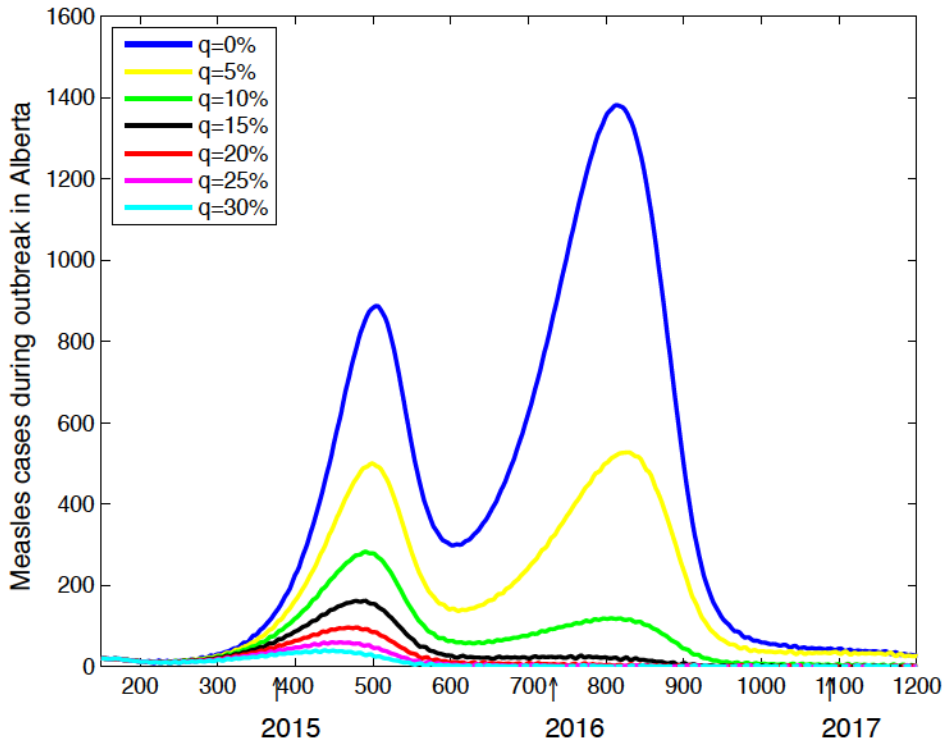


Figure 3.8: Fraction of infectives $I(t)$ with different vaccination rate with parameters $\delta = 1/64/365/day, a = 52/365/day, \nu = 52/365/day, g = 1/16/365/day$ and initial values $S_0 = 0.056,\ E_0 = 0.144,\ I_0 = 0.5$ e-05, R_0=0.5 e-05.

Obviously, the higher the vaccination rate is, the less severe the outbreak will be. However, the government and hospitals need to spend a lot of money if they peruse a very high vaccination rate. If a lower vaccination rate can lead to a similar result with a high vaccination rate, our money will be more worth spending if we control it at the lower level. For instance, if both 90% and 95% can prevent an outbreak, then 90% is a more reasonable rate to control.

Figure 3.9 plots vaccination rate versus both outbreak peaks of $I(t)$. We can see that $I(t)$ decreases faster with small vaccination rate q. Also, the second outbreak is more sensitive to q since vaccination has longer time to affect the outbreak peak than the first outbreak.

From the figure, can see that if q is greater then 35%, both the first and second outbreak is prevented.

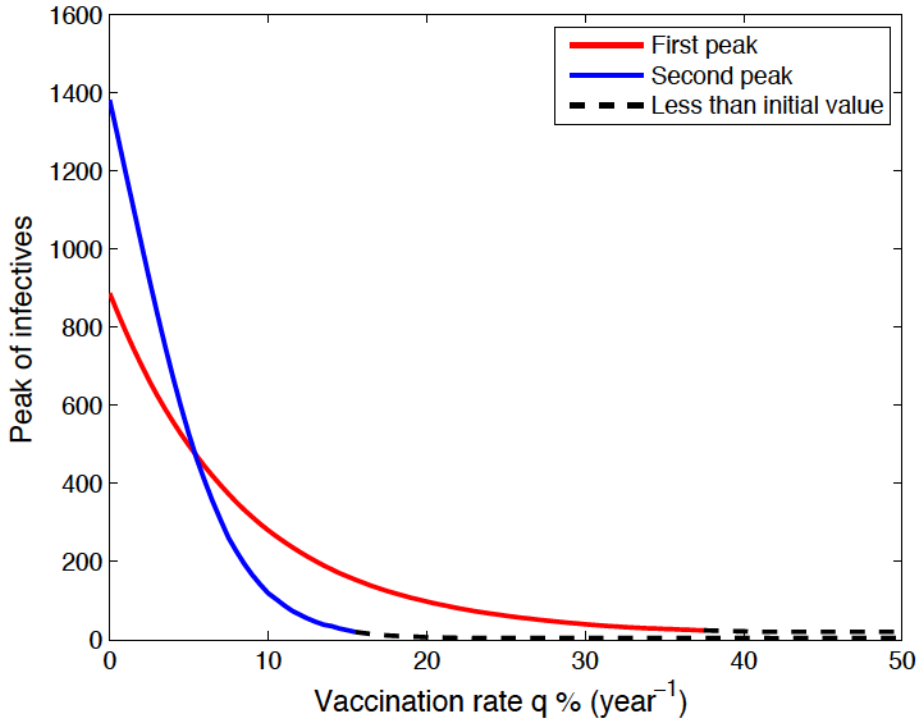


Figure 3.9: The outbreak peak of $I(t)$ as q varies from 0 to 40% per year with parameters $\delta = 1/64/365/day, a = 52/365/day, \nu = 52/365/day, g = 1/16/365/day$ and initial values $S_0 = 0.056,\ E_0 = 0.144,\ I_0 = 0.5$ e-05 , R_0=0.5 e-05 .

The dashed lines is figure 3.9 is when the maximum value of outbreaks is the initial value.

Table ?? lists the predicted outbreak peaks and dates. We further confirm our guess that outbreak amplitude is more sensitive with smaller vaccination rate. We can see that it is more efficient when q is smaller. For example, when q increases from 0 to 5%, the first outbreak decreases by 389, and when q increases from 35% to 40%, the first outbreak decreases by 7.

From table 3.4 we can see more clearly how measles cases will decrease as

vaccination rate q increases. The outbreak will not change after q is larger than 35% since $I(t)$ will decrease as soon as the pathogen was introduced.

Change of the vaccination rate	First peak deduction	Second peak deduction
$0\% \rightarrow 5\%$	389	835
$5\% \rightarrow 10\%$	218	407
$10\% \rightarrow 15\%$	118	94
$15\% \rightarrow 20\%$	64	16
$20\% \rightarrow 25\%$	36	5
$25\% \rightarrow 30\%$	22	3
$30\% \rightarrow 35\%$	12	0
$35\% \rightarrow 40\%$	7	0
$40\% \rightarrow 45\%$	1	0
$45\% \rightarrow 50\%$	0	0

Table 3.4: The quantitative impact of vaccination rate on the measles cases.

Over all, we would suggest that the government should control the vaccination rate at least greater than 5%, it is better that they can increase q to 35%. The government should not bother to increase q to be greater than 35%.

3.5 Conclusion

In this chapter, we consider pediatric vaccination. Again, we analyze the boundedness, positivity and stability of this model. It has the same positivity and boundedness properties. Also, there are two equilibria, one disease-free equilibrium and one endemic equilibrium. When the basic reproduction number $R_0 < 1$, the disease-free equilibrium is locally asymptotically stable and the endemic equilibrium is not feasible, i.e. the steady state is negative. When $R_0 > 1$, the disease-free equilibrium is unstable and the endemic equilibrium is locally asymptotically stable. The only difference with chapter 2 is that the threshold here is $R_0 = \frac{a\beta\delta(1-p)}{(a+g)(g+\delta)(\nu+g)} = 1$.

From sensitivity analysis, we know that the transmission rate is more important in controlling the endemic level of infectives at where pediatric vaccination rate is

higher while it is among the lowest in controlling an outbreak time and amplitude.

We extend both the prevalence algorithm and the incidence algorithm for the $SEIRA$ model. We verify that the dominant frequencies are not affected by vaccination which strengthens our common belief that the transmission of measles is driven by seasonality and school holidays.

The Alberta measles study suggested that vaccination rate of 35% per year is the optimum rate in preventing an outbreak as well as money saving.

Chapter 4

Discussion

4.1 Concluding remarks

In this thesis, we presented an infectious disease model that could better describe the childhood infectious diseases. Mathematical and numerical investigations have revealed a number of biologically and mathematically significant results.

Stability analysis shows that when $R_0 \geq 1$, the $SEIRA$ system goes to the endemic equilibrium and when $R_0 < 1$, it goes to the disease-free equilibrium where $R_0 = \frac{a\beta\delta(1-p)}{(a+g)(g+\delta)(\nu+g)}$ is the basic reproduction number.

Sensitivity analysis proves the importance of quarantining patients to prevent an epidemic outbreak while birth and removal rate are the most important factors in controlling the endemic level of patients which indicts the importance of medication treatment. The transmission rate is more important in controlling the endemic level of infectives at where pediatric vaccination rate is higher while it is among the lowest in controlling an outbreak time and amplitude.

Studying the inverse method of extracting the transmission rate proves the advantages the incidence algorithm has over the prevalence algorithm. The dominant

frequencies of 1 and 3 per year is not affected by vaccination which strengthens our belief that the transmission of measles is driven by seasonality and school holidays, and seasonality is more dominant, especially in large cities. Also, cycles of the transmission rate is synchronized in different cities, i.e. different cities have the dominant frequencies.

Case study of measles outbreak in the province of Alberta verifies the efficiency of wildlife vaccination. For measles outbreak in Alberta, we would suggest that the government should control the vaccination rate at least greater than 5% per year, and it is better to increase it to 35% per year.

4.2 Future directions

Estimating the initial values

Estimating of the initial value is a great challenge. From the inverse method, we can see that initial value of $\beta(t)$ and $\beta'(t)$ can, to a huge extent, determine the behavior of the transmission rate. Initial value of $S(t)$ will not affect the pattern of the transmission rate, but the level of the transmission rate $\beta(t)$. It is a challenge to estimate it since it is depended on the vaccination policy during the past decade.

Data processing

Real disease data sets are used to test whether a mathematical model is reasonable or to extract information from the model. In practice, errors exist both in the data itself or in the process of manipulating data. If errors or the effect of errors are minimized, we can obtain better and more accurate results. More work should be put on how to minimize errors or to derive a better algorithm which will generate less errors.

Multi schedules vaccination

Vaccination is efficient in preventing an endemic outbreak. However, successful vaccination may fail to develop successful immunity. In many countries, especially developed countries, program of multi doses vaccination are conducted. Double-dose measles vaccination are introduced in Canada.

According to Alexander et al. (2006) [1], the population can be classified as Susceptible (S), Vaccinated (S_ν), Infectious (I), and Booster Vaccinated (V) who are immune for life. Therefore the system can be expressed as

$$\begin{aligned}
\frac{dS}{dt} &= (1-p)\delta - \beta SI - \delta S - \xi S + \mu S - \nu, \\
\frac{dS_\nu}{dt} &= p\delta + \xi S - (1-\alpha)\beta S_\nu I - (\delta + \rho + \mu)S_\nu, \\
\frac{dI}{dt} &= \beta SI + (1-\alpha)\beta S_\nu I - (\delta_\gamma)I, \\
\frac{dV}{dt} &= \rho S - \nu + \gamma I - \delta V.
\end{aligned}$$

where p is pediatric vaccination rate, δ is death/bith rate, α is represents the efficiency of the vaccine, μ is the waning rate following the pediatric vaccination ,$1/\gamma$ is infectious period, and ρ and ξ are the rates of booster vaccine to previously vaccinated and susceptible ratio, respectively.

Bibliography

[1] Alexander, M. E., Moghadas, S. M., Rohani, P., & Summers, A. R. (2006). Modelling the effect of a booster vaccination on disease epidemiology. Journal of mathematical biology, 52(3), 290-306.

[2] Anderson, R. M., May, R. M., & Anderson, B. (1992). Infectious diseases of humans: dynamics and control (Vol. 28). Oxford: Oxford university press.

[3] Bjørnstad, O. N., Finkenstädt, B. F., & Grenfell, B. T. (2002). Dynamics of measles epidemics: estimating scaling of the transmission rates using a time series SIR model. Ecological Monographs, 72(2), 169-184.

[4] Clarkson, J. A., & Fine, P. E. (1985). The efficiency of measles and pertussis notification in England and Wales. International journal of epidemiology, 14(1), 153-168.

[5] Earn, D. J., Rohani, P., Bolker, B. M., & Grenfell, B. T. (2000). A simple model for complex dynamical transitions in epidemics. Science, 287(5453), 667-670.

[6] Economic dash board, Alberta,Canada.

[7] Fenner, F., Henderson, D. A., Arita, I., Jezek, Z., Ladnyi, I. D., & World Health Organization. (1988). Smallpox and its eradication/F. Fenner...[et al.].

[8] Gallagher, The Reverend John A. (1936). The Irish Emigration of 1847 and Its Canadian Consequences. Canadian Catholic Historical Association Report, University of Manitoba

[9] Hadeler, K. P. (2011). Parameter identification in epidemic models. Mathematical biosciences, 229(2), 185-189.

[10] "Health. De-coding the Black Death". BBC. 3 October 2001

[11] Retrieved August,13, 2014 at: http: //en.wikipedia.org/wiki/Pandemic#cite_ref-2

[12] Infectious disease data. Available at http://people.biology.ufl.edu/bolker/measdata.html.

[13] International Infectious Diseases Data Archive.

[14] Keeling, M. J., & Rohani, P. (2008). Modeling infectious diseases in humans and animals. Princeton University Press.

[15] Kermark, M. D., & Mckendrick, A. G. (1927). Contributions to the mathematical theory of epidemics. Part I. In Proc. R. Soc. A (Vol. 115, No. 5, pp. 700-721).

[16] Lombard, M., Pastoret, P. P., & Moulin, A. M. (2007). A brief history of vaccines and vaccination. Revue Scientifique et Technique-Office International des Epizooties, 26(1), 29.

[17] London, W. P., & Yorke, J. A. (1973). Recurrent outbreaks of measles, chickenpox and mumps I. Seasonal variation in contact rates. American Journal of Epidemiology, 98(6), 453-468.

[18] Mummert, A., Weiss, H., Long, L. P., Amigó, J. M., & Wan, X. F. (2013). A perspective on multiple waves of influenza pandemics. PloS one, 8(4), e60343.

[19] Pollicott, M., Wang, H., & Weiss, H. (2012). Extracting the time-dependent transmission rate from infection data via solution of an inverse ODE problem. Journal of biological dynamics, 6(2), 509-523.

[20] Rothman, K. J. (2012). Epidemiology: an introduction. Oxford University Press. Chicago

[21] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., ... & Tarantola, S. (2008). Global sensitivity analysis: the primer. John Wiley & Sons.

[22] "Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003". World Health Organization (WHO)

[23] The weekly OPCS (Office of Population Censuses and Surveys) reports, the Registrar General's Quarterly or Annual Reports & various English census reports 1948-1967.

[24] Tidd, C. W., Olsen, L. F., & Schaffer, W. M. (1993). The case for chaos in childhood epidemics. II. Predicting historical epidemics from mathematical models. Proceedings of the Royal Society of London. Series B: Biological Sciences, 254(1341), 257-273.

[25] Wolfram Mathematical Documentation Center

[26] Xia, Y., Bjørnstad, O. N., & Grenfell, B. T. (2004). Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. The American Naturalist, 164(2), 267-281.

[27] "Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003". World Health Organization (WHO)

[28] Retrieved July 27, 2014 at: http://www.who.int/topics/measles/en/

[29] Retrieved July 28, 2014 at: http://www.health.alberta.ca/documents/Guidelines-Measles-2013.pdf

[30] Alberta Health Interactive Health Data Application (IHDA) data.

[31] Alberta Health, Public Health Notifiable Disease Management Guidelines, November 2013.

[32] Alberta Health Service, retrieved Aug 12, 2014 at: http://www.albertahealthservices.ca/9842.asp