MATHEMATICAL MODELLING OF THE SPREAD OF MOSQUITO-BORNE DISEASES AND CHILDHOOD INFECTIONS

ZHAO SHI

M.Phil.

The Hong Kong Polytechnic University
2018

The Hong Kong Polytechnic University
Faculty of Applied Science and Textiles
Department of Applied Mathematics

Mathematical Modelling of the Spread of Mosquito-borne Diseases and Childhood Infections

Shi Zhao

A thesis submitted in partial fulfilment of the requirements for
the degree of Master of Philosophy
January, 2018

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____(signed)

_____Zhao Shi_____(name of student)

Dedication

Dedicate to my parents, those beloved, and the very interest I hold to epidemiology.

Abstract

Introduction Emerging infectious diseases (EIDs) in recent years have captured worldwide attention due to their potential for the rapid spread between countries and continents. In this thesis, we are going to explore the characteristics and try to explain the patterns of infectious diseases including Yellow Fever (YF) in Luanda, Angola 2015-16, Zika Virus Diseases (ZVD) in Northeastern (NE) Brazil 2015-16, Japanese Encephalitis (JE) in Hong Kong 2004-16 and Varicella (or chickenpox) in Shenzhen, China 2013-15, in order to further interpret the features and key factors of infectious diseases including potential impact factors (e.g., climatic factors, vector abundance, human behaviors and vaccination programs ... etc.). We also combined the game theoretical framework with an epidemic model (based on the compartmental SIR model) to study the decision-making process regarding to travelling during an outbreak, and to investigate the effects of travel strategies on local disease control.

Data Collection Cases data were collected from various public domains: Center of Health Protection (CHP) of Hong Kong government, Yellow Fever situation reports of the World Health Organization (2016), Shenzhen Centers for Disease Control and Prevention (CDC) and Minister of Health of Brazil. Besides, regional climatic data were also collected for analysis.

Methods On the basis of classical models in epidemiology, we constructed innovative compartmental models and agent-based models for specified infectious diseases and for particular research goals. For Yellow Fever (YF), a novel compartmental model was built up, which includes both host and vector populations and time-dependent vector abundance. In addition, we also considered the local vaccination campaign. For Zika Virus Diseases (ZVD), we also established compartmental model with hosts' and time-dependent vectors' populations, we model ZVD epidemics according to local GBS time series (in Northeastern Brazil, where ZVD hit hardest among the world from 2015-16) and we studied the relationship between the possible infectivity of asymptomatic infection and the final ZVD infection attack rate (IAR). For Japanese Encephalitis (JE), as the ratio of pig population and human JE cases was also explored in the same level, we build up an epidemiology model among local pig population and connect to human cases with a spill-over rate. In the model, we considered long-term mosquitoes, pigs and humans dynamic, we studied the"skip-and-resurge" of JE epidemics in Hong Kong and hypothesize that "new JEV strain invaded Hong Kong around 2011". For Varicella, an agent-based model is constructed to study the varicella infection among school children, and simulate the effects of different school-based vaccination programs. Last but not least, game theory is employed to model the individual decision-making process. A game theoretical framework is combined with an epidemic model to study the decision-making process regarding to travelling during an outbreak. The group optimal strategy that maximizes overall population utility is also computed.

Results For YF, the local vaccination campaign saved 5.1-fold more people from death (with 73 death reported originally), the possible human reaction to recent YF deaths (i.e., the death-driven transmissibility) is likely to explain the transmission pattern of the YF epidemic in Luanda, and we report very low YF IAR (0.09-0.15%) and high YF cases reporting ratio (71%). For ZVD, we found the exceeding local GBS time series can explain the first

ZVD epidemics (the first wave) of NE Brazil in 2015-16 and the infectivity of asymptomatic infections are positively related to the ZVD IAR of 2015-16. For JE, the simple mathematical model can re-generate the long-term JE epidemics in Hong Kong, we report without vectors JEV cannot maintain among swine, the dramatical decrease of local living pigs was likely to be responsible for "skip" of JE from 2006-10, and we show high confidence in the hypothesis that "the resurge of JE since 2011 was likely due to new strain invaded Hong Kong". For varicella, our agent-based (or school-based) model fits the observed cases data well and introducing school-based vaccination program can effectively prevent large-scale varicella outbreaks (particularly during summer). At last, for the epidemiological travelling game, we find perfect agreement between individual and group optimal strategies for a range of epidemiologically and economically plausible values. However, in regions where disagreement occurs, the conflict between the individual optimum (corresponding to a "voluntary entrance" scheme) and the group optimum (a "restricted entrance" scheme) is often extreme. In this region, model outcomes are highly sensitive to small changes in the infection transmissibility and traveller costs/benefits.

Conclusion Infectious disease is a great threat to human health all over the world, with the ability to spread among the population. Simple ODE equations and compartmental model plays an important role in studying the spreading pattern and transmission of infectious diseases. Our mathematical frameworks have significant theoretical value for exploring infectious diseases, improving our understanding of the dynamics and helping us to take appropriate strategies. Regarding to the travelling game theory framework, we conclude that a conflict between individually optimal and group optimal travel strategies during an outbreak may not occur under many scenarios, but in other cases, extreme conflicts could emerge suddenly even under slight changes in epidemiological or economic conditions.

Publications arising from the Thesis

Here lists the publications arising from this thesis by the thesis submission date:

- Zhao S, Stone L, Gao D, He D. (2018) Modelling the large-scale yellow fever outbreak in Luanda, Angola, and the impact of vaccination. PLoS Neglected Tropical Diseases. 12(1):e0006158. https://doi.org/10.1371/journal.pntd.0006158.
- Tang X, Zhao S, Chiu APY, Ma H, Xie X, et al. (2017) Modelling the transmission and control strategies of varicella among school children in Shenzhen, China. PLoS ONE. 12(5):e0177514. https://doi.org/10.1371/journal.pone.0177514.

Acknowledgements

The endeavor of work is a non-isolated process. I am grateful to those who have supported me during my M.Phil study and would like to hereby acknowledge their assistance.

I would like to express my sincere appreciation to my Chief Supervisor, Dr. Daihai He, for his patience, enlightening guidance, generously financial support and constant encouragement, without which this work would not have been possible. For his unwavering support, I am truly grateful. Throughout these years, I have learned a lot from his rigorous attitude and passion to scientific research, and most importantly, being a decent man.

My appreciation also goes to Dr. Yijun Lou, Prof. Lewi Stone, Prof. Chris T. Bauch, Dr. Lin Yang, Prof. Daozhou Gao and Prof. Shigui Ruan, for their spiritually and academically supports during my M.Phil study.

The indispensable support of my M.Phil experience has been the guidance and kindness of my Co-Supervisor, Prof. Ka-fai Yiu, Cedric, who has been a source of inspiration and encouragement.

My thanks also goes to my parents, for their constantly love, encouragement and support.

I would also like to thank my friends for encouraging me during my study, as well as research fellows at the Hong Kong Polytechnic University.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In recent years, emerging infectious diseases (EIDs) have attracted wide public attention because of their potential for spread. Because of their epidemic potential, they are threats to both global and local public health. It is of great importance to ascertain the transmission and spatial pattern of these EIDs, and to implement effective measures to control the spread. In this thesis, we are going to explore the characteristics and try to explain the patterns of infectious diseases including Yellow Fever (YF) in Luanda, Angola 2015-16, Zika Virus Diseases (ZVD) in Northeastern (NE) Brazil 2015-16, Japanese Encephalitis (JE) in Hong Kong 2004-16 and Varicella (or chickenpox) in Shenzhen, China 2013-15. These diseases have invaded countries and continents and taken its toll around the world. This chapter will provide background information of these EIDs and also introduce the key statistical inference approach we applied in modelling frameworks. We will give a brief introduction to the commonly used models, establish the objectives of this thesis, key biological parameters of infectious diseases, and summarize some main findings and conclusions in this thesis.

1.1 Backgrounds on Vector-borne Diseases

In this thesis, vector-borne diseases are referring to the infectious diseases that are transmitted via certain vectors. The possible vector can be in living (e.g., mosquito, tick, etc.) or non-living (e.g., water, air, etc.) forms, but commonly, the vectors are referring to some living creatures. Similarly, in this thesis, vector-free diseases are referring to the infectious diseases that can be transmitted without vectors. These kinds of diseases are

usually transmitted based on contact (e.g., sexually or non-sexually). If the "vector" of one disease is in "non-living" form (e.g., air-borne diseases), one can regard this disease as a vector-free disease because the disease transmission is almost wherever there is a infectious host.

An infectious disease could also transmitted in both vector-borne and vector-free paths (e.g., Zika virus). Some other diseases also include additional part in their transmission path (e.g., Japanese encephalitis virus), for example, reservoirs (in most situations, they are animals).

1.2 Brief Introduction on Compartmental Models in Epidemiology

Compartmental models divide the population into different compartments, in this section, we will firstly introduce the simplest SIR model, then introduce the simplest SIR-SI model for vector-borne diseases and finally, use SIR-SI model as an example to derive the basic reproduction number ($\mathcal{R}_0$). These two model is the most direct and simplest way to study the epidemics of most infectious diseases.

1.2.1 SIR model

Susceptible-infectious-recovered (SIR) model is one of the most basic and commonly used compartmental models in epidemiology [144]. The standard SIR model divides into susceptible (S), infectious (I), removed (R) classes and many others under specific assumptions. S indicates the class of individuals who are susceptible to the disease but are not infected yet. I are those have been infected and are able to spread the disease by contacting the susceptible (S). R are those have been removed from the disease transmission compartments (either by recovery, death or leaving the system...etc.) and have gained lifelong immunity. For most of the vector-free diseases, their transmission pattern can be simply modelled by SIR model (see Fig 1.1 as the diagram of SIR model). As we are focusing on the host population (e.g., humans) when modelling vector-free diseases, we denote S_h, I_h and R_h as susceptible, infectious and removed respectively. The model is as follows (see Eqn.

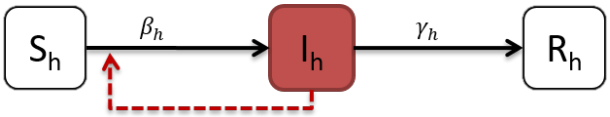


Figure 1.1: Diagram of SIR model. Red compartment represents infected class.

(1.1)):

$$\begin{cases} \frac{dS_h}{dt} = -\beta_h \cdot \frac{I_h}{N_h} \cdot S_h \\ \frac{dI_h}{dt} = \beta_h \cdot \frac{I_h}{N_h} \cdot S_h - \gamma_h I_h \\ \frac{dR_h}{dt} = \gamma_h I_h \end{cases} \tag{1.1}$$

where $N_h = S_h + I_h + R_h$ represents the number of population. β_h denotes the effective contact rate (i.e., term β_h and be model in detail as the product of contact rate and the successful transmission probability per contact) between susceptible and infectious. γ_h denotes the average removing rate as γ_h^{-1} can be regarded as the mean removing period from infectious class to removed class.

1.2.2 SIR-SI model

Similar to the standard SIR model, for vector-borne diseases (e.g., arbovirus diseases), we add vectors' classes into the model and further divide into two parts: susceptible (S_v) and infectious (I_v) (see Fig 1.2 as the diagram of SIR-SI model). Vector-borne diseases can

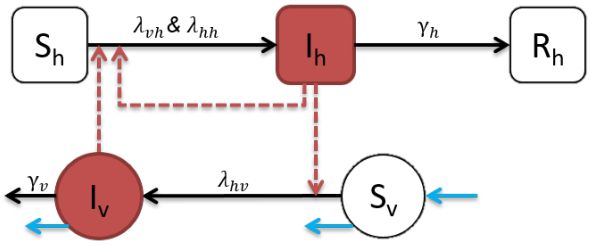


Figure 1.2: Diagram of SIR-SI model. Black arrows represent infection status transition paths, red dashed arrows represent transmission paths, and the blue arrows represent the natural birth and death pathway. Square compartments represent host classes and circular compartments represent vectors classes. Red compartments represent infectious classes.

be transmitted by either bite of vectors (from vector to host and from host to vector) or by direct contact between hosts (i.e., like the vector-free transmission), thus, the model is as

follows (see Eqn. (1.2)):

$$\begin{cases} \frac{dS_h}{dt} = -(\lambda_{vh} \cdot \frac{I_v}{N_h} + \lambda_{hh} \cdot \frac{I_h}{N_h}) \cdot S_h \\ \frac{dI_h}{dt} = (\lambda_{vh} \cdot \frac{I_v}{N_h} + \lambda_{hh} \cdot \frac{I_h}{N_h}) \cdot S_h - \gamma_h I_h \\ \frac{dR_h}{dt} = \gamma_h I_h \\ \\ \frac{dS_v}{dt} = b_v N_v - \lambda_{hv} \cdot \frac{I_h}{N_h} \cdot S_v - \mu_v S_v \\ \frac{dI_v}{dt} = \lambda_{hv} \cdot \frac{I_h}{N_h} \cdot S_v - (\gamma_v + \mu_v) I_v \end{cases} \tag{1.2}$$

where $N_h = S_h + I_h + R_h$ and $N_v = S_v + I_v$ represents the number of hosts' and vectors' population respectively. Subscripts $_h$ denotes hosts' classes and $_v$ denotes vectors' classes. λ_{xy} denotes the effective transmission rate from x to y, specially, λ_{hh} represents the same transmission patch as term β_h in Eqn. (1.1). γ_h and γ_v denote the removing rate of hosts and vectors respectively. b_v and μ_v denote the natural birth and death of vectors' population.

1.2.3 Basic reproduction number

The basic reproduction number, $\mathcal{R}_0$, is the expected number of secondary cases produced by one typical infection joining in a completely susceptible population [68, 221]. When $\mathcal{R}_0 < 1$, the disease would die out in long run. While if $\mathcal{R}_0 > 1$, the disease would spread among the population and may cause a pandemic.

According to [221], a systematic procedure to calculate the $\mathcal{R}_0$ by solving the dominant eigenvalue (i.e., the eigenvalue with the largest real part) of the next generation matrix ($\mathbf{G}$). Regarding the next generation matrix $\mathbf{G} = \mathbf{F}\mathbf{V}^{-1}$, matrix $\mathbf{F}$ is the new infection matrix and matrix $\mathbf{V}$ is the infection transfer matrix. The entry of i-th row and j-th column of matrix $\mathbf{F}$: $F_{i,j} = \frac{\partial \mathcal{F}_i}{\partial x_j}$ with $\mathcal{F}_i$ is the i-th equation of $\mathcal{F}$ and x_j is the j-th variable of the vector of infected classes. The entry of i-th row and j-th column of matrix $\mathbf{V}$: $V_{i,j} = \frac{\partial \mathcal{V}_i}{\partial x_j}$ with $\mathcal{V}_i$ is the i-th equation of $\mathcal{V}$ and x_j is the j-th variable of the vector of infected classes. $\mathcal{F}$ is the vector of rate of transmission (i.e., the changing rates from infectious to non-infectious classes) and $\mathcal{V}$ is the vector of rate of transition (i.e., the changing rates among infectious classes). $\mathbf{F}$ is the Jacobian of $\mathcal{F}$ and $\mathbf{V}$ is the Jacobian of $\mathcal{V}$, and we derive $\mathbf{F}$ and $\mathbf{V}$ under the disease free equilibrium [221].

For the SIR model (see Eqn. (1.1)), $\mathcal{F} = \left(\beta_h \cdot \frac{I_h}{N_h} \cdot S_h\right)$ and $\mathcal{V} = (\gamma_h I_h)$, thus, as the vector of infected classes is only (I_h), $\mathbf{F} = (\beta_h)$ and $\mathbf{V} = (\gamma_h)$, and we have the next generation matrix $\mathbf{G} = \mathbf{F}\mathbf{V}^{-1} = \left(\frac{\beta_h}{\gamma_h}\right)$. According to $\mathbf{G}$, we derive the basic reproduction number $\mathcal{R}_0 = \frac{\beta_h}{\gamma_h}$.

Moreover, for the SIR-SI model (see Eqn. (1.2)), compartment I_h and I_v are regarded as the infected classes (i.e., should be considered in the vector of infected classes). We have $\mathcal{F} = \begin{pmatrix} (\lambda_{vh} \cdot \frac{I_v}{N_h} + \lambda_{hh} \cdot \frac{I_h}{N_h}) \cdot S_h \\ \lambda_{hv} \cdot \frac{I_h}{N_h} \cdot S_v \end{pmatrix}$ and $\mathcal{V} = \begin{pmatrix} \gamma_h I_h \\ (\gamma_v + \mu_v) I_v \end{pmatrix}$. Then, $\mathbf{F} = \begin{pmatrix} \lambda_{hh} & m\lambda_{vh} \\ \lambda_{hv} & 0 \end{pmatrix}$ and $\mathbf{V} = \begin{pmatrix} \gamma_h & 0 \\ 0 & \gamma_v + \mu_v \end{pmatrix}$. $\mathbf{G} = \mathbf{F}\mathbf{V}^{-1} = \begin{pmatrix} \frac{\lambda_{hh}}{\gamma_h} & \frac{m\lambda_{vh}}{\gamma_v + \mu_v} \\ \frac{\lambda_{hv}}{\gamma_h} & 0 \end{pmatrix}$.

Therefore, for SIR-SI model, $\mathcal{R}_0 = \frac{\lambda_{hh} \cdot (\gamma_v + \mu_v) + \sqrt{[\lambda_{hh} \cdot (\gamma_v + \mu_v)]^2 + 4m\gamma_h \cdot (\gamma_v + \mu_v)\lambda_{hv}\lambda_{vh}}}{2\gamma_h \cdot (\gamma_v + \mu_v)}$. According to [111], the basic reproduction number of vector-borne transmission is $\mathcal{R}_v = \sqrt{\frac{m\lambda_{hv}\lambda_{vh}}{\gamma_h \cdot (\gamma_v + \mu_v)}}$ and basic reproduction number of vector-free (i.e., from host to host) transmission is $\mathcal{R}_h = \frac{\lambda_{hh}}{\gamma_h}$. The basic reproduction number is given as $\mathcal{R}_0 = \frac{\mathcal{R}_h + \sqrt{\mathcal{R}_h^2 + 4\mathcal{R}_v^2}}{2}$. Furthermore, if we only consider the vector-free transmission (i.e., by ignoring the effect of vector-borne transmission, we set $\mathcal{R}_v \to 0^+$), we can see the $\mathcal{R}_0$ of SIR-SI model (see Eqn. (1.2)) is equivalent to the $\mathcal{R}_0$ of SIR model (see Eqn. (1.1)).

1.3 Partially Observed Markov process

Partially observed Markov process (POMP, i.e., a Markov process with unobserved states), also known as Hidden Markov Model (HMM), is a state-space model. The time series being modelled is assumed to be a POMP model (see [133, 146, 147]).

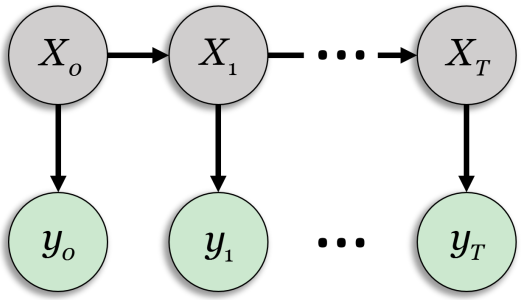


Figure 1.3: Diagram of Partially Observed Markov process.

Fig 1.3 shows the structure of POMP, where $(X_0, X_1, X_2, \cdots, X_T)$ are unobserved and modelled as the real number of cases, and $(y_0, y_1, y_2, \cdots, y_T)$ are simulated and modelled as the reported number of cases. Based on POMP structure, iterated filtering [147] with plug-and-play statistical inference framework [133] is applied to solve the fitting problem and calculate the maximum likelihood estimate (MLE) of parameters. In addition, plug-and-play likelihood-based statistical inference framework is implemented according to following steps (this part will also be detailedly explained in each modelling chapters):

- A stochastic simulation model is developed by the fixed-time-step Euler-multinomial algorithm.
- Probability measurement is assumed to follow a specific probability (e.g., negative binomial or Poisson) distribution.
- Sequential Monte Carlo (SMC) is employed for likelihood estimation and iterated filtering method is conducted to obtain MLEs.
- The secondary small-sample-size corrected Akaike's Information Criterion (AICc) or Bayesian Information Criterion (BIC) are applied to quantify the tradeoff between the goodness-of-fit of a model and its complexity [205].
- The confidence intervals (C.I.s) of parameters are estimated by using the method of profile likelihood [132, 133].
- The partial rank correlation coefficients (PRCCs) are adopted for models' sensitivity analysis [111].

The above model framework is implemented by using `R` package "`POMP`" [146]. We define Markov process $X(t)$ (i.e., the actual time series at time t_n) depends on the state at time t_{n-1} and model parameters (denoted as vector Θ) as:

$$X(t_n) = f(X(t_{n-1}); \Theta) \tag{1.3}$$

where Θ is the vector of parameters under estimation, $X(t_n)$ is the state of time series at time t_n and function f corresponds to the compartmental models. Initial stages $X(t_0)$ at some time $t_0 < t_1$ is specified.

Fig 1.4 presents detailed steps of iterated filtering. More detailed information about R package "POMP" can be found via [146, 147].

Algorithm 3: Iterated filtering: `mif(P, start=`θ_0`, Nmif=`M`, Np=`J`, rw.sd=`$\sigma_{1:p}$`, ic.lag=`L`, var.factor=`C`, cooling.factor=`a`)`, using notation from Table 1 where `P` is a class 'pomp' object with defined `rprocess`, `dmeasure`, `init.state`, and `obs` components.

input: Starting parameter, θ_0; simulator for $f_{X_0}(x_0;\theta)$; simulator for $f_{X_n|X_{n-1}}(x_n \,|\, x_{n-1};\theta)$; evaluator for $f_{Y_n|X_n}(y_n \,|\, x_n;\theta)$; data, $y^*_{1:N}$; labels, $I \subset \{1,\dots,p\}$, designating IVPs; fixed lag, L, for estimating IVPs; number of particles, J, number of iterations, M; cooling rate, $0 < a < 1$; perturbation scales, $\sigma_{1:p}$; initial scale multiplier, $C > 0$.

for m *in* $1{:}M$ **do**
- Initialize parameters: $[\Theta^F_{0,j}]_i \sim \text{Normal}\left([\theta_{m-1}]_i, (Ca^{m-1}\sigma_i)^2\right)$ for i in $1{:}p$, j in $1:J$.
- Initialize states: simulate $X^F_{0,j} \sim f_{X_0}(\,\cdot\,;\Theta^F_{0,j})$ for j in $1{:}J$.
- Initialize filter mean for parameters: $\bar{\theta}_0 = \theta_{m-1}$.
- Define $[V_1]_i = (C^2+1)(a^{m-1}\sigma_i)^2$.
- **for** n *in* $1:N$ **do**
  - Perturb parameters: $[\Theta^P_{n,j}]_i \sim \text{Normal}\left([\Theta^F_{n-1,j}]_i, (a^{m-1}\sigma_i)^2\right)$ for $i \notin I$, j in $1:J$.
  - Simulate prediction particles: $X^P_{n,j} \sim f_{X_n|X_{n-1}}(\,\cdot\,|X^F_{n-1,j};\Theta^P_{n,j})$ for j in $1:J$.
  - Evaluate weights: $w(n,j) = f_{Y_n|X_n}(y^*_n|X^P_{n,j};\Theta^P_{n,j})$ for j in $1:J$.
  - Normalize weights: $\tilde{w}(n,j) = w(n,j)/\sum_{u=1}^{J} w(n,u)$.
  - Apply Algorithm 2 to select indices $k_{1:J}$ with $\mathbb{P}[k_u = j] = \tilde{w}(n,j)$.
  - Resample particles: $X^F_{n,j} = X^P_{n,k_j}$ and $\Theta^F_{n,j} = \Theta^P_{n,k_j}$ for j in $1:J$.
  - Filter mean: $[\bar{\theta}_n]_i = \sum_{j=1}^{J} \tilde{w}(n,j)[\Theta^P_{n,j}]_i$ for $i \notin I$.
  - Prediction variance: $[V_{n+1}]_i = (a^{m-1}\sigma_i)^2 + \sum_j \tilde{w}(n,j)([\Theta^P_{n,j}]_i - [\bar{\theta}_n]_i)^2$ for $i \notin I$.
- **end**
- Update non-IVPs: $[\theta_m]_i = [\theta_{m-1}]_i + [V_1]_i \sum_{n=1}^{N} [V_n]_i^{-1}\left([\bar{\theta}_n]_i - [\bar{\theta}_{n-1}]_i\right)$ for $i \notin I$.
- Update IVPs: $[\theta_m]_i = \frac{1}{J}\sum_j [\Theta^F_{L,j}]_i$ for $i \in I$.

end

output: Monte Carlo maximum likelihood estimate, θ_M.

complexity: $\mathcal{O}(JM)$

Figure 1.4: Algorithm of iterated filtering. Figure is from [147].

1.4 Outline of this thesis

Chapter 2 presents a novel compartmental model was built up to study the transmission pattern of yellow fever (YF) in Luanda, Angola from Dec 2015 to Aug 2016, the model includes both host and vector populations and time-dependent vector abundance. The model also considered the local vaccination campaign and the possible transmissibility of asymptomatic infections. According to simulation results, the local vaccination campaign

saved 5.1-fold more people from death (with 73 death reported originally), the possible human reaction to recent YF deaths (i.e., the death-driven transmissibility) is likely to explain the transmission pattern of the YF epidemic in Luanda, and we report very low YF IAR (0.09-0.15%) and high YF cases reporting ratio (71%).

In chapter 3, we build up an epidemiology model among local pig population and connect to human cases with a spill-over rate in order to study the long-term "skip-and-resurge" of Japanese Encephalitis (JE) epidemics in Hong Kong. The model considered the long-term mosquitoes, pigs and humans dynamic, and hypothesized that "new JEV strain invaded Hong Kong" around 2011. Model results indicate that the simple mathematical model can re-generate the long-term JE epidemics in Hong Kong, we report without vectors JEV cannot maintain among swine, the dramatical decrease of local living pigs was likely to be responsible for "skip" of JE from 2006-10, and we show high confidence in the hypothesis that "the resurge of JE since 2011 was likely due to new strain invaded Hong Kong".

In chapter 4, a compartmental model is established with hosts' and time-dependent vectors' populations, Zika Virus Diseases (ZVD) epidemics, in Northeastern (NE) Brazil 2015-16, is modelled according to local GBS time series and this chapter studies the relationship between the possible infectivity of asymptomatic infection and the final ZVD infection attack rate (IAR). Model results show the exceeding local GBS time series can explain the first ZVD epidemics (the first wave) of NE Brazil in 2015-16 and the infectivity of asymptomatic infections are positively related to the ZVD IAR of 2015-16.

In chapter 5, an agent-based (or school-based) model is constructed to study the varicella infection among school children, and simulate the effects of different school-based vaccination programs. This chapter reports that the proposed agent-based model fits the observed cases data well and introducing school-based vaccination program can effectively prevent large-scale varicella outbreaks (particularly during summer).

In chapter 6, a game theoretical framework is combined with an epidemic model to study the decision-making process regarding to travelling during an outbreak. Travellers can play an important role in infectious disease transmission, acting both as a source of case imports and contributing to the susceptible pool at their destination. To balance the economic benefits of travelling against potential infection risks, game theory is employed to model the individual decision-making process. The game theoretical framework is combined

with an epidemic model (based on the compartmental SIR model) to investigate the effects of travel strategies on local disease control. The group optimal strategy that maximizes overall population utility is also computed. For this epidemiological travelling game, we find perfect agreement between individual and group optimal strategies for a range of epidemiologically and economically plausible values. However, in regions where disagreement occurs, the conflict between the individual optimum (corresponding to a "voluntary entrance" scheme) and the group optimum (a "restricted entrance" scheme) is often extreme. In this region, model outcomes are highly sensitive to small changes in the infection transmissibility and traveller costs/benefits. Simulations show how uncontrolled traveller inflow can cause an unexpected large-scale outbreak when the disease risk level suddenly rises even a small amount, although government travel restrictions according to group optimal levels can effectively control the outbreak in this situation. We conclude that a conflict between individually optimal and group optimal travel strategies during an outbreak may not occur under many scenarios, but in other cases, extreme conflicts could emerge suddenly even under slight changes in epidemiological or economic conditions.

Chapter 7 includes summary and introduction of future works based on this thesis.

Chapter 2

Modelling Yellow Fever in Luanda, Angola

Yellow fever (YF), transmitted via bites of infected mosquitoes, is a life-threatening viral disease endemic to tropical and subtropical regions of Africa and South America. YF has largely been controlled by widespread national vaccination. Nevertheless, between December 2015 and August 2016, YF resurged in Angola, quickly spread and becoming the largest YF outbreak for the last 30 years. Recently, YF resurged again in Brazil (December 2016). Thus, there is an urgent need to gain better understanding of the transmission pattern of YF.

The present study provides a refined mathematical model, combined with modern likelihood-based statistical inference techniques, to assess and reconstruct important epidemiological processes underlying Angola's YF outbreak. This includes the outbreak's attack rate, the reproduction number ($\mathcal{R}_0$), the role of the mosquito vector, the influence of climatic factors and the unusual but unnoticed appearance of two-waves in the YF outbreak. The model explores actual and hypothetical vaccination strategies, and the impacts of possible human reactive behaviors (e.g., response to media precautions).

While there were 73 deaths reported over the study period, the model indicates that the vaccination campaign saved 5.1-fold more people from death and saved from illness 5.6-fold of the observed 941 cases. Delaying the availability of the vaccines further would have greatly enhanced the epidemic in terms of cases and deaths. The analysis estimated a mean $\mathcal{R}_0 \approx$ 2.6-3.4 and an YF attack rate of 0.09-0.15% (% population infected by YF) over the

whole period from December 2015 to August 2016. Our estimated initial and upper bound $\mathcal{R}_0$ are in line with previous studies.

Unusually, $\mathcal{R}_0$ oscillated in a manner that was out-of-phase with the weekly death reports. High recent numbers of deaths were associated with periods of relatively low disease transmission and low basic reproduction number, and vice-versa. The time-series of Luanda's YF cases suggest the outbreak occurred in two waves, a feature that would have become far more prominent had there been no mass vaccination. The waves could possibly be due to protective reactive behavioral changes in the population affecting the mosquito population. The second wave could well be an outcome of the March-April rainfall patterns in the 2016 El Niño year by creating ideal conditions for the growth of the mosquito vectors.

The modelling framework is a powerful tool for studying future YF epidemic outbreaks, and provides a basis for future vaccination campaign evaluations.

2.1 Introduction

Yellow fever (YF) is a life-threatening viral disease endemic to tropical regions of Africa and South America. The disease is transmitted in urban areas primarily via the bites of infected female *Aedes aegypti* mosquitoes, which is also the vector of Zika, dengue and chikungunya viruses [12, 45, 174]. Rural and intermediate YF are transmitted by sylvatic and peri-domestic *aedes* species in Africa. For those infected with YF, the disease incubates in the first 3-6 days of onset, after which there is an abrupt "period of infection" of intense viremia lasting for 3-4 days (fever, weakness, headache, nausea, muscle pain) [172]. This is followed by a period of remission in which the symptoms reduce and settle, and most infected individuals recover at this stage. Thus some 70-85% of individuals infected cases are asymptomatic or have at most a mild case of YF. However, 15-25% of patients relapse and move to a "period of intoxication" characterized by abdominal pain, vomiting, jaundice (yellow skin and eyes) and often culminating in death. The case-fatality-ratio (CFR) in this latter subset is understood to be approximately 20% among the general population, and 50% among hospitalized cases [172], although the CFR is well known to be highly variable, and dependent on the particular circumstances. Like Ebola, YF is classified as a viral hemorrhagic fever, although it is responsible for a 1000-fold more illness and death than

Ebola [174].

In 2016, YF resurged in Angola to become the largest YF outbreak on record over the last 30 years [55]. In swift response, almost all global stocks of the YF vaccine were exhausted by April 2016. Similar to the Angolan experience, YF recently resurged in Brazil in December 2016, where it continues to expand towards the Atlantic coast in regions not previously deemed at risk (as of March 16, 2017) [45]. Thus there is an urgent need to gain a better understanding of the transmission patterns of YF. Here we develop a mathematical model to help identify the key epidemic processes behind the Angolan outbreak in 2015-16, and the impact and effectiveness of the vaccination campaign.

The first case of YF in Angola were seen on December 5, 2015 but reported in the media only on 20 January 2016 [238]. By November 2016, the large YF epidemic of Angola and the Democratic Republic of Congo resulted in 962 confirmed infections including 393 reported deaths [45]. YF is vaccine preventable and the vaccine can confer long-lasting immunity. The vaccine is suitable for individuals of age 9 months or older. As such, the Angolan government initiated a vaccination campaign to prevent the spread of yellow fever on the first week of February 2016 [28, 45]. More than 10 million doses were needed for the whole country [238]. The center of the outbreak was in Angola's capital, Luanda province. Estimates suggest that vaccination coverage of Luanda province was 38.0% at the end of January 2016, and reached 80.0% by mid-March 2016, and 93.0% by mid-June 2016 [8, 9, 27, 45].

Fig. 2.1 graphs the epidemic curve of YF case numbers (probable and confirmed; as defined in Data section) in Luanda province as obtained from the WHO [8, 9]. The graph peaks in February 2016, when large-scale vaccination was introduced, and then followed by a period of rapid decline in case numbers. Despite the major vaccination effort, the epidemic proved tenacious rather than die out as predicted, and persisted for a sustained period of time forming a long "tail" in reported case numbers from April to August (see Fig. 2.1). Also unusual is the minor peak in case numbers that occurred in May, followed soon after by an increase in deaths, despite the pressure of the vaccination and control efforts. By modelling YF time series of Luanda, our goal is to reconstruct the important epidemiological processes that help explain these different and sometimes nonintuitive features. The model allows estimation of the attack rate of the outbreak, and the basic reproduction number ($\mathcal{R}_0(t)$), which was changing during the epidemic. Moreover, the model is able to explore the role of the mosquito vector, and the unusual waves of the YF outbreak, which we find would have

become even more apparent had there been no vaccination. Some exploration of the role of climatic variables is also possible.

As it is well known, the basic reproduction number ($\mathcal{R}_0$) is an important measure to a disease's transmissibility, and is one of the first parameters that need to be estimated in any epidemiological study. Recall that $\mathcal{R}_0$ is defined as the number of secondary cases a single typical infected individual infects over the lifetime of the disease [50]. A recent study estimated $\mathcal{R}_0$ to lie between 5.2-7.1 at the early stage of the 2016 YF outbreak in Angola [240]. However, $\mathcal{R}_0$ was found to decrease with time as the epidemic proceeded. Kraemer *et al.* [151] estimated $\mathcal{R}_0$ to be 4.8 (95% C.I.: 4.0-5.6) for Angola, although this was possibly an over-estimate given reporting rates were not stable. In summary, the literature suggests that YF is highly transmissible with direct estimates of the reproduction number being $\mathcal{R}_0 \approx 5$, which is almost double that of pandemic influenza ($\mathcal{R}_0$ is from 1.5 to 3.6 [116, 159, 170, 234]) and Ebola ($\mathcal{R}_0$ is from 1.2 to 2.0 [104, 118, 184, 220]). In this chapter, our analysis uses modern statistical inference techniques to estimate $\mathcal{R}_0$ from the time-series in Fig. 2.1. Unlike other modelling studies, our procedure also examines how reactive protective behavior (e.g., insecticide, vector-control, travel restrictions possibly in response to news and media precautions), may lead to changes in $\mathcal{R}_0$, and allows us to explore the implications of this reaction.

Any model of YF must take into account that most infected individuals are asymptomatic or mild-symptomatic (individuals who show only fever but not jaundice) [12, 88, 173, 175, 182, 224], making the disease difficult to detect and under-reported in the first phases. With only a slight abuse of terminology, it simplifies the modelling that follows, to classify mild symptomatic individuals as though they were asymptomatic cases. Thus asymptomatic cases refer to all individuals who do not have severe YF. It is well understood that asymptomatic YF infections can be infectious [224] and therefore may act as "silent sources" of YF virus [224]. Asymptomatic infections, thus, have the potential to play an important role in disease transmission. It was previously understood that 6 out of 7 YF infections could be asymptomatic[175]. However, a recent meta-analysis based on 11 independent studies, suggested that the asymptomatic ratio should be 55% [142]. Given the lack of information on the proportion and infectivity of asymptomatic YF cases, we examine a number of different relevant scenarios.

To the best of our knowledge, this is the first detailed modelling of YF that includes

both the host and vector populations, and the asymptomatic and severe (those exhibiting fever and jaundice) cases in the host populations. Previous models that assessed vaccination impact on YF have not included these fundamental components and pathways in a comprehensive approach. By fitting the time-series of the Angola outbreak, its evolution over time and its curtailment with vaccination, it becomes possible to statistically infer key model parameters. This, in turn, makes it possible to simulate alternative "what if" scenarios, and examine what might have happened under different vaccination schemes.

2.2 Data and Methods

2.2.1 Data and case definitions

We study time-series of YF cases from the province of Luanda of Angola with a population of 6,543,000 in 2016 [8, 31, 45]. The African Health Observatory (AHO) published weekly YF data for Luanda province reporting 941 (confirmed and probable) cases and 73 deaths over the study period from December 5, 2015 to August 18, 2016

Probable cases (see [42]) are those "with acute onset of fever, with jaundice appearing within 14 days of onset of the first symptoms and one of the followings: (i) presence of yellow fever IgM antibody in the absence of YF immunization within 30 days before onset of illness; or (ii) positive postmortem liver histopathology; or (iii) epidemiological link to a confirmed case or an outbreak.' Confirmed cases are defined as those positive to serological or PCR testing.

Similar to the WHO [8] and Kraemer *et al.* [151], both (weekly) probable cases and confirmed cases are grouped together and are referred to simply as "YF cases" or equivalently "severe cases" in this chapter. YF vaccination coverage in Luanda province, obtained from AHO reports, increased from 38% on February 2, 2016 when the vaccination campaign started to 93% on August 18, 2016 (see Fig. 2.1) [8] (also see section 2.3.4 for more details of the local vaccination coverage data). The vaccination coverage was determined by a linear interpolation of reported data (see the blue squared line in Fig 2.1).

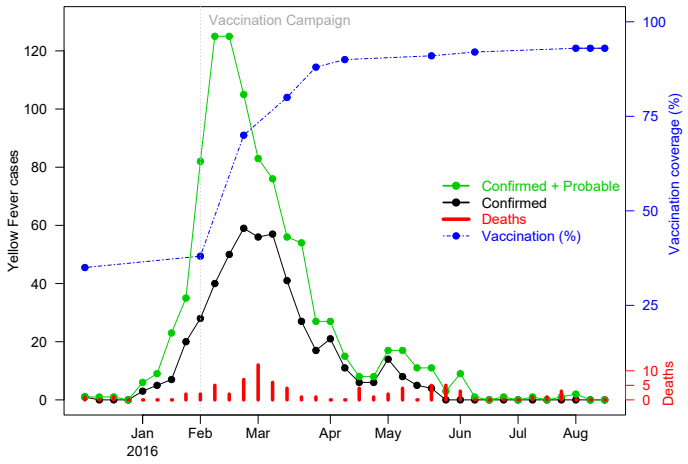


Figure 2.1: The observed yellow fever outbreak in Luanda from December 5, 2015 to August 18, 2016. Green dots denote the sum of probable and confirmed cases, black dots denote confirmed cases, red bars denote the confirmed death, and blue squares denote vaccine coverage. The vertical grey dashed line denotes the time point when the YF vaccination campaign was initiated.

2.2.2 Methods

Yellow Fever Model

Since YFV is not spread by human-to-human-transmission, the standard SIR type modelling approaches (which are based solely on human-to-human transmission) are inappropriate. Instead, we use a vector-host model of YFV transmission, as illustrated in Fig 2.2, which is based on well-known models of mosquito-borne diseases (dengue, Zika virus, etc.). With this two-host model, we are able to explore the impact of different control strategies (such as vaccination, reducing mosquito abundance and human exposure to mosquito) which could not be examined with approaches that fail to incorporate vector-host dynamics.

The model applied the following notations. For human host populations, S_h represents the number of susceptible individuals, E_h is the number of individuals exposed to YF but not the cases infected, A_h represents the number of asymptomatic cases, I_h the number of severe infectious individuals, and T_h the number of individuals in the toxic stage. Finally, R_h individuals have either recovered from the disease and/or have been vaccinated.

For the human host population, there are two main transmission pathways, as seen in Fig. 2.2.

$H_1)$ $S_h \rightarrow E_h \rightarrow I_h \rightarrow T_h \rightarrow R_h$

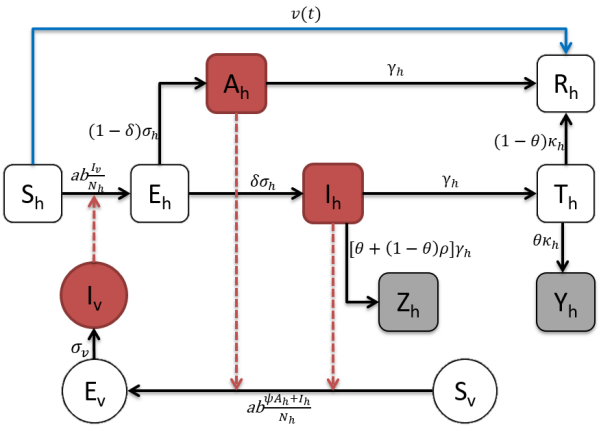


Figure 2.2: Flowchart of the yellow fever model. Black arrows represent infection status transition paths, red dashed arrows represent transmission paths and the blue arrow represents the vaccination pathway. Square compartments represent host classes and circular compartments represent vector classes. Red compartments represent infectious classes, and gray compartments are the simulated weekly reported cases (Z_h) and deaths (Y_h).

The susceptible hosts become exposed to YF by the bites of infectious mosquitoes, harbor the virus (move to the Exposed class) and eventually become infected (move to Infected class), enter the toxic stage (move to Toxic class) and then either eventually recover (move to Recovered class) and remain immune, or in the case of 5-50% in this stage (as specified by the CFR), die from the disease (D_h).

H_2) $S_h \to E_h \to A_h \to R_h$

In this second pathway, susceptible hosts become exposed to YF by the bites of infectious mosquitoes, harbor the virus and eventually become infected but only asymptomatic. The latter usually recover and gain future immunity.

In the above scheme, the infected and asymptomatic hosts may both infect mosquitoes should they be bitten, as shown in Fig. 2.2. However, asymptomatic cases have a reduced transmissibility, ψ, when compared to that of a typical severe case. Also, note that individuals in the toxic phase no longer have viremia [174], and therefore cannot be infectious [47, 172, 174].

The vector population has only a single pathway:

V_1) $S_v \to E_v \to I_v$

As in usual notation, S_v is the number of susceptible mosquitoes. It is assumed that vertical transmission of YF virus in the mosquito population is relatively small, and could reasonably

be neglected for the purpose of this model [72, 92, 174]. It is also understood that mosquitoes are relatively unaffected by the mosquito-borne viruses [136].

Based on the above descriptions, we formulate an ordinary differential equations (ODEs) system that matches the scheme as illustrated in Fig. 2.2:

$$\begin{aligned}
S_h' &= -ab\frac{I_v}{N_h}S_h - v(t-t_0) \\
E_h' &= ab\frac{I_v}{N_h}S_h - \sigma_h E_h \\
A_h' &= (1-\delta)\sigma_h E_h - \gamma_h A_h \\
I_h' &= \delta\sigma_h E_h - \gamma_h I_h \\
T_h' &= \gamma_h I_h - \kappa_h T_h \\
R_h' &= v(t-t_0) + \gamma_h A_h + (1-\theta)\kappa_h T_h \\
D_h' &= \theta\kappa_h T_h
\end{aligned} \tag{2.1}$$

$$\begin{aligned}
S_v' &= B_v(t) - ac\frac{\psi A_h + I_h}{N_h}S_v - \mu_v S_v \\
E_v' &= ac\frac{\psi A_h + I_h}{N_h}S_v - \sigma_v E_v - \mu_v E_v \\
I_v' &= \sigma_v E_v - \mu_v I_v
\end{aligned}$$

$$\begin{aligned}
Y_h^{(i)} &= \int_{\text{week } i} \theta\kappa_h T_h \, dt \\
Z_h^{(i)} &= \int_{\text{week } i} [\theta + \rho\cdot(1-\theta)]\cdot\gamma_h I_h \, dt
\end{aligned}$$

Here, $v(t)$ represents the vaccination rate at time t, and t_0 is the mean time period from receiving vaccination to acquiring full immunity. $Y_h^{(i)}$ represents the weekly recorded deaths due to YF for the i-th week. It is calculated as an integral which effectively sums the weekly number of toxic phase individuals (T_h) who leave or are removed from the toxic class ($\kappa_h T_h$) over one week. Only a fraction (θ) of the latter dies, where θ is the CFR for severe cases. Similarly the variable $Z_h^{(i)}$ denotes the weekly recorded observed cases. This is determined through the term $[\theta + \rho\cdot(1-\theta)]$ which collects the deaths (through θ) and the non-fatal severe cases via the severe case reporting ratio ρ (more detailed discussion regarding YF case counting can be found in section 2.4.5). It is assumed that all deaths are reported, and that the fitting procedure can estimate the reporting ratio of severe cases ρ.

In terms of the total host and vector population sizes (N_h and N_v), the following relations must hold:

$$N_h = S_h + E_h + A_h + I_h + T_h + R_h + \sum_i Y_h^{(i)} \longrightarrow \text{constant} \tag{2.2}$$

$$N_v = S_v + E_v + I_v \tag{2.3}$$

In our model, $N_v = N_v(t)$ is time-dependent in a manner that is controlled by the mosquito birth rate $B_v(t)$ and death rate $\mu_v(t)$, namely $N_v' = B_v(t) - \mu_v(t) \cdot N_v$. Following [111, 132], we suppose that

$$N_v(t) = m(t) \cdot N_h \tag{2.4}$$

Here the constant $N_h = 6,543,000$ is the number of humans in Luanda province. The parameter $m(t)$ is the time-dependent ratio of the mosquito-to-human populations that needs to be estimated. It is assumed that $m(t)$ is an exponential cubic spline function of time with number of nodes n_m (see section 2.2.4). Nodes are distributed uniformly over the time-domain with values (m_i) that are estimated but restricted to lie between 0 and 20. The range was chosen to reflect reality (where $m = 20$ implies $\mathcal{R}_0 = 10$, which is beyond the $\mathcal{R}_0$ upper limit for YF).

Model Parameters

The model is parameterized from prior knowledge of YF, and uses parameter values that are accepted in the literature. Table 3.1 summarises all model parameters and their ranges. Table 2.2 summarises parameter values for the different scenarios and model estimates (discussed below).

With regard to parameters for the host population in Eqn. (2.1), σ_h^{-1} and γ_h^{-1} represent the host latent and infectious period respectively, with both being approximately 4 days. The latent period also indirectly allows for a four-day reporting delay. Symptoms appear when patients leave the latent class, but are reported only when they leave the infectious class which is a four-day period. The toxic phase duration κ_h^{-1} is set to eight days.

The parameters on mosquitoes were taken from the dengue literature, where mosquito dynamics is also modelled. Since dengue and YF virus belong to the same family of viruses

Table 2.1: Summary of parameters.

Parameter	Notation	Value/Range	Unit/Remark	Source
mosquito biting rate	a	0.3 - 1.0	per vector·day	[51]
transmission probability from vector to host	b	0.10 - 0.75	per bite	[51]
transmission probability from host to vector	c	0.30 - 0.75	per bite	[80]
vaccination rate	v	0 - 0.043	per day	[8, 45]
host latent period	σ_h^{-1}	3 - 6	days	[12, 45, 141]
vector latent period	σ_v^{-1}	8 - 12	days	[172]
non-severe case relative infectivity	ψ	0.1 - 0.5	Nil	-
host infectious period	γ_h^{-1}	3 - 4	days	[12, 45]
severe case proportion	δ	15%	Nil	[12, 142]
severe case CFR	θ	0% - 50%	see text	[12, 41, 45, 172]
toxic case duration	κ_h^{-1}	7 - 10	days	[12, 45]
vector lifespan	μ_v^{-1}	4 - 35	days	[51, 80]
severe case reporting ratio	ρ	1% - 99%	Nil	-
initial susceptible host	$S_h.0/N_h$	0.62	fixed, Nil	[8]

(i.e., *flaviviridae*) and share the same type of vectors (i.e., *Aedes aegypti*), we follow the practice of previous studies and assume they have similar parameter values. Specifically, parameter values for the mosquito biting rate (a), the transmission probabilities (b, c), and the mosquito lifespan (μ_v^{-1}) were taken from Massad *et al.* [168], as indicated in Table 3.1. The vector latent period σ_v^{-1} and lifespan μ_v^{-1} were taken as 10 and 20 days respectively.

Basic Reproduction Number

Here we derive the basic reproductive number $\mathcal{R}_0$ for the vector-host model Eqn. (2.1). Following standard procedures [68, 221], the rate of transmission (i.e., the changing rates from infectious to non-infectious classes) is given by, $\mathcal{F}$:

$$\mathcal{F} = \begin{pmatrix} ab\frac{I_v}{N_h}S_h \\ 0 \\ 0 \\ ac\frac{\psi A_h + I_h}{N_h}S_v \\ 0 \end{pmatrix}$$

and the rate of transition (i.e., the changing rates among infectious classes), $\mathcal{V}$:

$$\mathcal{V} = \begin{pmatrix} \sigma_h E_h \\ (\delta-1)\sigma_h E_h + \gamma_h A_h \\ -\delta\sigma_h E_h + \gamma_h I_h \\ (\sigma_v + \mu_v) E_v \\ -\sigma_v E_v + \mu_v I_v \end{pmatrix}$$

Then, we have two Jacobian matrices, of which F is the Jacobian of $\mathcal{F}$ and V is the Jacobian of $\mathcal{V}$, and we derive F and V under the disease free equilibrium [221] (DFE, by setting the proportions of susceptible classes to be 100%),

$$F = \begin{pmatrix} 0 & 0 & 0 & 0 & ab \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & ac\psi m & acm & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.5}$$

where $F_{i,j} = \frac{\partial \mathcal{F}_i}{\partial x_j}$ with $\mathcal{F}_i$ is the i-th equation of $\mathcal{F}$ and x_j is the j-th variable of the vector (E_h, A_h, I_h, E_v, I_v),

$$V = \begin{pmatrix} \sigma_h & 0 & 0 & 0 & 0 \\ (\delta-1)\sigma_h & \gamma_h & 0 & 0 & 0 \\ -\delta\sigma_h & 0 & \gamma_h & 0 & 0 \\ 0 & 0 & 0 & (\sigma_v + \mu_v) & 0 \\ 0 & 0 & 0 & -\sigma_v & \mu_v \end{pmatrix} \tag{2.6}$$

where $V_{i,j} = \frac{\partial \mathcal{V}_i}{\partial x_j}$ with $\mathcal{V}_i$ is the i-th equation of $\mathcal{V}$ and x_j is the j-th variable of the vector (E_h, A_h, I_h, E_v, I_v). Then, we have the next generation matrix of our model,

$$FV^{-1} = \begin{pmatrix} 0 & 0 & 0 & \frac{ab\sigma_v}{\mu_v(\sigma_v+\mu_v)} & \frac{ab}{\mu_v} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{acm}{\gamma_h} \cdot (\psi + \delta - \psi\delta) & \frac{acm}{\gamma_h} \cdot \psi & \frac{acm}{\gamma_h} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.7}$$

Since $\mathcal{R}_0$ is the spectral radius of FV^{-1} [221], the basic reproduction number is

$$\mathcal{R}_0 = \sqrt{(\psi + \delta - \psi\delta) \cdot \frac{a^2 bcm}{\gamma_h} \cdot \frac{\sigma_v}{\mu_v(\sigma_v + \mu_v)}} \tag{2.8}$$

where m is the ratio of mosquito-to-human population. $\mathcal{R}_0(t)$ is calculated as a function of time, by assuming the mosquito-to-human population ratio to be $m = m(t)$ (see Eqn. (4.2)), which varies with time in a manner that may be estimated by the model fitting procedures.

When the population is not fully susceptible, it is it is a common practice to make use of $\mathcal{R}_{\text{eff}}$, the effective $\mathcal{R}_0$ given by $\mathcal{R}_{\text{eff}}(t) = \mathcal{R}_0(t) \cdot S_h(t)$, to describe the ability of a virus to invade the host population [143]. $\mathcal{R}_{\text{eff}}$ incorporates both the changes in the intrinsic ability of the virus, the characteristics of the mosquito vector as well as the availability of human host susceptibles. We note that the vaccination campaign can only reduce the availability of human susceptibles but not the intrinsic transmissibility of the virus. Vector control (e.g., mosquito fogging) can reduce the transmissibility of the virus, by reducing its vector of transmission. In our figures we plot both $\mathcal{R}_0(t)$ and $S(t)$. From these, it is simple to obtain $\mathcal{R}_{\text{eff}}(t)$.

Vaccination

The term $v(t)$ appearing in the equation for susceptible host dynamics (Eqn. (2.1)), represents the time-dependent vaccination rate of the host population (blue arrow in Fig 2.2). It is determined by considering the equation for susceptible dynamics (from Eqn. (2.1)): $dS_h/dt = -abI_v/N_h - v(t - t_0)$

Thus the rate of vaccinating people is $v(t)$ and the total number of people who would normally be vaccinated by time t is $V(t) = \int_0^t v(x)dx + V_0$. Note that t_0 is the mean time taken for an individual to gain full immunity after being vaccinated. Averaging the data reported by WHO and CDC [12, 45] gives $t_0 = 20$ days. The overall cumulative vaccination for Luanda province, as reported by WHO, is plotted in Fig. 2.1 as a percentage of the total population N_h. That is, the y-axis plots $V(t)/N_h \times 100$. Using this graph and the relation $dV(t))/dt = v(t)$ allows us to reconstruct $v(t)$ which we use when we numerically integrate Eqn. (2.1).

The constant V_0 denotes pre-existing immunity of the population at the beginning of the

2015-16 YF outbreak. As the attack rate of YF is typically very small and no major outbreaks occurred in Angola since 1988, we suppose that previously built-up immunity is relatively small and population immunity waned significantly over the next 27 years. Nevertheless, it was assumed that at the beginning of each simulation, $V_0 = 38\%$ of the population is already vaccinated to be consistent with WHO estimates. [8]. The 38% includes both the outcome of EPI (WHO's Expanded Program on Immunization) vaccination and immunity remaining from the mass campaign in 1988. (This is incorporated by setting $S_h.0/N_h = 0.62$ as in Table 3.1.) Wu *et al.* [240] assumed that initial immunity was equivalent to 28% vaccination coverage, which under-estimates the WHO data. Section 2.3.4 provides further information about YF vaccination doses.

The approach for modelling vaccination is adapted from our previous work on influenza [129], and avoids the need for inclusion of a separate vaccinated compartment in the model which would result in unnecessary additional complexity.

Different YF vaccination intervention scenarios are compared in order to evaluate the effectiveness of the actual national vaccination campaign. The best-fitting model to the data and the actual vaccination coverage will be taken as the "baseline scenario" as experienced in Luanda, which was initiated on February 2, 2016. This will be compared to three other additional hypothetical intervention scenarios.

- Actual vaccination campaign as experienced in Luanda (baseline scenario);
- 60, 120 and $\geqslant$ 180 days delay of vaccination campaign (hypothetical intervention scenarios

The total observed cases, as well as the total deaths, are evaluated by the model for each vaccination scenario.

The 180-day delay period, in fact, represents a "no-vaccination" scenario. When taking into account the extra 20 days required for vaccination to be effective, anyone vaccinated 180 days after February 2, 2016, will not gain any protective effect from the vaccination given the observation study period is only 200 days. Thus, any scenario with a delay greater than 180 days is equivalent to a no-vaccination scenario.

Fatality per infection and Case-Fatality-Ratio

Monath *et al.*, [172] estimated the fatality per infection of YF for the whole population to be in the range of 3-7.5% [172], and are variable across time and location. In an earlier study, Monath *et al.* estimated the fatality per infection to be in the range of 1-15% for Nigerian villages [177]. For severe YF cases, the Case-Fatality-Ratio (CFR) resulting in death (θ) is 20-50% or higher [115, 172, 188, 219], although the CFR is well known to be highly variable, and dependent on particular circumstances.

Given the large proportion of infected but asymptomatic YF cases, the accurate fatality per infection in Luanda cannot be determined without a comprehensive serological study, which to our knowledge has not been undertaken. But the Cases-Fatality Ratio (CFR, cases refer to severe cases) can be immediately approximated as the ratio of confirmed deaths to the confirmed (and probable) cases. For the data of Luanda province, the CFR is approximately 7.76%, and thus substantially lower than 20-50%. (Similar low estimates were noted by the WHO reports throughout the epidemic.) Moreover, because of reporting errors, we would expect the CFR to be even lower than this empirical estimate.

In this chapter, we choose CFR = 6% for our main simulations. But we have also carefully explored other possibilities. For example, in section 2.4.3, we run our fitting procedure to actually estimate the CFR given the data and find CFR = 4%. We also discuss what might be expected if the mortality data is under-reported (see section 2.4.4). More data and research is needed to gain a better understanding of the CFR and to check whether and how it changes over the study period.

Finally, we note that in our model, the fatality per infection is given by $\delta \cdot \theta$, where the CFR = θ, while δ is the proportion of severe cases. In this chapter, we consider either fixing the CFR to a value considered realistic or inferring the CFR from the data itself, as an extra parameter.

Asymptomatic infections

As reported by the CDC, "*asymptomatic or clinically unapparent infection is believed to occur in most YFV infections*" [12]. A case is defined as an asymptomatic infection only if it is confirmed strictly to have no symptoms, but is nevertheless found infectious as confirmed

by RNA or serological tests. It was previously believed that 6 out of 7 YF infections could be asymptomatic[175]. Recently, a meta-analysis based on 11 studies suggested that the asymptomatic ratio should be 55% and mild cases 33% (without jaundice), the rest 12% are severe cases[142]. As mentioned, to simplify presentation all mild cases (without jaundice) were considered to belong to the asymptomatic class. Two scenarios were examined:

1 : 85% asymptomatic ($\delta = 15\%$) and weak infectivity ($\psi = 0.1$)

2 : 85% asymptomatic ($\delta = 15\%$) and strong infectivity ($\psi = 0.5$)

The proportion of 85% for asymptomatic infections based on [142]. Scenarios 1 & 2 differ only in terms of their weak ($\psi = 0.1$) or strong ($\psi = 0.5$) infectivity. The results for scenarios 1 and 2 (with fixed CFR = 6%) are presented in the main text, while results of flexible CFR can be found in the section 2.4.3.

2.2.3 Plug-and-Play Inference Framework

The YF outbreak in Luanda is modelled as a Partially Observed Markov process (POMP) and makes use of the Iterated Filtering and plug-and-play likelihood-based inference frameworks to fit the data [111, 133, 147]. These are modern state-of-the-art statistical methodologies developed for fitting complex epidemiological datasets, and seeking Maximum Likelihood Estimate (MLE) for model parameters (`R` package "`POMP`" is available at [146]).

Bayesian Information Criterion (BIC) is employed as a criterion for model comparison, and quantifies the tradeoff between the goodness-of-fit of a model and its complexity [205]. The simulations made use of the Euler-multinomial integration method with the time-step fixed to be one day [48, 133].

The model is first fitted to the observed YF cases and deaths, given knowledge of the true vaccination coverage. The mosquito abundance is assumed to be unknown but time-dependent, and is reconstructed. We allow the basic reproduction number of our model to be time-dependent, given that the mosquito abundance is not fixed and given that human behavior can impact $\mathcal{R}_0(t)$ and change over the study period.

The parameter fitting and inference process are carefully checked, thereby giving high confidence that the fits of the observed time-series are accurate for reasons that are consistent

with the true underlying epidemiological processes rather than artificial model over-fitting. We conducted tests to find the best-fit model. For each asymptomatic scenario, we studied 10 different values of n_m (degrees of freedom in the $m(t)$), and compared them with BIC. BIC quantifies the trade off of the goodness of fitting of the model and the complexity of the model — penalizing models with more variables. A smaller BIC implies a better-fit model. For the best-fit model, the profile of maximum log likelihood was calculated as a function of the reporting ratio (see section 2.4.1 for further details). The profile found is always a reasonably smooth function. The model was run 1,000 times with the estimated parameters, and the median of the model simulation matched the reported weekly cases. Thus we can be confident that the maximization of models log likelihood converged and the estimation is consistent.

The simulated weekly reported cases Z_t are modelled by Eqn 2.1. The corresponding weekly observed cases, C_t, as given by the WHO, are assumed to follow a Negative-Binomial (NB) distribution as

$$C_t \sim \text{NB}\left(n = \frac{1}{\tau}, p = \frac{1}{1+\tau Z_t}\right) \qquad \text{with} \quad \text{mean} : \mu_t = Z_t \tag{2.9}$$

where τ denotes an over-dispersion parameter that needs to be estimated.

The weekly observed deaths, D_t, and the corresponding weekly simulated deaths, Y_t, are similarly related. Finally, the overall log-likelihood function, l, is given by

$$l(\Theta|C_1,\ldots,C_N;D_1,\ldots,D_N) = \sum_{t=1}^{T} \ln\left[L_t^{(C)} \cdot L_t^{(D)}\right] \tag{2.10}$$

where Θ denotes the parameter vector under estimation, and $L_t^{(C)}$ and $L_t^{(D)}$ are the probability measurement functions associated with C_t vs. Z_t, and D_t vs. Y_t, respectively. T denotes the total number of weeks during the study period.

The confidence intervals (C.I.) of parameters are estimated based on parameters' ranges in Table 3.1, using the method of profile likelihood confidence intervals [132, 133]. This is demonstrated in section 2.4.1 for the severe case reporting rate ρ. Parameter estimation and statistical analysis are conducted using `R` (version 3.3.3).

2.2.4 Fitting $m(t)$ with BIC

We performed extensive testing on fitting the model with different degrees of freedom in $m(t)$, and calculated the BIC as a function of the number of nodes in $m(t)$ from 2 to 10. This ensured that a wide range of possible shapes or profiles for $m(t)$ were explored, including possible constant, and monotonically increasing or decreasing profiles. The best model fit was chosen based on the smallest BIC, and the profile maximum log likelihood was determined as a function of the reporting ratio. The MLL profiles were plotted as a function of ρ, and yielded a smooth curve, which is an indicator of convergence. We also calculated the MLL (or BIC) as a function of n_m, the MLLs match. Thus we are confident that our MLL is the true maximum. The convergence of the maximization of log likelihood of the model given the data is guaranteed.

The median of 1,000 stochastic simulations of the best-fitting model matched the observed data, which indicates that it is the best model in all situations we explored.

All of the above steps were repeated (which involves fitting of dozens of models/parameter setting) under four different asymptomatic assumptions (as outlined in the main text). The possible impacts of climate and human behavioral responses (death driven vector control measure) were also considered (see section 2.3.3). Thus the total computational effort is huge.

We considered using three well-known test indices: AIC, BIC, DIC but ultimately decided on using BIC. This is because first, DIC has a number of known problems we prefer to avoid (see `http://avansp.github.io/2014/11/02/DIC-AIC-BIC.html`). Second, BIC is more appropriate than AIC in this chapter here, since the size of the data is relatively small. In the small sample size situation, a secondary AIC (i.e., AICc) can be considered [73]. Out investigation have made it clear that both BIC and AICc lead to the same conclusions in our examples.

2.2.5 Sensitivity Analysis

The Partial Rank Correlation Coefficients (PRCCs) are adopted for the model's sensitivity analysis [111]. Firstly, 1,000 random samples are taken for each model parameter from uniform distributions with parameter ranges as set out in Table 3.1. After that, for every

random parameter sample set, the YF model was simulated to obtain the target biological quantities (e.g., $\mathcal{R}_0$ and the total number of deaths in this chapter). Finally, PRCCs were calculated between each parameter and target biological quantities.

2.3 Results and Discussion

2.3.1 Model Fitting

The results for the best-fitting model under the two scenarios are shown in Fig 2.3. The model simulation median (of 1,000 simulations) of YF cases in Luanda is plotted in red and matches well with the observed patterns seen in weekly reported cases, both before and after the national vaccination campaign. The two scenarios (for asymptomatic infectivities) both model the data with almost the same goodness-of-fit with a ΔBIC ≈ 2 (see section 2.4.2 for the simulation results of strong infectivity scenario, i.e., scenario 2). That is, the observed and model time series are not significantly different for the two levels of infectivity [205]. As such the infectivity of asymptomatic cases cannot be accurately inferred from these data sets.

In Table 2.2, the over-dispersion τ, is notably small, which indicates the measurement model is close to a Poisson distribution (i.e., minor over-dispersion in measurement noise). This implies the reporting efforts (i.e., reporting ratios) were reasonably stable over time.

The analysis estimated a mean $\mathcal{R}_0 \approx$ 2.6-3.4 and an estimated YF attack rate of the whole period to be 0.09-0.15% (% population infected by YF) from December 2015 to August 2016. Our estimated initial and upper bound $\mathcal{R}_0$ are in line with previous studies.

Asymptomatic cases were not reported, and they might be considered as a completely hidden variable. However, if the number of asymptomatic cases is very large (e.g., if the asymptomatic-to-symptomatic ratio is 6:1 or 7:1) with a weak infectivity but full immunity, this will indirectly slow down the transmission of YF in the later stages, due to herd immunity built up by these silent asymptomatic cases. If their infectivity is strong, this will increase the difficulty to control the outbreak.

The model simulations of weekly deaths also fit the observed data well over the period

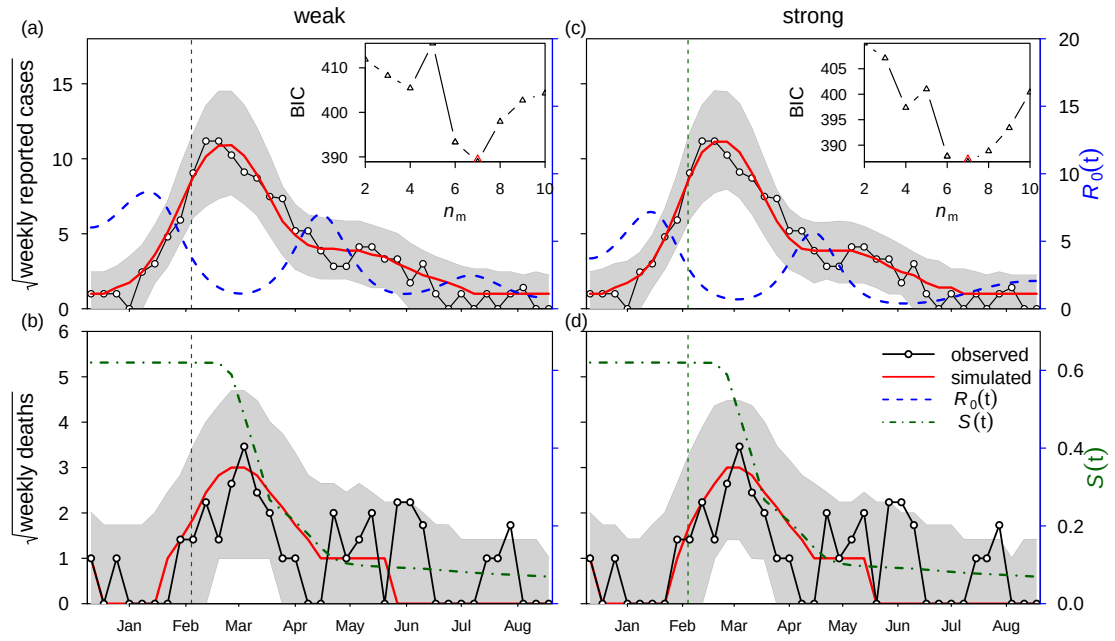


Figure 2.3: Model fitting results under two scenarios: scenario 1 in panels (a,b) and scenario 2 in panels (c,d). Black line with circles denote reported cases (in (a), in the form of square-root) and reported deaths (in (b), in the form of square-root), and red line denotes model simulation median. Blue dashed line denotes the fitted basic reproduction number, $\mathcal{R}_0(t)$, and the green dashed line shows the calculated host susceptible proportion, $S(t)$. Shaded region represents 95% bound of 1,000 model simulations. Vertical dashed line indicates the start date of the vaccination campaign. Inset panel shows BIC as a function of the number of nodes (n_m). The lowest BIC is attained at $n_m = 7$ in both scenarios, which is used in the main panel. Parameter values are listed in Table 2.2.

of the main epidemic until the end of April 2016. While the simulated median (red line) does not predict the two relatively small and erratic peaks at the beginning of June and end of July, nevertheless they fit reasonably within the 95% bounds. Note that similar peaks in death numbers appear in the delayed vaccination scenarios Figs. 2.9 b,d,f, where case numbers are higher)

The observed YF deaths are relatively "noisy" compared to the continuously observed YF cases (see red bars versus green dotted line in Fig 2.1), which might be due to the lower case numbers involved or possibly spatial variation of the YF CFR. The same holds for individual model simulations. The total number of deaths was only 6% out of all reported cases (CFR=6%), and 71% of the deaths appeared during the first wave. Although we cannot fit the final erratic mortality waves with high accuracy, our estimate of the total number of deaths is still a very good approximation. As can be seen in Table 2.3, the model's simulated cumulative death toll matches well with the observed death toll.

Table 2.2: Parameter Summary for two scenarios. $X.0$ denotes $X(t = 0)$, which is the number individuals in X class at the beginning of the study period.

Parameters	Notation	weak	strong	Type
mosquito biting rate	a (per day)	0.5	0.5	fixed
transmission probability from vector to host	b	0.4	0.4	fixed
transmission probability from host to vector	c	0.5	0.5	fixed
host latent period	σ_h^{-1} (days)	4	4	fixed
host infectious period	γ_h^{-1} (days)	4	4	fixed
toxic case duration	κ_h^{-1} (days)	8	8	fixed
vector latent period	σ_v^{-1} (days)	10	10	fixed
vector lifespan	μ_v^{-1} (days)	20	20	fixed
severe case proportion	δ	0.15	0.15	fixed
non-severe case relative infectivity	ψ	0.1	0.5	fixed
severe case CFR	θ	0.06	0.06	fixed
number of nodes	n_m	7	7	estimated
severe case reporting ratio	ρ	0.71	0.72	estimated
mean $m(t)$	$\langle m(t) \rangle$	6.34	2.79	estimated
over-dispersion	τ	0.002	0.0045	estimated
initial susceptible host	$S_h.0/N_h$	0.62	0.62	fixed
initial exposed host	$E_h.0/N_h$	3e-07	3.2e-07	estimated
initial non-severe host	$A_h.0/N_h$	3e-07	3.2e-07	estimated
initial severe host	$I_h.0/N_h$	3e-07	3.2e-07	estimated
initial toxic host	$T_h.0/N_h$	3e-07	3.2e-07	estimated
initial recovered host	$R_h.0/N_h$	0.38	0.38	fixed
initial susceptible mosquito	$S_v.0/N_h$	11.66	3.62	estimated
initial exposed mosquito	$E_v.0/N_h$	1.66e-06	1.93e-06	estimated
initial infectious mosquito	$I_v.0/N_h$	1.66e-06	1.93e-06	estimated
mean basic reproductive number	$\langle \mathcal{R}_0 \rangle$	3.41	2.57	estimated
infection attack rate (%)	AR	0.09	0.15	estimated
maximum log likelihood	MLL	-166.48	-165.4	estimated
Bayesian Information Criterion	BIC	388.9	386.75	estimated

Parameter estimates including the basic reproduction number $\mathcal{R}_0$, mean mosquito-to-human ratio $\langle m(t) \rangle$, and disease attack rate are listed in Table 2.2. Our estimated mean $\mathcal{R}_0 \approx 3.0$ with excursions to $\mathcal{R}_0 \approx 6.0$ matches well other studies in the literature (see Introduction). The very low attack rate is an outcome of the prompt and effective control measures of the Angolan government [8, 45, 238].

The estimated mean mosquito-to-human ratio, $\langle m(t) \rangle$, is in line with previous work by Gao *et al.* (see [110]). The estimated reporting ratio for severe cases, $\rho = 70\%$, is reasonable, given the easily recognizable symptoms (jaundice), and the control effort by the government, which managed to push the vaccination coverage to more than 90% of its population within a very short period of time.

Association between the spread of YF and local climatic factors has been discussed frequently in previous studies [154, 194]. As such, we explored the possibility that local temperature and rainfall are potential factors that consistently influence the long-term transmission dynamics. However, the temperature was found to have no significant effect while rainfall only had a minor long-term effect. A detailed discussion is given in the section 2.5. The possible reasons could be: i) local precipitation is relatively minor during the study period (but concentrated in March), and the weather is continuously hot and dry, and ii) duration of the outbreak is short and other factors (control measures and human reaction) played a more prominent role. Nevertheless, in what follows, we consider the possibility that the sporadic March rainfall patterns played an important role.

Estimating the Confidence Interval of $\mathcal{R}_0(t)$

Figs 2.4 and 2.5 show the 95% confidence interval (CI) for the estimated $\mathcal{R}_0$, obtained by calculating the 95% CI of $m(t)$, which is obtained from calculating the profile maximum log likelihood of the model as a function of value of each node of $m(t)$. The width of the CI became very wide in the last 2/7 (i.e., 28.57%) of the study period because both case numbers and the number of deaths became relatively small and noisy.

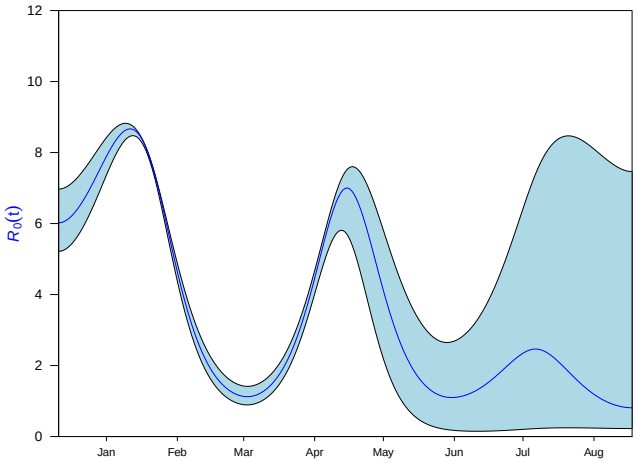


Figure 2.4: The Confidence Interval (C.I.) estimation plot of $\mathcal{R}_0$ under scenario 1.

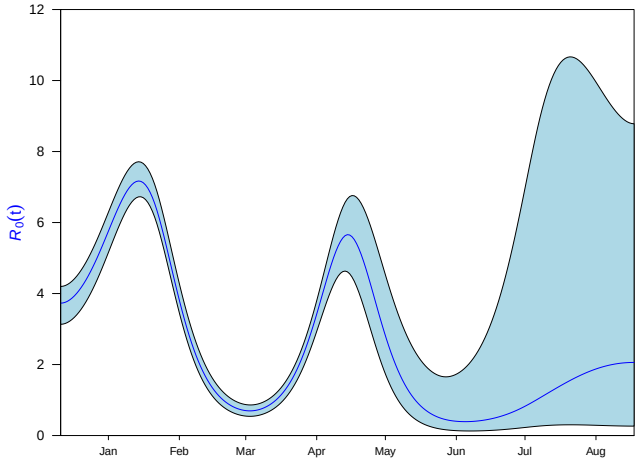


Figure 2.5: The Confidence Interval (C.I.) estimation plot of $\mathcal{R}_0$ under scenario 2.

Oscillations in $\mathcal{R}_0(t)$

For the two scenarios, the analysis revealed that $\mathcal{R}_0(t)$ oscillated in the interval [0.8,8.5] over the 37 weeks with mean $\mathcal{R}_0 \approx 3.4$ of weak infectivity, and $\mathcal{R}_0 \approx 2.6$ for strong infectivity. The estimated $\mathcal{R}_0$ (see blue dashed line in Fig 2.3) is approximately from 5.0-8.5 (see CI estimate of $\mathcal{R}_0(t)$ on section 2.3.1), which is in line with Wu *et al.*'s [151, 240] analysis of the early stage of the epidemic ($\mathcal{R}_0$ is from 5.2-7.1).

We now show that the oscillation in $\mathcal{R}_0(t)$ is a robust feature of the dataset based on Fine and Clarkson's [101] effective and well-known methodology. Fine and Clarkson plotted $Q(t) = C_{t+1}/C_t$ (i.e., next weeks reported cases divided by this weeks – as used in many classic studies of childhood infectious diseases), as a function of time t. This index is usually regarded as the transmissibility of human-to-human infectious diseases but can be similarly exploited for vector-borne diseases. In Fig 2.6 we find strong waves in $Q(t)$ (solid green) that match the oscillations observed in $\mathcal{R}_0(t)$ (dashed blue). This suggests that the waves observed in $\mathcal{R}_0$ are most likely a feature embedded in the data and are not artificial.

Unusually, the wave-like pattern of $\mathcal{R}_0$ in time is out-of-phase with the weekly reported death rate, as seen in Fig. 2.3. Thus larger number of deaths are associated with periods of low disease transmission (i.e., low $\mathcal{R}_0$), and vice-versa. This is biologically reasonable, and could well be due to protective behavioral changes such as usage of insecticide, mosquito repellent, movement restrictions, cordon sanitaire and general vector control. Thus, we hypothesize that the behavioral response (e.g., usage of mosquito insecticides, repellents and nets) to cases and deaths witnessed by Luanda's population, reduced disease transmission

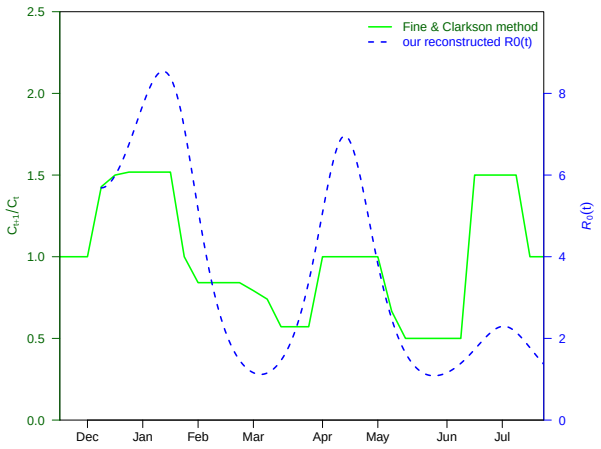


Figure 2.6: Reconstructed transmission rate $Q(t)$ via Fine & Clakrson's method (with a three-week shift to the left) and our estimated $\mathcal{R}_0(t)$.

in periods of high mortality. Such a process would drive the YF case numbers to follow dynamics that differ substantially from the standard SIR epidemic curve, and possibly even induce waves. A similar phenomenon was reported for the 1918 influenza pandemic [66, 128].

In practice, the vector control efforts (i.e. "mosquito-fogging") was case-driven, and fogging was implemented in localities where cases and mortality were reported, it is not unreasonable to link mortality to the mosquito-abundance and the transmission rate (see a discussion of this in section 2.6). In section 2.3.3, we describe a simple model simulation in which $m(t)$ is controlled by death rates only, and it is possible to fit the YF datasets just as convincingly as our earlier fits where $m(t)$ is a free cubic spline best fitting function.

A closer examination of Fig. 2.3 helps explain the oscillations in $\mathcal{R}_0(t)$ and their implications. A few weeks after the initiation of the vaccination program in February 2016, the YF epidemic is seen to curtail its rapid growth and then diminish over March. Unusually, the epidemic does not rapidly crash to extinction as the usual (SIR) epidemic curve would predict. Rather, YF cases reduce gently over the next five months in a plateau from April to September. This behavior is uncharacteristic of the standard SIR epidemic model, and appears to be an outcome of the oscillations observed in the time-series estimates of $\mathcal{R}_0$. Thus the epidemic begins to rapidly decline and turns-around only after $\mathcal{R}_0$ has reduced considerably (i.e., after the peak of $\mathcal{R}_0$ in January). A similar decreasing trend in $\mathcal{R}_0$ over Feb-Mar 2016 was also noted by [240]. This implies that although vaccination was important for the main epidemic's demise, the February decline in $\mathcal{R}_0$ is also likely to have played an important role. Interestingly, the sustained period of cases in April (rather than the expected

drop to disease extinction) occurs when $\mathcal{R}_0$ increases again. Since theoretically, $\mathcal{R}_0$ should remain unaffected by vaccination, we suppose the changes in $\mathcal{R}_0$ are likely due to changes in other factors (eg., control measures and behavior, or possibly weather patterns) that may influence the vector population and its transmission.

Generation Time The generation time (GT, i.e., the time between two successive infections) equals the sum of latent and infectious periods of the host [213], and the sum of latent and infectious periods of the mosquito. However, mosquitoes have a short lifespan and die before the loss-of-infectiousness of YF. Precisely, the life of adult (female) mosquito includes three parts: pre-infection, latent and infectious period. In this chapter, we fix the mean lifespan (L_v) of a mosquito to be 20 days, and the mean latent period to be 10 days. For the pre-infection period, we use the formula of mean age at infection (for typical childhood infections): $L_v/(1+\mathcal{R}_0)$ [143]. If we fix $\mathcal{R}_0 = 3.0$ (see Table 2.2), the mean age at infection of mosquito is 5 days. Thus, the sum of latent and infection periods of mosquito will be $20 - 5 = 15$ days. The sum of latent and infectious periods of human is $4 + 4 = 8$ days. Thus, the GT of YF is 23 days, which explains the time delay between the maxima of $\mathcal{R}_0(t)$ and the maxima of reported weekly cases. GT will be between 18 and 28 days.

2.3.2 Potential causes of the second wave of $\mathcal{R}_0(t)$

For most years in Luanda, the rainy season is between November-May but the most accumulation of rain occurs in March-April. The year 2016 was an El Niño year and it brought dramatic and unpredictable flooding events especially in the March-April period, thereby leading to conditions ideal for the growth of mosquito populations. As in the 1971 YF outbreak in Luanda, local water-storage containers (mainly the larger ones) serving the community but also in most homes, accounted for 85% of the *Ae. aegypti* larval breeding. As vividly described by Moreira [178] that "*The 2016 outbreak coincided with unusually heavy rains and a severe El Niño weather pattern. We are also suffering from an economic crisis and poor sanitary conditions. All these factors created a fertile environment for an increase in the mosquito population. The outbreak reached its peak in February and has been declining since (i.e., population numbers, not $\mathcal{R}_0$). We have much more vaccine now (in September 2016) than we had earlier in the epidemic. The response interventions are*

involving communities successfully. The dry season arrived in May (2016) and since then the mosquito population has diminished."

Thus after the peak of the YF outbreak had passed, and the vaccination program was in progress, the local March-April El Niño rains were enhancing mosquito breeding conditions. It is not unsurprising that simultaneously one of Luanda's largest malaria epidemics ever was underway (see "*During the first quarter of 2016, the number of cases of malaria increased dramatically to 1,531,629, up from 980,192*" of [34]). We suggest that these conditions were responsible for the unusual and robust second wave that is observed in the time series of $\mathcal{R}_0(t)$.

Since we only model a single province, Luanda, and the YF transmission spread relatively rapidly throughout the province, we might assume the effects of spatial heterogeneity are likely to be minimal. However, we do not possess sufficiently detailed data to perform a careful analysis of spatial effects, and effects at the micro-scale may be important as in other diseases such as dengue. Kraemer *et al.* [151] have discussed the importance of spatial effects for YF over all provinces of Angola, and there is a possibility that geographic waves generated from surrounding provinces could play some part in the appearance of multiple YF waves in Luanda province. Hence future work and more comprehensive data are needed to examine these possibilities.

2.3.3 Underlying Oscillation in Basic Reproductive Number

Import of New Susceptibles

In order to explain the multiple waves in $\mathcal{R}_0$, one might speculate that there were additional inflows of new susceptibles in the later stage of the epidemic. But we show that even if there are, the estimated $\mathcal{R}_0$ still exhibits features of multiple waves. We examined the addition to the system of a continuous inflow (daily 10,000 after Apr 09) or a square-wave inflow (daily 20,000 for the period Mar 30 - Apr 29 and Jun 08 - Jul 07) of imported host susceptibles (as well as cross-border infections from confirmed cases outside Luanda province). These input of new susceptibles resulted in little difference to the baseline scenario (see Fig. 2.7. Parameters of scenario 1 were used here).

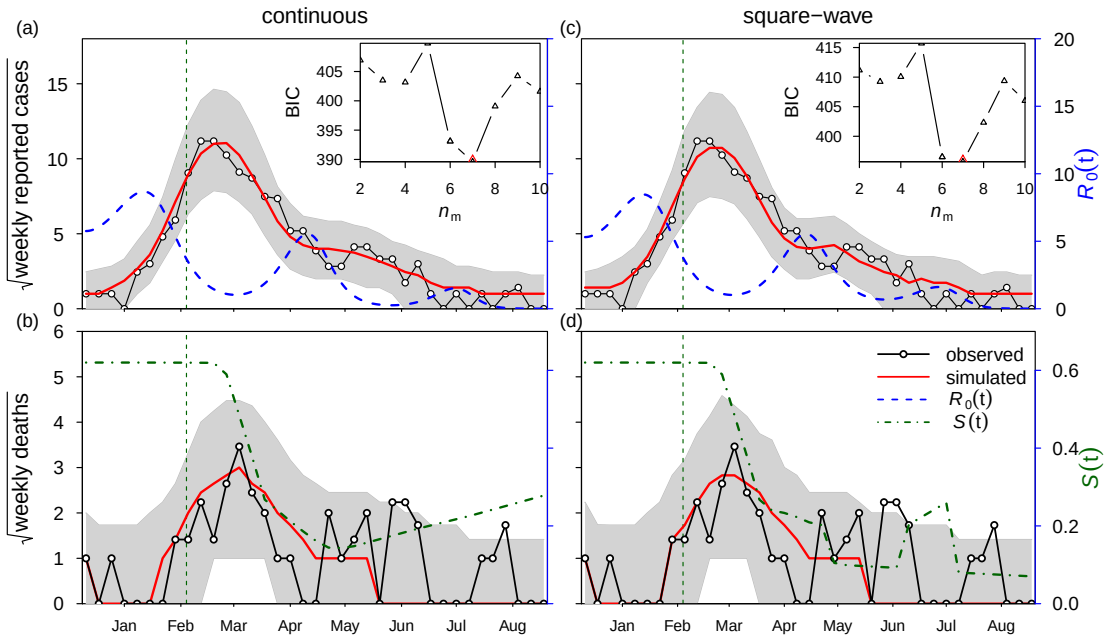


Figure 2.7: Fitting results for model including a continuous (a,b) or square-wave(c,d) inflow of new susceptibles. Weekly reported cases (a,c) and deaths cases (b,d) in Luanda province, Angola. Black line with circles denotes reported cases, and red line denotes model simulation median, blue dashed line denotes the fitted basic reproduction number, $\mathcal{R}_0$, and green dashed line shows the calculated host susceptible proportion, $S(t)$. Shaded region represents 95% bound of 1000 simulations. The vertical dashed line indicates the starting date of the vaccination campaign. Inset panel shows the BIC as a function of the number of nodes (n_m). The lowest BIC is attained at $n_m = 7$, which is used in the main panel.

Modelling Human reaction to mortality

We have seen that the reproductive number $\mathcal{R}_0(t)$ exhibits multiple waves. An examination of the weekly YF mortality reveals possibly related oscillations. High weekly mortality corresponds to high transmission rate later in the following week. Given that in practice, the vector control (mosquito-fogging) was case-driven (fogging was implemented in localities where cases and mortality were observed), it is not totally unreasonable to link mortality to the mosquito-abundance and the transmission rate (see a discussion of this in section 2.6). In this section, we use a simple model simulation to demonstrate this possibility.

We replace the cubic spline function (with 7 nodes) for $m(t)$ (Eqn. (4.2)) with a simple function based on the YF mortality and obtain almost identical results. We set:

$$m(t) = m_{\text{base}} + k \cdot \exp\Big[- D_h(t - t_{\text{lag}})\Big] \tag{2.11}$$

Here m_{base} is a constant term, k is a parameter controlling the strength of the death-induced human reaction, $D_h(t)$ is the yellow fever deaths of week t and t_{lag} is the lag time for the population response in reaction to mortality levels. The fitting results for this simple human behavior model are shown in Figure 2.8, with $t_{\text{lag}} = 1$ week fixed. We see that a simple behavior model achieved very similar fitness as our original model in Fig. 2.3. Thus we illustrated that death-driven oscillation in $\mathcal{R}_0$ is possible.

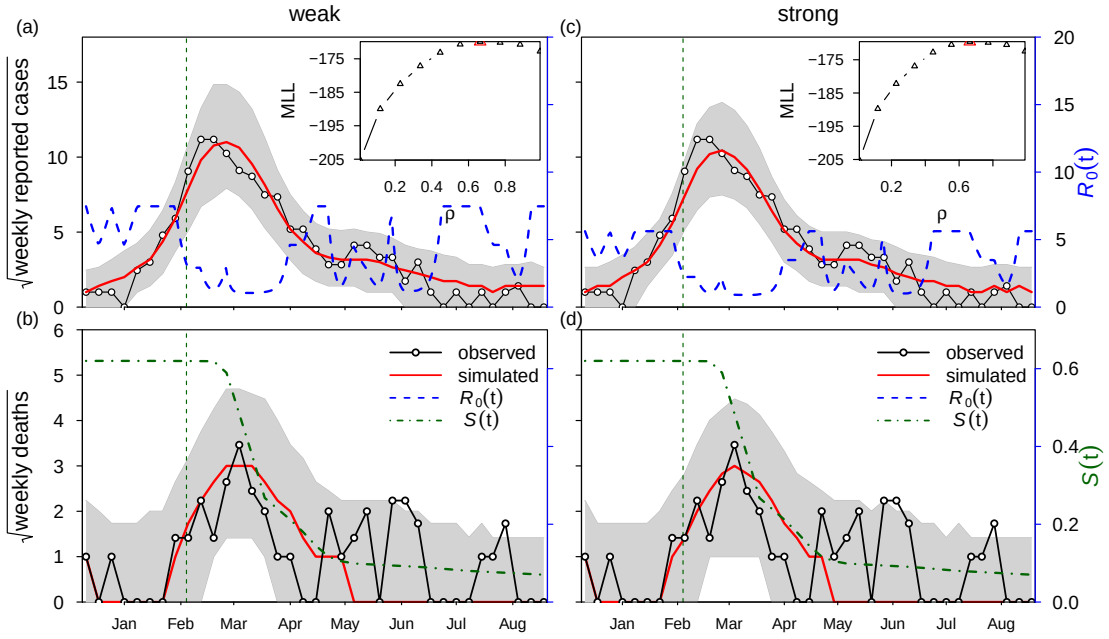


Figure 2.8: Fitting results implementing humans reaction to mortality (with a time lag of one week, $t_{\text{lag}} = 1$), scenario 1 (a,b) and scenario 2 (c,d). The black line with circles denotes reported cases, and red line denotes model simulation median, blue dashed line denotes the fitted basic reproduction number, $\mathcal{R}_0$, and green dashed line shows the calculated host susceptible proportion, $S(t)$. The shaded region represents 95% bound of 1000 simulations. The vertical dashed line indicates the start date of the vaccination campaign. Inset panel shows the MLL as a function of the mild YF infections counting ratio (ρ). The highest MLL is used in the main panel.

2.3.4 Vaccine usage

As discussed in the main text, we examined scenarios when an identical vaccination scheme to that observed in Luanda was delayed by 60 and 120 days. Here we show the number of doses given in these delayed schemes is almost identical to the baseline scheme varying at most by 5% (for delay of 120 days; 3.44 million doses). The 180-day delay is included because, as explained in the text, it is so late that it is effectively a no-vaccination scenario (see below).

The calculations are based on Luanda having a population size of 6,543,000, with 38% of the population having immunity before the epidemic (November 2015). The baseline scenario characterized what happened in practice between December 2015 and August 2016. By August 2016, 93% of the population was vaccinated according to the WHO reported baseline scenarios: $(0.93 - 0.38) \times 6543000 = 3,598,650$ person doses used.

When the vaccination scheme was delayed by 60 days, in August 2016, 91% of the population was vaccinated. Thus,

Delay by 60 days: $(0.91 - 0.38) \times 6543000 = 3,467,790$ person doses used.

Similarly, when the vaccination scheme was delayed by 120 days, in August 2016, 90.5% of the population was vaccinated. Thus,

Delay by 120 days: $(0.905 - 0.38) \times 6543000 = 3,435,075$ person doses used.

Delay by 180 days: $(0.54 - 0.38) \times 6543000 = 1,046,880$ person doses used.

The total reported observed cases, as well as total deaths, are evaluated by the model for each different vaccination scenario.

A 180-day delay period is considered a no-vaccination scenario. When taking into account the extra 20 days required for the vaccine to be effective, anyone vaccinated 180 days after the December 2015, will not gain any protective effect from the vaccination given the observation study period is only 200 days. The impact of the vaccination goes beyond the study period. Anyone vaccinated in the last twenty days will change their status from susceptible to recovered after 15 August 2016 (i.e., the end of the study period).

2.3.5 Vaccination and Vector Control

There are many possible ways to evaluate the effects of a delayed vaccination campaign when compared to the baseline scenario that was implemented in practice in Luanda. The approach followed here is to simply delay the exact same baseline scenario (in terms of doses per week) by a fixed time interval until the end of the observation period arrives. It is difficult to extend beyond the observation period without introducing an unacceptable rate of errors. This can be seen in the large confidence intervals for $\mathcal{R}_0$ towards the end of the observation period (see section 2.3.1).

The results of 60, 120, and 180 days delay of the vaccination campaign for the 2016 yellow fever outbreak are presented in Fig 2.9 for the scenario 1 ($\psi = 0.1$). The total reported cases and total deaths are calculated for four vaccination scenarios (including the baseline) and outcomes are listed in Table 2.3. The 180-day delay is included because it gives an impression of what might happen when vaccination is unavailable, as mentioned.

The baseline scenario (actual vaccination or 0-delay) results in an estimated 73 deaths associated with YF in the study period, which matches the observed number. With a 60-day delay to the vaccination roll-out, YF deaths saved were 2.2-fold of the observed number (see Fig. 2.9 (a,b)). With a 120-day delay, the YF death saved were 4.5-fold of the observed number (see Fig. 2.9 (c,d)). With a 180-day delay, YF deaths saved were 5.1-fold of the observed number (see Fig. 2.9 (e,f)). The latter result is a good approximation to what might have occurred if there were no vaccination campaign in Luanda up to August 2016. All of these results show that delaying the vaccination campaign would have greatly enhanced the epidemic in terms of infectious cases and mortality. We also investigated the "vaccination delay" situation under different scenarios (see sections 2.4.2 and 2.4.3), and found our main results of "deaths prevented" largely holds. In addition, we also considered the scenario of "what if deaths were underreported" (i.e., there was a constant proportion of YF deaths not reported, see sections 2.4.4 for details), we report our main results are also robust.

A clear feature of the simulated outcomes with delayed vaccination (red lines in Fig. 2.9), is the noticeable second wave of YF cases and deaths that appear. This feature becomes even more prominent in a situation of no vaccination (see Fig. 2.9e and 2.9f). Returning to Figure 2.1, we also see strong signs from the observed time series of YF in Luanda, that the outbreak may have indeed occurred over two waves. Hence, even with Luanda's large-scale vaccination campaign, the multiple-wave feature is noticeable in the observed time series, which implies considerable fluctuations in the driving force ($\mathcal{R}_0$) or other factors.

The multiple-waves may also be in part due to the interference of multiple disease control strategies. The control efforts on mosquito eradication increased as the number of YF cases and deaths increased. After the national vaccination campaign began in the first week of February, 2016, and reported death decreased, the mosquito eradication effort may have also slowed down (possibly due to limited funding and resources). Additionally, human reaction in responding to YF deaths (i.e., "recent death-driven" human reaction) could also be involved in this complex wave-like behavior, as studied in the context of other infectious

diseases [66, 128], and discussed further in 2.6.1.

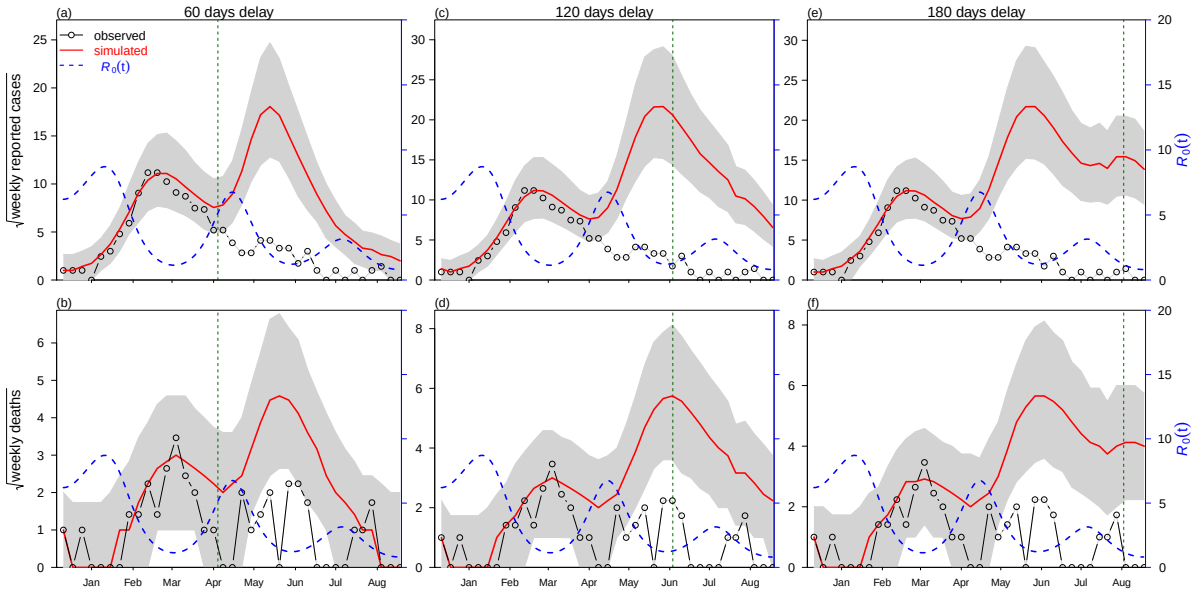


Figure 2.9: Simulation results of scenario 1 under three deferred vaccination campaign scenarios: 60 day delay in panels (a,b), 120 day delay in panels (c, d) and 180 day delay in panels (e,f). Black line with circles denote reported cases (eg., in (a), in the form of square-root) and reported deaths (in (b), in the form of square-root), red line denotes model simulation median and blue dashed line is the fitted basic reproduction number, $\mathcal{R}_0$. Shaded region represents 95% range of 1,000 simulations. The vertical dashed line represents initiation of the vaccination campaign. The number of nodes, $n_m = 7$, is adopted.

Table 2.3: Impacts of vaccination campaign delay under scenario 1: weak infectivity.

Scenario	Total reported cases	Total deaths
Observed	941	73
Baseline model	1026 [540 , 1797]	77 [35 , 139]
60 days delay	3143 [1604 , 5584]	233 [119 , 411]
120 days delay	5450 [2751 , 9611]	400 [203 , 724]
180 days delay	6242 [3139 , 10919]	444 [226 , 787]

2.3.6 Sensitivity Analysis

Results of sensitivity analysis are presented in Fig. 2.10 and indicate how model parameters impact the basic reproduction number $\mathcal{R}_0$ and the death toll. $\mathcal{R}_0$ is most sensitive to vector biting rate (a) and the vectors' lifespan (μ_v^{-1}), indicating the importance of the mosquitoes role in disease transmission. The total deaths are considerably sensitive to the proportion of severe cases (δ), the case-fatality rate of severe cases (θ) and the initial number of susceptibles (i.e. the ratio $S_h.0/N_h$).

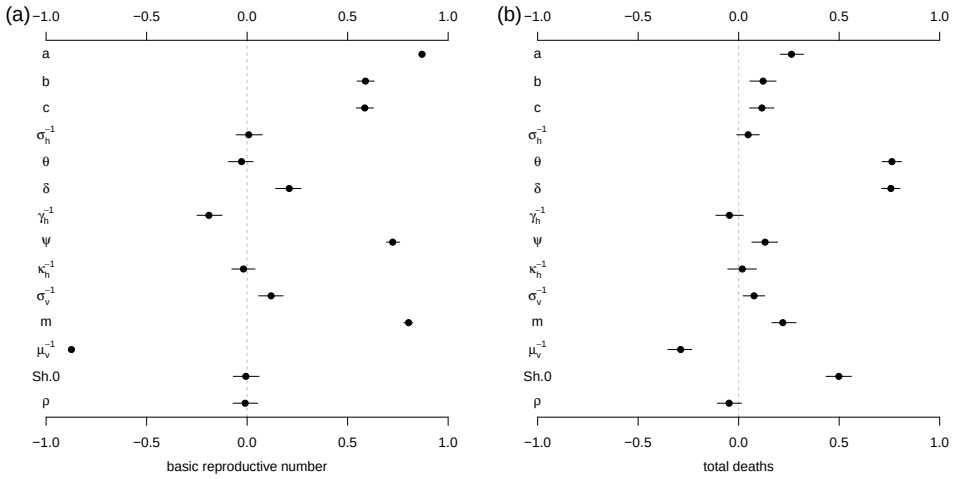


Figure 2.10: The Partial Rank Correlation Coefficients (PRCCs) of basic reproduction number (panel (a)) and total deaths (panel (b)) with respect to model parameters. $S_h.0$ denotes the initial susceptible ratio ($S_h.0/N_h$). The black circle is the estimated correlation and the bar represents the 95% C.I.. The ranges of parameters are given in Table 3.1.

2.4 Different model scenarios

2.4.1 Confidence Interval of ρ

Fig. 2.11 shows the maximum log likelihood as a function of reporting ratio ρ. The 95% confidence interval of ρ is (0.52, 0.95) for weak infectivity and (0.51, 0.99) for strong infectivity.

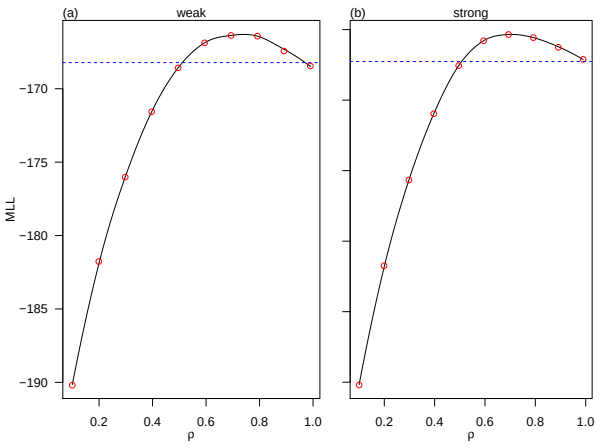


Figure 2.11: Maximum log likelihood (MLL) as a function of the reporting ratio ρ. Blue dashed lines are the threshold of 95% significance level ($\frac{1}{2}\chi^2_{1;95\%}$ below the max. of the MLL function). Panel (a) is under weak infectivity and panel (b) strong infectivity.

2.4.2 Scenario 2: strong infectivity $\psi = 0.5$

Fitting results of scenario 2 (strong infectivity) are shown in Fig. 2.3 (c,d). Parameter estimates are listed in Table 2.2. The estimated basic reproduction number, $\mathcal{R}_0$ (see blue dashed line in Fig. 2.12, rises to over 8.5 at the first peak and exhibits two major waves during the study period. The estimated reporting ratio for severe YF cases is similar to the weak infectivity scenario (see Table 2.2).

The results of 60, 120 and 180 days delay of the vaccination campaign are presented in Fig. 2.12 and Table 2.4.

Table 2.4: Impacts of vaccination campaign delay under strong infectivity scenario.

Scenario	Total reported cases	Total deaths
Observed	941	73
Baseline model	984 [540 , 1537]	80 [41 , 131]
60 days delay	3026 [1651 , 4717]	238 [128 , 378]
120 days delay	4522 [2465 , 7219]	354 [194 , 568]
180 days delay	4682 [2541 , 7312]	362 [199 , 587]

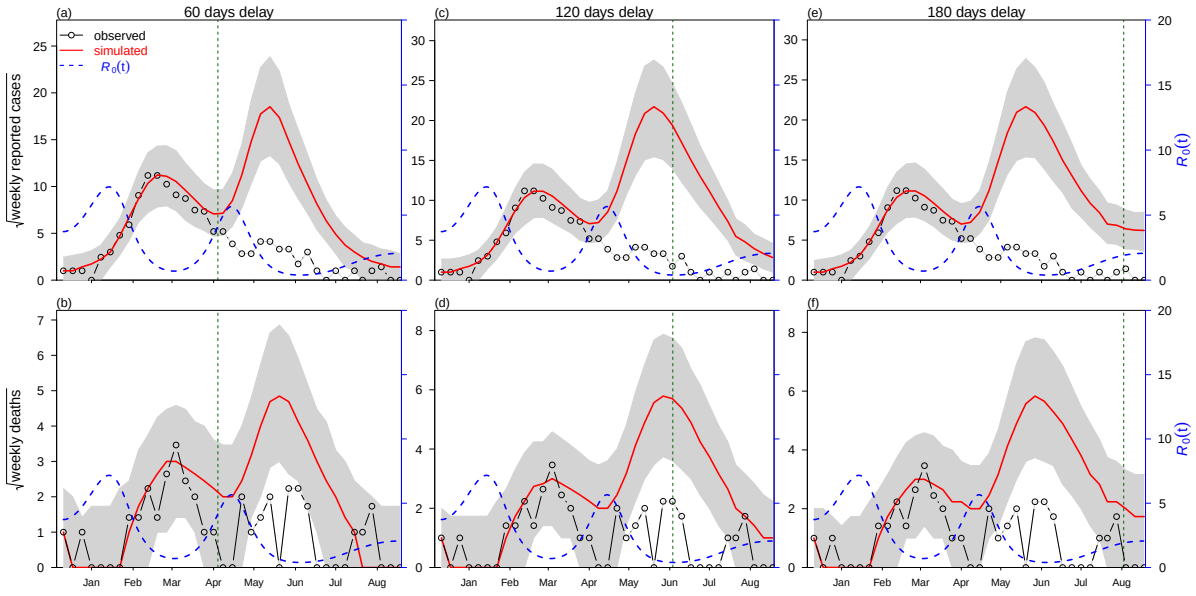


Figure 2.12: Simulation results (strong infectivity) under deferred vaccination campaign for 60 days in panels (a,b), 120 days in panels (c, d) and 180 days in panels (e,f). The observed cases (in the form of square-root) given by black circles, model simulation median (in the form of square-root) is in red and the fitted basic reproduction number, $\mathcal{R}_0$, is blue dashed line. Shaded region represents 95% range of 1,000 simulations. The vertical dashed line is the time point when the vaccine campaign started. The number of nodes, $n_m = 7$, is adopted.

For $\psi = 0.95$, we also examined the outcome when non-severe case relative infectivity was extremely large (or strong), namely $\psi = 0.95$. The results so obtained were very similar

to results for $\psi = 0.5$. The mean $\mathcal{R}_0$ reduced to 2.47 from 2.57, and the infection attack rate increased to 0.13% from 0.12%. The deaths prevented was almost the same.

2.4.3 Fitting CFR

Here we give results for fitting CFR, rather than fixing CFR to the constant $\theta = 0.06$. The model simulation results are summarized in Table 2.5, Table 2.6, Table 2.7, and Figs. 2.13, 2.14 and 2.15.

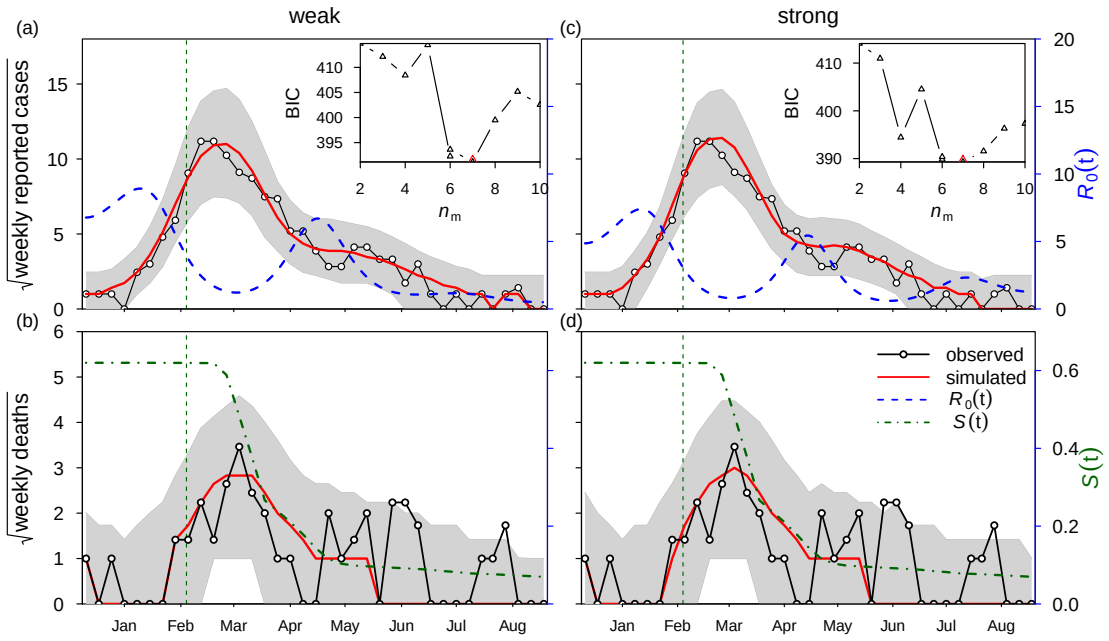


Figure 2.13: Model fitting results for fitting CFR with weak infectivity (a,b) and strong infectivity (c,d) of non-severe cases. Panels' information and color code are same as in Fig. 2.3. Parameter values are listed in Table 2.5.

Weak infectivity scenario With a 60-day delay to the vaccination roll-out, YF deaths saved were 2.3-fold of the observed number (see Table 2.6). With a 120-day delay, the YF death saved were 4.4-fold of the observed number (see Table 2.6). With a 180-day delay, YF deaths saved were 4.7-fold of the observed number (see Table 2.6).

Strong infectivity scenario With a 60-day delay to the vaccination roll-out, YF deaths saved were 2.3-fold of the observed number (see Table 2.7). With a 120-day delay, the YF

Table 2.5: Parameters summary for fitting CFR. $X.0$ denotes $X(t=0)$, which is the number individuals in X class at the beginning of the study period.

Parameters	Notation	weak	strong	Type
mosquito biting rate	a (per day)	0.5	0.5	fixed
transmission probability from vector to host	b	0.4	0.4	fixed
transmission probability from host to vector	c	0.5	0.5	fixed
host latent period	σ_h^{-1} (days)	4	4	fixed
host infectious period	γ_h^{-1} (days)	4	4	fixed
toxic case duration	κ_h^{-1} (days)	8	8	fixed
vector latent period	σ_v^{-1} (days)	10	10	fixed
vector lifespan	μ_v^{-1} (days)	20	20	fixed
severe case proportion	δ	0.15	0.15	fixed
non-severe case relative infectivity	ψ	0.1	0.5	fixed
severe case CFR	θ	0.04	0.04	estimated
number of nodes	n_m	7	7	estimated
severe case reporting ratio	ρ	0.51	0.46	estimated
mean $m(t)$	$\langle m(t) \rangle$	6.68	3.12	estimated
over dispersion	τ	0.0047	0.0037	estimated
initial susceptible host	$S_h.0/N_h$	0.62	0.62	fixed
initial exposed host	$E_h.0/N_h$	2.9e-07	5e-07	estimated
initial non-severe host	$A_h.0/N_h$	2.9e-07	5e-07	estimated
initial severe host	$I_h.0/N_h$	2.9e-07	5e-07	estimated
initial toxic host	$T_h.0/N_h$	2.9e-07	5e-07	estimated
initial recovered host	$R_h.0/N_h$	0.38	0.38	fixed
initial susceptible mosquito	$S_v.0/N_h$	16.26	6.06	estimated
initial exposed mosquito	$E_v.0/N_h$	1.24e-06	1.33e-06	estimated
initial infectious mosquito	$I_v.0/N_h$	1.24e-06	1.33e-06	estimated
mean basic reproductive number	$\langle \mathcal{R}_0 \rangle$	3.37	2.82	estimated
infection attack rate (%)	AR	0.16	0.12	estimated
maximum log likelihood	MLL	-165.44	-164.52	estimated
Bayesian Information Criterion	BIC	391.14	389.3	estimated

Table 2.6: Impacts of vaccination campaign delay (fitting CFR) with weak infectivity of non-severe cases.

Scenario	Total reported cases	Total deaths
Observed	941	73
Baseline model	991 [535 , 1659]	76 [36 , 138]
60 days delay	3142 [1698 , 5255]	240 [126 , 403]
120 days delay	5287 [2839 , 8885]	396 [201 , 676]
180 days delay	5514 [2974 , 9170]	413 [223 , 709]

death saved were 4.1-fold of the observed number (see Table 2.7). With a 180-day delay, YF deaths saved were 4.4-fold of the observed number (see Table 2.7).

There were differences in estimates of death prevented, but after considering the width

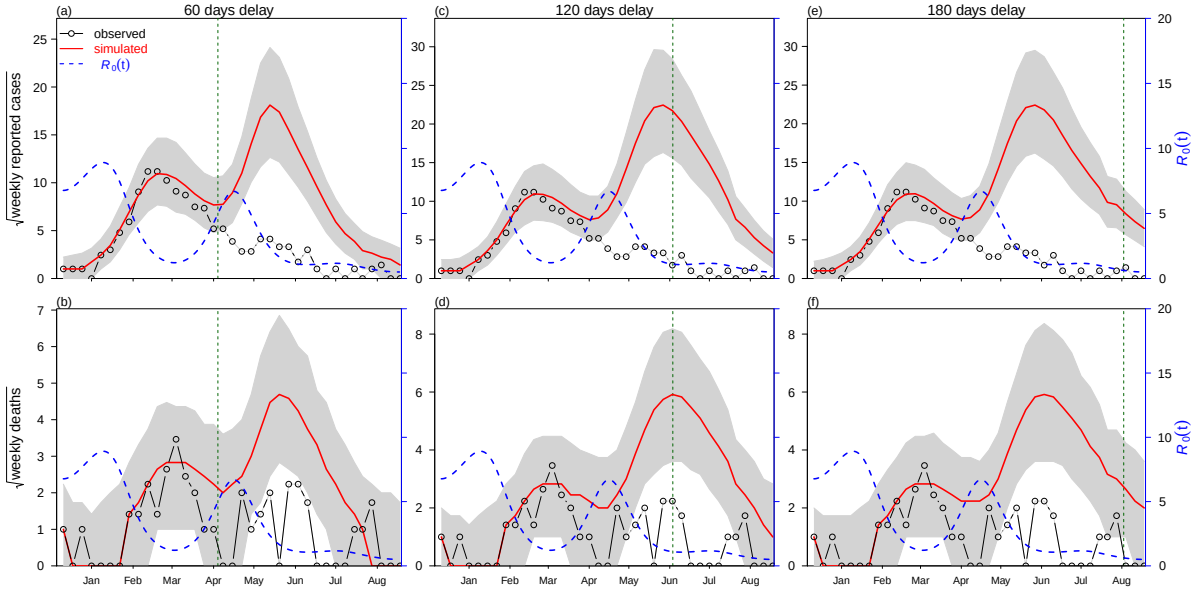


Figure 2.14: Simulation results for fitting CFR (weak infectivity) under delayed vaccination campaign for 60 days in panels (a,b), 120 days in panels (c, d) and 180 days in panels (e,f) delay. The case is black dotted line, model simulation median is in red and the fitted basic reproduction number, $\mathcal{R}_0$, is blue dashed line. Shaded region represents 95% range of 1,000 simulations. The vertical dashed line is the time point when the vaccination campaign started. Number of nodes, $n_m = 7$.

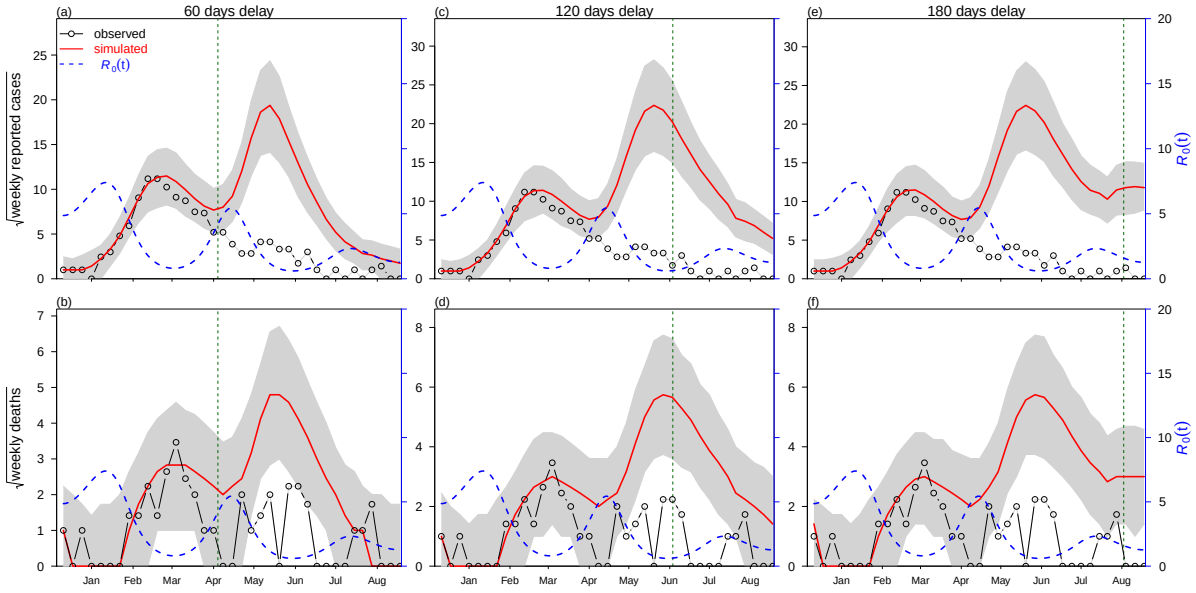


Figure 2.15: Simulation results for fitting CFR (strong infectivity) under delayed vaccination campaign for 60 days in panels (a,b), 120 days in panels (c, d) and 180 days in panels (e,f) delay. The case is black dotted line, model simulation median is in red and the fitted basic reproduction number, $\mathcal{R}_0$, is blue dashed line. Shaded region represents 95% range of 1,000 simulations. The vertical dashed line is the time point when the vaccination campaign started. Number of nodes, $n_m = 7$.

Table 2.7: Impacts of vaccination campaign delay (fitting CFR) with strong infectivity of non-severe cases.

Scenario	Total reported cases	Total deaths
Observed	941	73
Baseline model	1046 [593 , 1582]	78 [40 , 122]
60 days delay	3290 [1859 , 5030]	238 [133 , 373]
120 days delay	5162 [2910 , 7837]	372 [211 , 580]
180 days delay	5638 [3205 , 8621]	397 [224 , 602]

of the 95% CI, the differences were small.

2.4.4 Assuming deaths were under-reported (CFR=15%)

To further test the robustness of our conclusion, we assume that deaths were under-reported. We assume the "actual" YF-related deaths were double the confirmed deaths, so that the "actual" death toll is $2 \times 73 = 146$. We then set the CFR = 15%, which is in line with previous studies in Africa (see "*the historic estimate of 200,000 cases and 30,000 deaths annually, which was based on serological survey data obtained from children in Nigeria between 1945 and 1971*" of Ref. [115]). Results are summarized in Fig. 2.16 and Tables 2.8 and 2.9. Our key conclusions still hold, namely the deaths prevented are approximately 5,6-fold times the reported "actual" death toll.

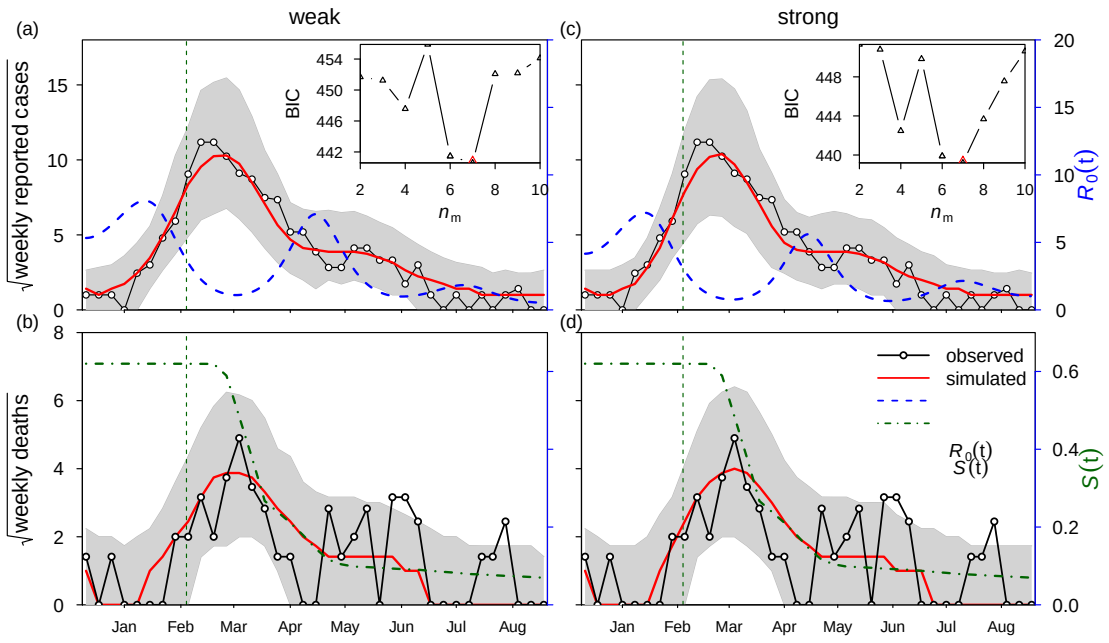


Figure 2.16: Fitting a model with mocked deaths data and CFR = 15%.

Table 2.8: Impacts of vaccination campaign delay under weak infectivity scenario and CFR = 15%.

Scenario	Total reported cases	Total deaths
Observed	941	146
Baseline model	938 [486 , 1669]	149 [76 , 264]
60 days delay	3425 [1736 , 6168]	538 [263 , 962]
120 days delay	5906 [2829 , 10529]	917 [450 , 1678]
180 days delay	6266 [3104 , 11220]	957 [481 , 1684]

Table 2.9: Impacts of vaccination campaign delay under strong infectivity scenario and CFR = 15%.

Scenario	Total reported cases	Total deaths
Observed	941	146
Baseline model	896 [431 , 1611]	148 [69 , 271]
60 days delay	3325 [1579 , 5846]	538 [250 , 984]
120 days delay	5508 [2676 , 9995]	910 [443 , 1618]
180 days delay	6146 [2932 , 11034]	978 [450 , 1779]

2.4.5 The basic equations - reporting severe cases

Given the definition for YF reported cases corresponds to severe cases (as opposed to all infections), we seek the most natural tapping point to identify them in the equations (see Fig. 2.2 and Eqn. (2.1)). We assume that to be identified as severe, they would have to reside in the infected compartment for some time. As such we have allowed case reporting to be proportional to the rate at which individuals move into the toxic phase ($\gamma \cdot I_h$). Note that similar results would be obtained had we let cases be proportional to arise ($\delta \cdot \sigma \cdot E_h$) but may induce a time shift of several days which is relatively small.

2.5 Impacts of Temperature and Rainfall

Climatic data for Luanda Angola were obtained from Weather Underground Organization (`https://www.wunderground.com`). Fig. 2.17 shows the weekly reported cases and deaths and daily mean air temperature and rainfall (in mm). Over the study period, the air temperature decreased and the rainfall concentrated in a few months. There is no evident association between temperature/rainfall and reported cases/deaths by inspection, except that the rainfall in April/May might be associated with the second wave in reported cases.

Daily temperature and rainfall are obtained for Luanda airport, Angola. We assume

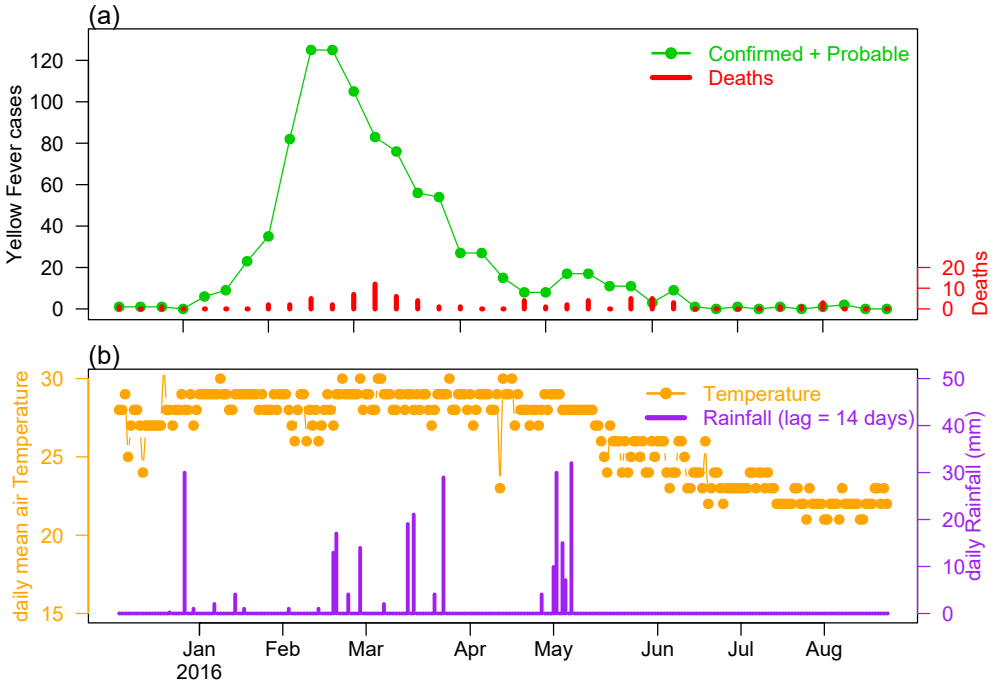


Figure 2.17: Weekly reported cases and deaths in Luanda Angola (a) and daily mean air temperature and rainfall (mm) (b).

the mosquito human population size ratio is:

$$m(t) = m_{\text{base}}(t) + \xi_1 \cdot [\text{TEMP}(t) - \min\{\text{TEMP}\}] + \xi_2 \cdot \text{RAIN}(t - t_d) \tag{2.12}$$

where $m_{\text{base}}(t)$ is an exponential cubic spline function with $n_m = 7$ nodes, TEMP (with minimum value, $\min\{\text{TEMP}\}$) and RAIN are daily temperature (in °C) and rainfall (in mm) respectively, and ξ_1 and ξ_2 are parameters to be estimated. A time delay $t_d = 14$ days is considered for rainfall. In this simulation, we integrated the model with a time-step-size of 1 day. If temperature or rainfall played a significant role, we expect to detect a non-zero ξ_1 or ξ_2 with a significantly improved fitting. But the fitting is not improved significantly (see Fig 2.18 with an estimated $\xi_1 = 0.05$ and $\xi_2 = 0.2$. It appears that rainfall has a slight and probably minor role. Thus we conclude that the effect of climate factors are likely to be weak. Rainfall could have a marginal role.

The results of correlation tests between climatic time series and YF time series (see Table 2.10) shows that none of the correlations are significant.

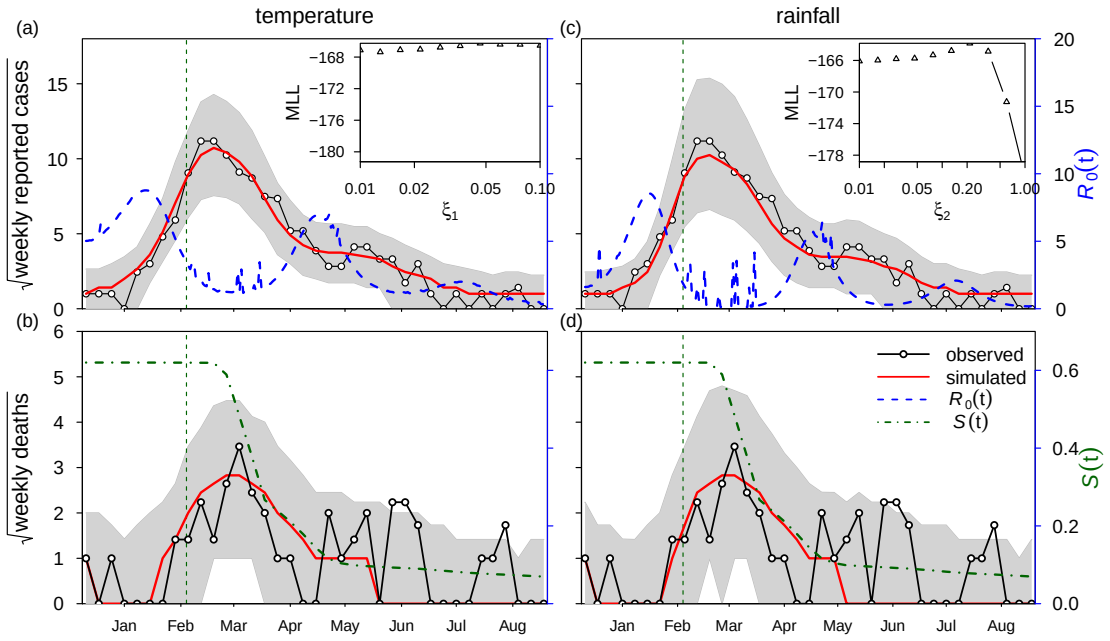


Figure 2.18: Fitting results with climate factors: local temperature (a,b) and rainfall (c,d). Inset panel shows the profile MLL as a function of ξ_1. Black line with circles denotes reported cases, and red line denotes model simulation median, blue dashed line denotes the fitted basic reproduction number, $\mathcal{R}_0$. Inset panel shows the MLL as a function of the temperature parameter (ξ_1). The highest MLL is used in the main panel.

Table 2.10: Correlation results

Correlations	weekly cases	weekly deaths
weekly temperature	-0.1756 (p-value=0.30)	-0.2451 (p-value=0.14)
weekly rainfall	-0.1659 (p-value=0.33)	-0.0740 (p-value=0.66)
weekly rainfall (lagged 14 days)	-0.0290 (p-value=0.86)	-0.1058 (p-value=0.53)

2.6 Impact of vector control and possible impact of climate

We discuss how controlling mosquito eradication programs could also substantially influenced mosquito populations. Note that the government programs on vector control are often driven by reported cases and deaths. When deaths were reported from a local community, the vector control team was sent on location for fogging the streets with chemical spray as a means of destroying mosquitos. Other factors controlling the vector population might also have an impact. Dr. Moreira has outlined that the WHO's response was organized into five different stages and that the reaction and support of the public were varied at different times from strong resistance to cooperation.

Dr. Moreira states: “The government set up a task force to lead the response and launched a five-part plan. The first part was active surveillance. We reinforced laboratory capacity to allow for early detection and notification of new cases.”

“The outbreak coincided with unusually heavy rains and a severe El Niño weather pattern. We are also suffering from an economic crisis and poor sanitary conditions. All these factors created a fertile environment for an increase in the mosquito population. The outbreak reached its peak in February and has been declining since. We have much more vaccine now in September 2016 than we had earlier in the epidemic. The response interventions are involving communities successfully. The dry season arrived in May 2016 and since then the mosquito population has diminished.”

The second cycle The rainy season is between November - May but the most accumulation of rain in Luanda occurs in March - April.

2.6.1 Details of control program

As discussed, major control efforts aimed to control the mosquito vector population *Aedes aegypti* (which transmits the YF virus) were ongoing during the observation period. The Mission Report of the European Centre for Disease Prevention and Control “Assessing the yellow fever outbreak in Angola” [5] outlines the control efforts.

Suspected case: inspection of the case’s residence and the surrounding homes within a 200 m radius; calculation of larval and adult infestation levels.

Confirmed case: indoor fogging of the case’s residence and surrounding homes within a 200 m radius.

Neighborhoods with a high infestation level: truck-fogging with cypermethrine, door-to-door visits to inspect potential breeding sites, Bti application, and informing household members about control measures for *Aedes* control.

In remote areas: fogging of residence and surroundings when a resident returned from Luanda. Information/education/communication messages on household control measures distributed through various channels to achieve community actions.

Bti or cypermethrine are not always available because purchase depends on the municipality's priority settings. Application is hampered by the high cost of fuel which needs to be mixed with the adulticide. The municipal vector control teams are often not aware of the main *Aedes* breeding sites or geographical high-risk pockets. *Aedes* infestation levels are reaching a Breteau index of 50 to 103 in some geographical pockets while in other areas no *Aedes* mosquitoes were present. Only *Aedes aegypti* has been identified, with neither *Aedes albopictus* nor *Aedes africanus* found. All in all, it remains difficult to determine the geographical distribution of the vector, the corresponding infestation levels, and the entomological risk for local transmission.

Vector control/entomology during the epidemic: As confirmation of cases is often only available after several days, it is essential to ensure that fogging is performed as soon as a suspected case is detected; waiting for case confirmation results in losing valuable time.

The vector control measure was driven by reported cases (see the above), and also by reported deaths, but also under limited resources. We, therefore, make the not unreasonable assumption that $\mathcal{R}_0$ may be driven by reported deaths, as has been done in the past in a number of studies [66, 128]. (It was thought that the CFR of YF could be as high as 47%. Thus it should not be surprising if there were panic at the initial stage of the outbreak.)

Dr. Moreira's account of vector control and vaccination

Dr. Moreira has outlined that the WHO's response was organized into five different stages and that the reaction and support of the public was varied at different times from strong resistance to cooperation [178].

Dr. Moreira states: "*The government set up a task force to lead the response and launched a five-part plan. The first part was active surveillance. We reinforced laboratory capacity to allow for early detection and notification of new cases. The second part was case management. With the help of Moreira Sans Frontieres, we developed case management guidelines and distributed these to different provinces because there is no specific treatment for yellow fever. We also provided health workers with a flowchart indicating six health facilities where people with severe disease should be referred. By February, about 30% of some 300 people with confirmed yellow fever had died. After implementing our plan, cases' fatality dropped to 11%. The third part was mass vaccination, and the fourth was integrated*

vector control measures to lower the density of the Aedes aegypti mosquito that carries yellow fever, as well as dengue, chikungunya and Zika viruses. The fifth part of the strategy was risk communication and social mobilization."

"*We began working in Viana, a marketplace attracting people from all over Angola. It is difficult to control the movement of people and the area has major sanitation problems. Mosquito density was high, so we carried out a mass distribution of larvicide and a social mobilization campaign to explain to people how to use it. We also did indoor and outdoor spraying with insecticides. Initially, we faced resistance: some people kept the larvicide at home and did not use it. So we asked community leaders to help us persuade people to join the campaign. As a result of these vector control measures, the mosquito density level fell substantially.*"

"*We have been conducting fixed post and mobile vaccination campaigns in 73 districts. By September 7, around 16 million individuals, 65% of the Angolan population, had been vaccinated with most of them in reactive campaigns, where local transmission was confirmed. In addition, 3 million of these individuals were vaccinated in a preventive campaign in August and September.*"

2.7 Conclusion

Using modern likelihood-based statistical inference techniques, it was possible to fit a vector-host epidemic model successfully to the surveillance data collected for the YF outbreak in Luanda, Angola in 2016. We were thus able to assess the success of the vaccination campaign as rolled out in Luanda. While there were in reality 73 deaths reported over the 37-week study period, the model showed that the vaccination campaign saved from death 5.1-fold of observed deaths and prevented from illness 5.8-fold of observed cases, over the study period, and no doubt much more if we were to extrapolate beyond the study period. This was determined by simulating Luanda's YF outbreak in the absence of any vaccination. The national vaccination campaign was also found to be timely, in that delaying the availability of the vaccination any further would have greatly enhanced the epidemic in terms of the number of YF cases and mortality.

The change in the number of YF cases over time in Luanda suggests the possibility

that the outbreak occurred in two waves over the 37-week study period. The modelling and sensitivity analysis demonstrated that this is a robust feature (see section 2.3.3), which would have become far more prominent had the vaccination campaign been reduced in intensity. The appearance of waves implies that $\mathcal{R}_0$ must oscillate to some degree in time. Reconstruction of the underlying dynamics reveals that $\mathcal{R}_0$ is strongly out of phase with mortality, so that $\mathcal{R}_0$ decreases when the number of deaths increase, and vice versa. Thus we hypothesize that the high death rates and the number of cases influenced Luanda's population behavioral response which in turn led to some reduction of disease transmission during the years of high mortality. Behavioral responses may typically involve using more insecticide, mosquito repellent, insecticide-treated bednets and broader vector-control programs, as outlined in section 2.6.1. In Luanda, it also involved cordon sanitaire, and movement restriction with the aim of reducing transmission through the wider population [46].

Such behavioral changes are able to modulate the basic reproduction number, which in turn can lead to waves in the YF case numbers. A similar phenomenon was reported for the deadly pandemic influenza (e.g., 1918 influenza pandemic with a fatality rate (per infection) 2% and an attack rate 1/3) [66, 128] but never in mosquito-borne diseases since either the CFR or the AR is typically low. Moreover, we showed how a simpler model that explicitly incorporated human behavior reproduces the observed data in Figure 2.1 (see section 2.3.3). This may be the first example of mortality-driven basic reproduction number in a mosquito-borne disease outbreak. While this possibility appears to hold in other epidemiological contexts (e.g., Spanish flu [66, 128]), it would be beneficial to check this further for vector-host systems. In the case of Luanda's YF outbreak, $\mathcal{R}_0$ is likely to have also been affected by the sporadic but heavy El Niño rainfall, which in turn could influence mosquito population numbers. Such a process could occur even if there is no visible long-term correlation between climate (rainfall) and the vector dynamics.

The modelling approach described here provides a basis for future vaccination campaign evaluations. Since the YF mortality appeared to lead to oscillations in the basic reproduction number ($\mathcal{R}_0$), this possibility should be considered in the development of short-term prediction tools of the spread of YF. The general approach should be of benefit in mitigating the spread and impact of YF outbreaks in the future.

Chapter 3

Modelling Japanese Encephalitis Virus in Hong Kong

Japanese encephalitis virus (JEV) is a mosquito-borne virus, which causes annual epidemics in Southern and Eastern Asia. Human JEV cases could be severe or deadly. Local JEV human cases were reported in Hong Kong annually in recent years with a skip between 2006 and 2010 (when non-local human cases were reported). However, the mechanism behind the "skip-and-resurgence" behavior of JEV in Hong Kong is unclear, which motivates the current study. This chapter formulates a refined mathematical model, combined with modern likelihood-based inference approaches to identify the main factors behind the vanishment and reemergence of JEV in Hong Kong. The model indicates that: (a) vector-free transmission (i.e., without mosquitoes) increases the size of JEV epidemics but is unlikely to maintain outbreaks among pigs; (b) the vanishment of JEV human case in 2006-10 could be due to the rapid decrease of local living pigs' population; (c) the JEV re-emergence could be due to the invasion of a new JEV strain in 2011, which, increases both the transmissibility of the virus and/or the spill-over ratio from reservoirs to hosts.

3.1 Introduction

Japanese encephalitis virus (JEV), from the family *flavivirus*, can cause Japanese encephalitis, one of the deadly infectious diseases popular in southern and eastern Asia, as the clinical fatality rate was reported as 20-30% [10, 24, 43, 53, 161, 165, 208] and over

35% of children [153]). JEV exists in a transmission cycle among mosquitoes, pigs and/or water birds which is transmitted to humans through bites from infected female mosquitoes of the *Culex* species (mainly *Culex tritaeniorhynchus*) [10, 16, 43]. Humans are incidental and dead-end hosts, and can not develop sufficient viraemia to infect feeding mosquitoes [10, 43, 137].

The vertical transmission of JEV exists between mosquitoes and their eggs [202, 203, 214]. Ricklin *et al.*'s recent study revealed the existence of swine-to-swine, the vector-free, transmission path among pigs [197]. Vectors' population increases during spring, peaks in summer and decreases during fall of each year [199, 201]. Therefore, vertical transmission and vector-free transmission are believed to be the driving source of next year's JEV transmission.

Reports from the government of Hong Kong suggest that the seroprevalence of JEV antibodies among swines throughout rainy season (May-July of each year) is up to 91% compared with 34% seroprevalence reported throughout the dry season [21]. The seroprevalence of JEV antibodies among swines is roughly from 80% to 90% in July and August of 2000-04 (see Fig 2A of [201]). Previous local serological surveys conducted by Hong Kong government indicate that 23.5% of pig farmers and 5.9% abattoir workers in Hong Kong are seropositive to JEV antibody, and, by contrast, 0% seropositive cases are reported from 30 blood donors [21]. Although JEV vaccine is reported to have a high effective protection rate (93.3% by 5 years after injection and estimated 85.5% by 10 years [91]), a review study estimates that only 2.4% JEV seropositive cases are found from 1,547 serum samples [155], which indicates that most inhabitants in Hong Kong are immunologically naive to JEV. One recent vaccine protection rate evaluation on the newly introduce genotype 5 (G5, the new wild JEV strain) JEV strain conducted by Cao *et al.* [75] concludes that current JEV vaccine (G3-based doses) has low protection rate, lower than 50% and even much lower for pediatric patients against G5 strain.

The recent press on JEV transmission via human blood transfusion in Hong Kong caught wide media coverage, which, to the best of our knowledge, is the first human-to-human JEV transmission (of serological level) reported in the world [35]. Counting on the two JEV cases found in June 2017, 17 local JEV cases in Hong Kong from 2011 to 2016 is considerably more than that from 2006 to 2010 (with no cases reported). More interestingly, the observed sustained JEV cases for recent seven years from 2011 to 2017 appear after a long period of no JEV cases from 2006 to 2010. Considering the decreasing tendency of

local living pigs from 350,000 in 2004 to 60,000 in 2017 [13, 15, 26, 29, 30], the decrease and vanishment of JEV cases is expected, but the JEV re-emergence is unexpected. It is pivotal to investigate the underlining mechanism of this skip-and-resurgence behavior.

Previous work on the skip-and-resurgence pattern of JEV during 1969 and 2004 by Riley *et al.* suggested the main reason of JEV vanishment from 1990-2002, with only 1 local JEV case reported in 1996, was likely due to the lack of local rice production [201] as the major vectors of JEV breed principally in rice fields [217], and they also proposed that the reemergence from 2003-04 of JEV could be due to increased awareness of emerging infectious diseases as there was a huge SARS epidemic in Hong Kong in 2002-03, and local government indeed strengthened their reporting efforts. However, the mechanism behind the most recent JEV skip-and-resurgence behavior in Hong Kong is unclear.

The re-emergence of local JEV cases (since 2011) in Hong Kong associated with the recent blood-transfusion induced infections poses serious public health issues. This chapter attempts to identify the main factors underlining the observed JEV skip-and-reemergence pattern, and attributed the reason as the pig rearing policy in 2006 and new JEV strain invasion around 2011. The proposed hypothesis will be justified to be plausible and reasonable through a simple mathematical model.

3.2 Theoretical Model

3.2.1 Data

The monthly JEV cases during December 2003 and May 2017 are obtained from [17, 21] for Hong Kong and [22] for mainland China. The weekly JEV cases from December 2003 to May 2017 of Taiwan are obtained from [23]. Furthermore, the monthly regional mosquito ovitrap index in Hong Kong from December 2003 to December 2016 is obtained from [20]. Figs. 3.1 and 3.2 show the annual pattern of reported JEV cases and regional mosquito ovitrap indexes in Hong Kong.

The population information of local living pigs in Hong Kong (Fig 3.3) are obtained from several governmental reports and local news reports [13, 15, 18, 19, 26, 29, 30, 36, 37]. The rapid decrease in the number of local living pigs, as shown in the vertical grey line in

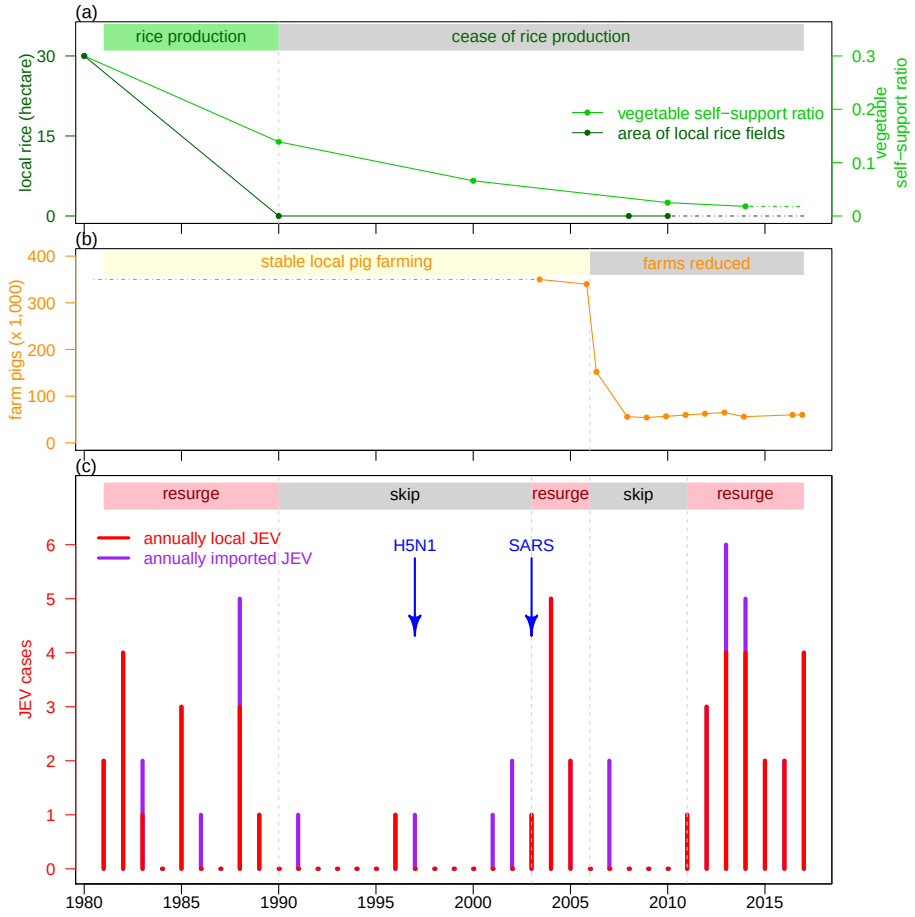


Figure 3.1: The skip-and-resurge pattern of Japanese encephalitis epidemics and its corresponding factors in Hong Kong from 1980 to 2017. Panel **(a)** shows the local vegetable self-support ratio and area (in hectare) of rice fields (dashed line represents estimation). Panel **(b)** shows the number of local farm pigs (dashed line represents estimation). Panel **(c)** shows the annually reported JEV (local in red and imported in purple).

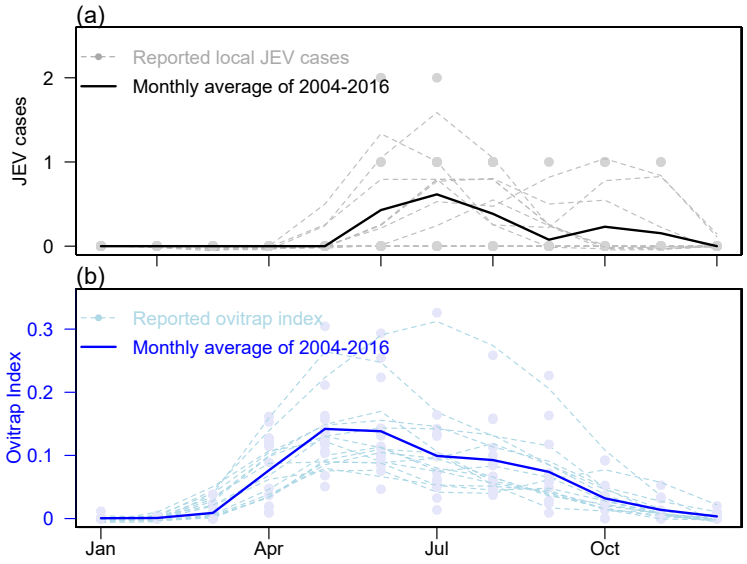


Figure 3.2: Averaged monthly reported local JEV clinical cases and ovitrap index in Hong Kong from 2004 to 2017. Panel (a) is the monthly reported local JEV cases in Hong Kong and panel (b) is the averaged monthly ovitrap index of regions around Yuen Long city (i.e., including Yuen Kong, Yuen Long and Tin Shui Wai). For both panels (a) and (b), lines in dark colors are the averaged annual data, the lines in light colors are the smoothed data and the dots in light colors are the reported data of each year from 2004-17.

Fig 3.3, is mainly due to the pig rearing licences surrender policy posed by the Hong Kong government in May 2006 [30], under which, a large part of pig farms would terminate rearing pigs and it was reported that 243 out of the total 265 pig farms applied for turning over the license.

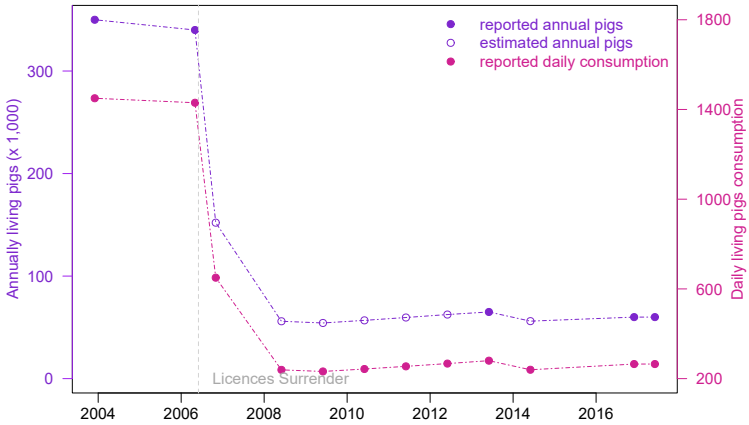


Figure 3.3: Local living pigs' population and daily local living pigs' consumption in Hong Kong from Jan 2004 to May 2017. The line and dots (and circles) in purple represent local living pigs' population (N_p). The line and dots in violet red represent daily local living pigs' consumption ($\nu_p N_p$). The vertical grey dashed line marks the time point when Hong Kong government triggered the pig rearing licences surrender policy.

3.2.2 JEV model

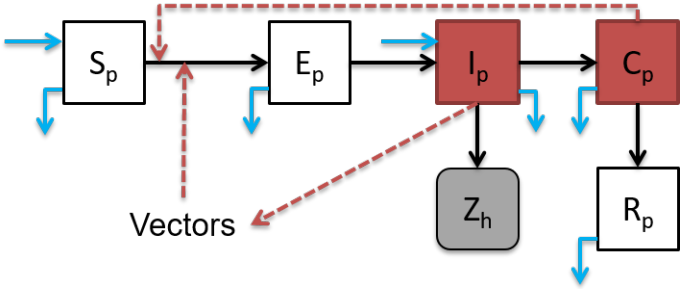


Figure 3.4: The JEV model diagram. Black arrows represent infection status transition paths, red dashed arrows represent transmission paths and light blue arrows represent the birth and death of reservoirs (including slaughter). Red compartments represent infectious classes, and grey compartment is simulated JEV human cases (i.e., Z_h or Z_i in Eqn. (3.2)).

Ricklin *et al.*'s recent study reveals the swine-to-swine, that is, the vector-free transmission path of the virus among pigs and also illustrates the existence of JEV convalescent period (JEV shedding phase) during which JEV can be found at swine oronasal, following the infection period (viraemic phase) during which JEV in swine serum is infectious to mosquitoes [197]. Therefore, the host population (pigs or swine) is categorized into five classes: susceptible, exposed, infectious, convalescent and recovered ones, with the respective size denoted as S_p, E_p, I_p, C_p and R_p. Moreover, we consider two transmission routes in the model, swine-to-swine and vector-borne transmissions. Fig 3.4 shows the model diagram of the JEV disease transmission paths among reservoirs, vectors and humans. Then the JEV transmission cycle can be described by the following system (see Eqn. (3.1)).

$$\begin{cases} S_p' = (1-\eta)\cdot B_p(t)\cdot N_p - \nu_p S_p - \left(\lambda_{vp} + \beta_p \cdot \dfrac{C_p}{N_p}\right) S_p, \\ E_p' = \left(\lambda_{vp} + \beta_p \cdot \dfrac{C_p}{N_p}\right) S_p - (\sigma_p + \nu_p) E_p, \\ I_p' = \eta \cdot B_p(t)\cdot N_p + \sigma_p E_p - (\gamma_p + \nu_p) I_p, \\ C_p' = \gamma_p I_p - (\delta_p + \nu_p) C_p, \\ R_p' = \delta_p C_p - \nu_p R_p. \end{cases} \tag{3.1}$$

The model parameters are summarized in Table 3.1. In this model, the total pig population

$$N_p = S_p + E_p + I_p + C_p + R_p$$

is time-dependent referring to the observed pigs population in Hong Kong (see the purple

dashed line in Fig 3.3). Since JEV human infections can not transmit disease to others [10, 43, 137], we can simply model human cases by directly referring to the pig infections with a variable spill-over ratio (ρ) in week i ρ_i (see Eqn. (3.2)):

$$Z_i = \int_{\text{week } i} \rho_i \gamma_p I_p \, dt. \tag{3.2}$$

Table 3.1: Model parameters. In this table, "v→p" and "p→h" denotes JEV is transmitted from vectors to pigs, and from pigs to humans, respectively.

Parameter	Notation	Value/Range	Remark/Unit	Source(s)
Force of transmission	λ_{vp}	time-dependent	v→p, per year	Eqn. (3.3)
Pig latent period	σ_p^{-1}	1-2	days	[148, 197]
Pig infection period	γ_p^{-1}	2-4	days	[24, 148, 197, 198, 237]
Pig convalescent period	δ_p^{-1}	1-4	days	[197]
Imported infection ratio	η	0.43%-1.45%	pigs, Nil	Eqn. (3.7)
Effective contact rate	β_p	0.0-0.4	pigs, per days	assumed
Pig living period	ν_p^{-1}	234.0	days	Eqn. (3.6)
Pig population	N_p	time-dependent	Nil	[13, 15, 21, 26, 29, 30]
Spill over ratio	ρ	time-dependent	p→h, Nil	Eqn. (3.4)

3.2.3 Parameter estimation

In this subsection, we are going to estimate the force of transmission from vectors to reservoirs λ_{vp}, spill-over rate from reservoirs to hosts ρ, living period of swine ν_p^{-1}, pig population N_p, imported infection ratio η.

Force of transmission from vectors to reservoirs λ_{vp}: Theoretically, one can express λ_{vp} as detailed as $\lambda_{vp} = a\vartheta_{vp} \cdot \frac{I_v}{N_p}$, where a is the mosquito biting rate, ϑ_{vp} is the successful JEV transmission probability per mosquito bite and I_v is the number of infected mosquitoes. However, in this paper, we are employing a vector-free modelling framework by simplifying λ_{vp} as a function of the mosquito ovitrap index over time. This is justified by observing that a and ϑ_{vp} are constant, while $\frac{I_v}{N_p}$ is roughly proportional to the ovitrap index. In short, we assume

$$\lambda_{vp} = k \cdot \omega(t) + b \tag{3.3}$$

where ω is the time series of ovitrap index in Hong Kong, k and b are model parameters under estimation. The constant b represents the contribution of the vertical transmission

from adult mosquitoes to their eggs since the average vertical transmission ratio of *Culex tritaeniorhynchus* is reported from 12% to nearly 100% [202, 214]. By using Eqn. (3.3), we also incorporate the case that although the ovitrap index becomes very low or approximately zero in the dry season, the transmission rate can still be non-zero because of the effect of vectors' vertical transmission.

Spill-over rate from reservoirs to humans ρ: In equation (3.2), the reported JEV human cases are proportional to swine infections weighted with a time-dependent spill-over rate (ρ) since human JEV infections can not transmit disease to others (i.e., there is no human-to-human transmission path) [10, 43, 137]. Since human cases are also largely related to the total number of vectors, we can further model the spill-over rate (ρ) as a function of ovitrap index:

$$\rho = \xi \cdot \omega(t - \tau) \tag{3.4}$$

where ξ is the infectivity strength parameter under estimation and τ is the summation of mosquito's incubation period (6-12 days [150, 150, 214, 222]), human latent period (5-13 days [10, 16, 24]) and case reporting delay. In this chapter, we fix $\tau = 15$ days for simplicity.

In this chapter, to investigate the mechanism of the observed re-emergence of JEV in Hong Kong after 2011, we are going to use the partitioned spill-over rate with the hypothesis that JEV re-emergence was due to the invasion of a new JEV strain. This hypothesis will be validated later through statistical approach. Theoretically, the spill-over ratio (ρ) under the new JEV strain invasion scenario is significantly higher than the non-invasion scenario, since the pig population is immune-naive to new strains for the first few years after the invasion. Following Tien *et al.*'s cholera modeling study [218], we assume the spill over ratio ρ under new JEV strain invasion scenario takes the following form

$$\rho = \begin{cases} \xi_1 \cdot \omega(t - \tau), & t < T_0 \\ \xi_2 \cdot \omega(t - \tau), & t \geqslant T_0 \end{cases} \tag{3.5}$$

where T_0 is the starting time instant when new JEV strain invaded. According to [218], high confidence level of the new strain invasion hypothesis is associated with $\xi_2 > \xi_1 > 0$ (or almost equivalently, the average spill-over ratio over time $\langle\rho\rangle$ before invasion should be significantly less than that after invasion). The normal situation without new strain invasion (Eqn. (3.4)) can be regarded as the special case when $\xi_1 = \xi_2$ in Eqn. (3.5).

Living period of pigs ν_p^{-1}: According to the Hong Kong government [18] and local press reports [19, 36], the pork consumption is around 265 living local/domestic pigs per day in Hong Kong from 2016-17. One can also compute, according to dated governmental report [30], that around 1,450 living local pigs were consumed per day in Hong Kong back to 2004. The total swine population was around 350,000 in 2004-05 [15, 29, 30], around 65,000 in 2012 [13] and around 60,000 during 2016-17 [26]. Therefore, the average living period of pigs $\langle \nu_p^{-1} \rangle$ can be estimated as

$$\begin{aligned} \text{During 2004-05}: \quad \langle \nu_p^{-1} \rangle &= \frac{N_p}{\text{daily consumptions}} = \frac{350000}{1450} \approx 241.38 \quad \text{days}, \\ \text{In 2012}: \quad \langle \nu_p^{-1} \rangle &= \frac{N_p}{\text{daily consumptions}} = \frac{65000}{275} \approx 236.36 \quad \text{days}, \\ \text{During 2016-17}: \quad \langle \nu_p^{-1} \rangle &= \frac{N_p}{\text{daily consumptions}} = \frac{60000}{265} \approx 226.42 \quad \text{days}. \end{aligned} \tag{3.6}$$

Taking the average of three numbers, we have the average living period of pigs $\nu_p^{-1} \approx 234$ days.

Pigs' population N_p: Given the Hong Kong's daily local living pig consumption is 600-700, with average 650 pigs per day, in 2007 [30] (Fig 3.3), we can infer that total number of pigs in 2007 is $N_p = 152100$ by using Eqn. (3.6) backwardly. The rapid decrease in the total rearing number of pigs (from 2006 to 2007) was mainly due to the reduced swine rearing licenses in earlier 2006 [30], which resulted in that 243 out of the total 265 pig farms applied for turning over the license. Although the daily living pigs consumption is not included in the model, given the pigs' living period, we can infer the population number according to the daily consumption amount.

Imported infection ratio among swine η: The imported infection ratio (η) can be computed as

$$\eta = \langle \text{IAR}_p \rangle \cdot \frac{\gamma_p^{-1} + \delta_p^{-1}}{\langle \nu_p^{-1} \rangle}, \tag{3.7}$$

where $\langle \text{IAR}_p \rangle$ is the average attack rate over the average living period of pigs $\langle \nu_p^{-1} \rangle$. Intuitively, the proportion of the infectious pigs among all imported living pigs (η) is averagely the the JEV infection attack rate (IAR_p) multiplies the probability of the infection is still on-going ($\frac{\gamma_p^{-1}+\delta_p^{-1}}{\langle \nu_p^{-1} \rangle}$). For example, if $\gamma_p^{-1} = 1.5$ days, $\delta_p^{-1} = 2.5$ days, $\langle \nu_p^{-1} \rangle = 234$ days and $\langle \text{IAR}_p \rangle \in [25\%, 85\%]$ [16, 145], we can estimate $\eta \in [0.43\%, 1.45\%]$.

3.3 Model Validation and Results

3.3.1 Fitting procedures

The JEV cases in Hong Kong are modeled as a Partially Observed Markov process (POMP), also known as hidden Markov model, with `R` package "`POMP`" available at [146]. The Iterated Filtering and plug-and-play likelihood-based inference frameworks are employed to fit the time series [111, 133, 147]. Furthermore, the Maximum Likelihood Estimate (MLE) is used to estimate model parameters. To quantify the tradeoff between the goodness-of-fit of a model and its complexity [205], Bayesian information criterion (BIC) is employed as a criterion for model comparison. Simulations are performed by deploying the Euler-multinomial integration method with a fixed time-step one day [48, 133].

The model is first validated with the observed JEV cases in Hong Kong, given knowledge of the swine population information. The mosquito abundance is time-dependent, smoothed over time line based on the local ovitrap index ω. The time-dependent force of transmission from vectors to reservoirs λ_{vp} and the spill-over rate ρ can then be estimated through ω.

The monthly observed cases, C_i, are assumed to follow a Poisson distribution (denotes Poi) with the mean Z_i, the real monthly cases modelled by Eq (3.2). Hence, we have

$$C_i \sim \text{Poi}\,(\lambda = Z_i) \qquad \text{with} \quad \text{mean} : \mu_i = Z_i.$$

Then the overall log-likelihood function l is given by

$$l(\Theta|C_1, \ldots, C_N) = \sum_{i=1}^{n} \ln f(C_i|C_{1:(i-1)}, \Theta)$$

where Θ denotes the parameter vector under estimation, $f(C_i|C_{1:(i-1)}, \Theta)$ is the posterior probability measurement function for C_i given $C_{1:(i-1)}$, which will be numerically computed by Sequential Monte Carlo (SMC, also known as particle filter) [133], and n denotes the total number of months during the study period.

The confidence intervals (C.I.) of parameters are estimated based on parameters' ranges in Table 3.1 by using the method of profile likelihood confidence intervals [132, 133]. Param-

eter estimation and statistical analysis are conducted using R (version 3.4.1).

3.3.2 Fitting results

In addition to showing simulation median, we also present annual mean of the model prediction by the approach in [74], since simulation mean demonstrates fitting results more consistently when the data are restricted as integers and relatively noisy.

The fitting results of the model under new JEV strain invasion scenario are shown in Fig 3.5. Estimated results of model parameters are summarized in Table 3.2. Although the long-term fitting is roughly acceptable due to numbers of local JEV cases are very noisy (either 0, 1 or 2 per month), the 95% simulation quantile interval covers all observed data, and the average annual pattern is consistent.

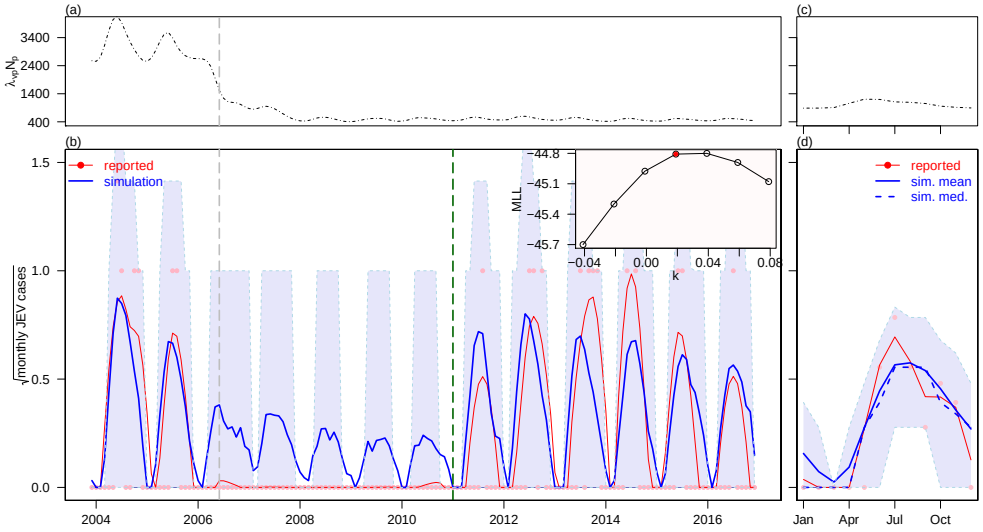


Figure 3.5: Fitting results of JEV local cases in Hong Kong from 2004 to 2016 under new JEV strain invasion scenario. Panel (a) and (b) are the scaled force of transmission (from vectors to pigs, scaled by the population size of pigs) and simulation results from 2004 to 2016 respectively. Panel (c) and (d) are the one-year-average scaled force of transmission and simulation results from 2004 to 2016 respectively. In panel (a) and (b), black dashed lines are the scaled force of transmission. In panel (b) and (d), blue lines are the simulation results, shaded regions are 95% quantile interval from simulation, pink dots are the reported (observed) JEV local cases and red lines are the smoothed (by *loess* function) reported JEV cases. The vertical grey dashed line marks the time point when Hong Kong government posed the pig rearing licences surrender policy. The vertical dark green dashed line marks the time point when the new JEV strain introduced to the pigs' population. The inset panel shows the maximum log-likelihood (MLL) values of different ks, the red dot with the highest MLL are selected for fitting in main panels.

We find BIC reduces over 28 units compared to the baseline no invasion scenario, which

Table 3.2: Summary table of parameter estimation under new JEV strain invasion scenario. X_{p0} denotes the initial proportion of class X_p.

Parameter	Notation	Value	Type	Initial status	Unit/Remarks
Average force of transmission	$\langle\lambda_{vp}\rangle$	0.0042	estimated	time-dependent	per year
Pig latent period	σ_p^{-1}	1.5	fixed	1-2	days
Pig infection period	γ_p^{-1}	3	fixed	2-4	days
Pig convalescent period	δ_p^{-1}	2.5	fixed	1-4	days
Imported infection ratio	η	1.0%	fixed	0.43%-1.45%	Nil
Effective contact rate	β_p	0.0011	estimated	0.0-0.4	per days
Pig living period	ν_p^{-1}	234	fixed	234	days
Pig population	N_p	-	-	time-dependent	pigs
Average spill over ratio	2004-10:$\langle\rho\rangle$	0.0002	estimated	time-dependent	before invasion
Average spill over ratio	2011-16:$\langle\rho\rangle$	0.0024	estimated	time-dependent	after invasion
Average ovitrap index	$\langle\omega\rangle$	0.0564	given	time-dependent	Nil
Initial susceptible	S_{p0}	0.6818	estimated	45-75%	Nil
Initial exposed	E_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial infectious	I_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial convalescent	C_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial recovered	R_{p0}	0.3152	estimated	25-55%	Nil
BIC	BIC	140.2633	estimated	-	Nil

is presented in 3.4.1. In order to avoid the over-fitting problem, we also conduct another invasion scenario where only the force of transmission λ_{vp} is partitioned, and find that BIC is in line with the main results (around 1 unit BIC difference implies almost equivalent goodness-of-fit). The result of partitioned force of infection with no partition on spill-over rate as in Eqn. (3.5) is presented in 3.4.2. Furthermore, an additional invasion scenario with time-dependent λ_{vp} and ρ is investigated in 3.4.3, and the difference of BIC is 4.55.

3.3.3 Estimate basic reproduction number of vector-free transmission

According to the next generation matrix method [68, 221], the basic reproduction number $\mathcal{R}_{pp}$ of pig-to-pig transmission, that is the vector-free transmission route, can be calculated as

$$\mathcal{R}_{pp} = \frac{\beta_p\gamma_p \cdot (\eta\nu_p + \sigma_p)}{[(1-\eta)(\gamma_p + \delta_p + \nu_p)\nu_p + \gamma_p\delta_p](\nu_p + \sigma_p)}. \tag{3.8}$$

It is easy to see that $\mathcal{R}_{pp} \approx \frac{\beta_p}{\delta_p+\nu_p}$ when imported infections are rare ($\eta \approx 0^+$), both the incubation and infection periods of JEV in pigs are negligible ($\sigma_p^{-1} \approx 0^+$ and $\gamma_p^{-1} \approx 0^+$). We estimate $\mathcal{R}_{pp}$ to be 0.0026 (95% C.I.: [0.00,0.30]) under the invasion scenario (see Fig 3.6).

Furthermore, the range of effective contact rate ($\beta_p \in [0.0, 0.4]$, see Table 3.1 and 3.2) is set correspondingly to $\mathcal{R}_{pp} \in [0.0, 1.0]$.

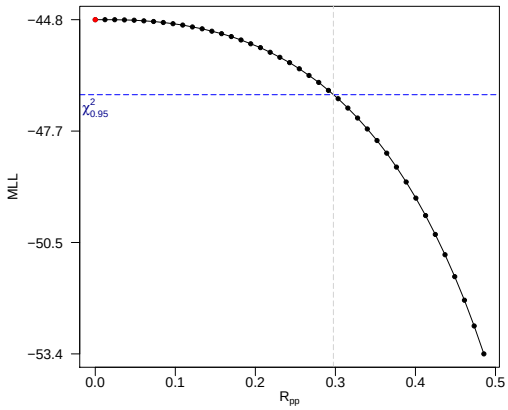


Figure 3.6: The estimation result of the basic reproduction number of pig-to-pig transmission $\mathcal{R}_{pp}$ under new JEV strain invasion scenario with variable ρ. The horizontal blue dashed line is the 95% confidence threshold from the profiles likelihood approach.

3.3.4 Critical community size

Nasell's study [181] formulated the approximation of critical community size (CCS) from simple compartmental models. Although Nasell's work is based on directly-transmitted (i.e., transmitted without vectors) diseases, we adopt the idea of CCS approximation formula for the vector-borne diseases, since the JEV model in this chapter does not include vectors' compartments (the effects of vector-borne transmission is modelled by parameter λ_{vp} in Eqn. (3.3)). The CCS approximation value can be formulated as

$$\text{CCS} \approx \frac{2\pi}{\ln 2} \cdot \frac{\mathcal{R}_0 \cdot \alpha_p^{1.5}}{(\mathcal{R}_0 - 1)^{1.5}} \tag{3.9}$$

where $\alpha_p = \frac{\gamma_p + \nu_p}{\nu_p}$ denotes the ratio of the average living period (rearing period for pigs ν_p^{-1}) to the average duration of infections. Further, the basic reproduction number of pig-to-pig transmission path $\mathcal{R}_{pp}$ in Eq (3.8) is relatively small compared to the basic reproduction number of vector-borne transmission path $\mathcal{R}_{vp}$. If we directly plug in the parameter values under new JEV strain invasion scenario in Table 3.2 to Eqn. (3.8), $\mathcal{R}_{pp}$ is only 0.0026, which is much less than $\mathcal{R}_{vp}$. $\mathcal{R}_{vp}$ is believed to be greater than 1.0, otherwise JEV would not spread during every rainy season, for example, Khan *et al.*'s recent modelling study

estimates $\mathcal{R}_{vp}$ to be 1.2 among pigs [145]. By using the next generation matrix approach [68, 111], the basic reproduction number is

$$\mathcal{R}_0 = \frac{\mathcal{R}_{pp} + \sqrt{\mathcal{R}_{pp}^2 + 4\mathcal{R}_{vp}^2}}{2}.$$

Therefore, we further have $\mathcal{R}_0 \approx \mathcal{R}_{vp}$ if we ignore the effect of $\mathcal{R}_{pp}$ (i.e., $\mathcal{R}_{pp} \approx 0^+$ as in [111]). If we fix $\nu_p^{-1} = 234.0$, set $\mathcal{R}_0 \in [1.10, 1.40]$ and $\gamma_p^{-1} \in [2.0, 4.0]$, then the relationship among α_p, $\mathcal{R}_0$ and CCS in Eqn. (3.9) can be illustrated in Fig 3.7.

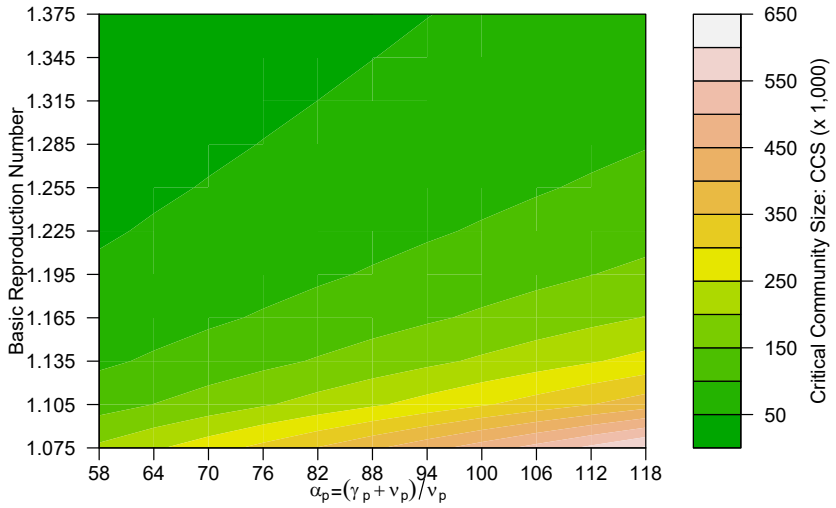


Figure 3.7: Contour plot of the relationship among critical community size (CCS), α_p and the basic reproduction number $\mathcal{R}_0$. Color code from green (the lowest CCS) to gray (the highest CCS) is shown on right.

We can further predict that CCS is around 150,000, which is the rough number of local living pigs in Hong Kong after the rearing license surrender policy [30] (note that the number of local living pigs was from 60,000 to 80,000 since 2008, see Fig 3.3), which could explain the local JEV case vanishment from 2006 to 2010 in Hong Kong.

Moreover, for the invasion scenario, we find that the force of transmission λ_{vp} can largely increase after the introduction of new JEV strain with the fixed spill-over rate ρ. The fitting results of the partitioned force of infection, with no partition on spill-over rate, are presented in the 3.4.2, which yields an almost equivalent goodness of fitting as in the previous case.

We also test the partitioned λ_{vp} and ρ scenario in 3.4.3. Although the fitting result is not as good as that predicted with the invasion scenario case, it is still significantly improved compared to the baseline no invasion scenario since BIC has improved 24 units. We find the

estimated force of transmission λ_{vp} also increases after introducing new JEV strain (see the Table in 3.4.3). Therefore, with increased $\mathcal{R}_0$, CCS level would become lower than the local living pigs' population level (see Fig 3.7), which explains the reemergence of JEV cases.

3.4 More Fitting Results Under Different Scenarios

3.4.1 Baseline fitting results

In this case, the force of transmission from mosquitoes to pigs is set as

$$\lambda_{vp} = k \cdot \omega(t) + b$$

where ω is the time series of ovitrap index in Hong Kong, k and b are model parameters under estimation. The baseline fitting results of the model are shown in Fig 3.8. Estimated results of model parameters are in Table 3.3.

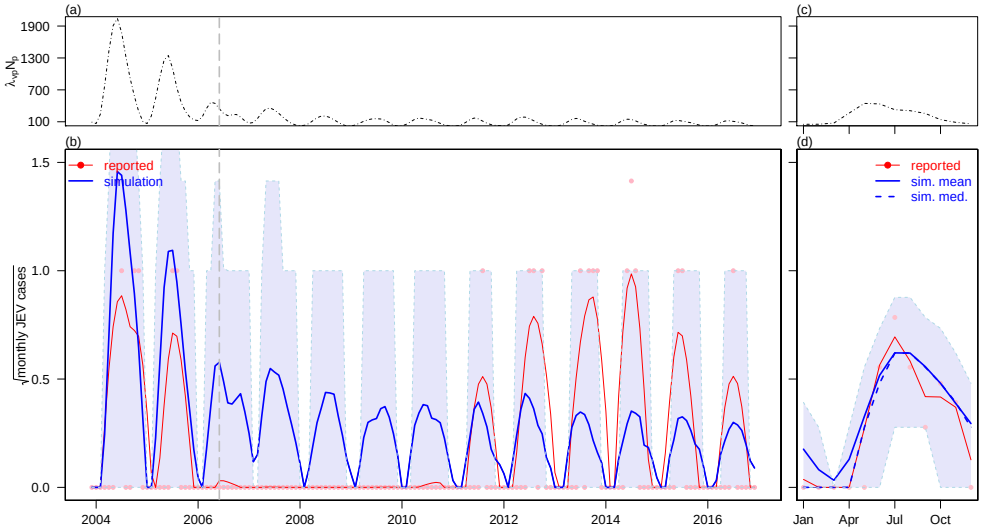


Figure 3.8: Fitting results of JEV local cases in Hong Kong from 2004 to 2016 under baseline scenario. Panels' information and color code are same as in Fig. 3.5.

Table 3.3: Summary table of model parameters' estimates under baseline scenario. X_{p0} denotes the initial proportion of class X_p.

Parameter	Notation	Value	Type	Initial status	Unit
Average force of transmission	$\langle\lambda_{vp}\rangle$	0.0019	estimated	time-dependent	per year
Pig latent period	σ_p^{-1}	1.5	fixed	1-2	days
Pig infection period	γ_p^{-1}	3	fixed	2-4	days
Pig convalescent period	δ_p^{-1}	2.5	fixed	1-4	days
Imported infection ratio	η	1.0%	fixed	0.43%-1.45%	Nil
Effective contact rate	β_p	0.0098	estimated	0.0-0.4	per days
Pig living period	ν_p^{-1}	234	fixed	234	days
Pig population	N_p	-	-	time-dependent	pigs
Average spill over ratio	$\langle\rho\rangle$	0.0008	estimated	time-dependent	Nil
Average ovitrap index	$\langle\omega\rangle$	0.0564	given	time-dependent	Nil
Initial susceptible	S_{p0}	0.5847	estimated	45-75%	Nil
Initial exposed	E_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial infectious	I_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial convalescent	C_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial recovered	R_{p0}	0.3123	estimated	25-55%	Nil
BIC	BIC	168.7009	estimated	-	Nil

3.4.2 Fitting results of partitioned force of transmission

In this case, fixing the spill-over ratio, we assume the force of transmission under new JEV strain invasion scenario is

$$\lambda_{vp} = \begin{cases} k_1 \cdot \omega(t) + b, & t < T_0 \\ k_2 \cdot \omega(t) + b, & t \geqslant T_0 \end{cases} \tag{3.10}$$

where T_0 is the starting time instant when new JEV strain joined the system. The situation with no invasion of new strains can be regarded as the special case that $k_1 = k_2$ in Eqn. (3.10).

The fitting results of the new JEV strain invasion scenario with variable force of transmission are shown in Fig 3.9. Estimated results of model parameters are in Table 3.4. The estimate of $\mathcal{R}_{pp}$ is 0.0053 (95% C.I.: [0.00,0.31]) under the invasion scenario with partitioned force of transmission (see Fig (3.10)).

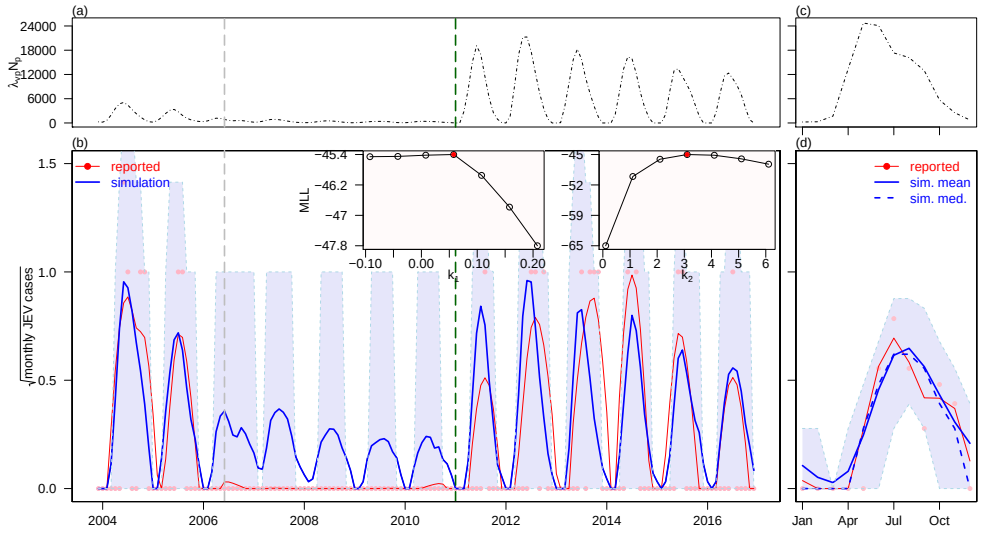


Figure 3.9: Fitting results of JEV local cases in Hong Kong from 2004 to 2016 under new JEV strain invasion scenario with variable force of transmission (λ_{vp}). Panels' information and color code are same as in Fig. 3.5.

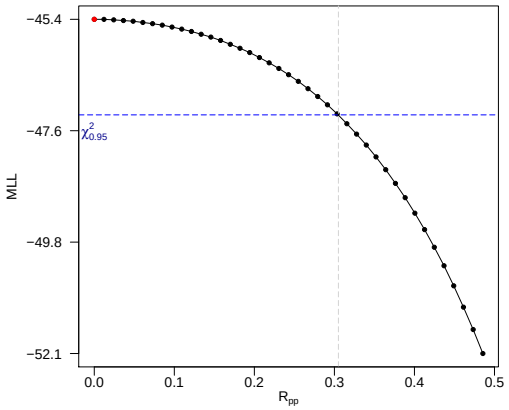


Figure 3.10: The estimation result of the basic reproduction number of pig-to-pig transmission ($\mathcal{R}_{pp}$) under new JEV strain invasion scenario with variable λ_{vp}. The horizontal blue dashed line is the 95% confidence threshold.

3.4.3 Fitting results of partitioned λ_{vp} and ρ

For partitioned λ_{vp} and ρ, similar to (3.10), we also assume ρ is increased after new strain invasion as:

$$\rho = \begin{cases} \xi_1 \cdot \omega(t-\tau), & t < T_0 \\ \xi_2 \cdot \omega(t-\tau), & t \geqslant T_0 \end{cases}$$

where τ is the time delay and T_0 is the starting time instant when new JEV strain joined the system. The force of transmission (λ_{vp}) is modelled as in Eqn. (3.10).

The fitting results of the new JEV strain invasion scenario with increased λ_{vp} and ρ after invasion are shown in Fig 3.11. Estimated results of model parameters are in Table

Table 3.4: Summary table of model parameters' estimates under new JEV strain invasion scenario with variable force of transmission (λ_{vp}). X_{p0} denotes the initial proportion of class X_p.

Parameter	Notation	Value	Type	Initial status	Unit/Remarks
Average force of transmission	2004-10:$\langle\lambda_{vp}\rangle$	0.0044	estimated	time-dependent	before invasion
Average force of transmission	2011-16:$\langle\lambda_{vp}\rangle$	0.1763	estimated	time-dependent	after invasion
Pig latent period	σ_p^{-1}	1.5	fixed	1-2	days
Pig infection period	γ_p^{-1}	3	fixed	2-4	days
Pig convalescent period	δ_p^{-1}	2.5	fixed	1-4	days
Imported infection ratio	η	1.0%	fixed	0.43%-1.45%	Nil
Effective contact rate	β_p	0.0022	estimated	0.0-0.4	per days
Pig living period	ν_p^{-1}	234	fixed	234	days
Pig population	N_p	-	-	time-dependent	pigs
Average spill over ratio	$\langle\rho\rangle$	0.0003	estimated	time-dependent	Nil
Average ovitrap index	$\langle\omega\rangle$	0.0564	given	time-dependent	Nil
Initial susceptible	S_{p0}	0.5767	estimated	45-75%	Nil
Initial exposed	E_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial infectious	I_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial convalescent	C_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial recovered	R_{p0}	0.4203	estimated	25-55%	Nil
BIC	BIC	141.2743	estimated	-	Nil

3.5. The estimate of $\mathcal{R}_{pp}$ is 0.014 (95% C.I.: [0.00,0.31]) under this scenario (see Fig 3.12).

3.5 Discussion

This chapter develops a simple mathematical model to investigate the mechanisms to derive the skip-and-resurgence pattern of Japanese encephalitis virus in Hong Kong. The critical community size (CCS) estimated through the model indicates that the pig rearing licence surrender policy in May 2006 could be responsible for the JEV vanishment during 2006-10. Compared to the results of baseline scenario (see 3.4.1), our hypothetical fitting results imply that the re-emergence of JEV in 2011 was likely due to the introduction of new JEV strains.

We estimate the basic reproduction number $\mathcal{R}_{pp}$ of pig-to-pig vector-free transmission route to be 0.0026 (95% C.I.: [0.00,0.30]). Although the vector-free JEV transmission exists [197], which can increase the epidemic size and prolongs the outbreak, JEV is unable to spread among pigs without vectors.

By implementing the Mann-Whitney U-test approach, Riley *et al.* suggests the JEV

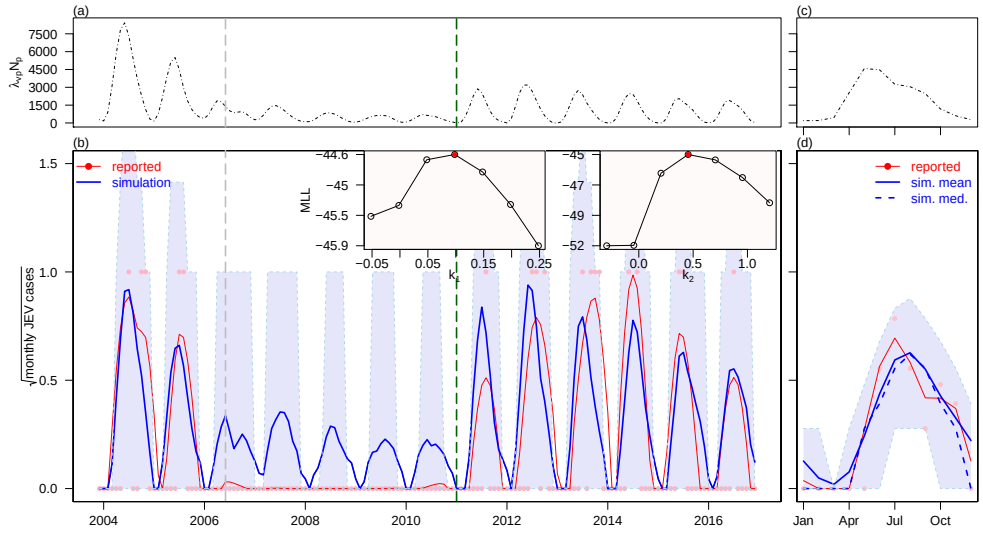


Figure 3.11: Fitting results of JEV local cases in Hong Kong from 2004 to 2016 under new JEV strain invasion scenario with both variable λ_{vp} and ρ. Panels' information and color code are same as in Fig. 3.5.

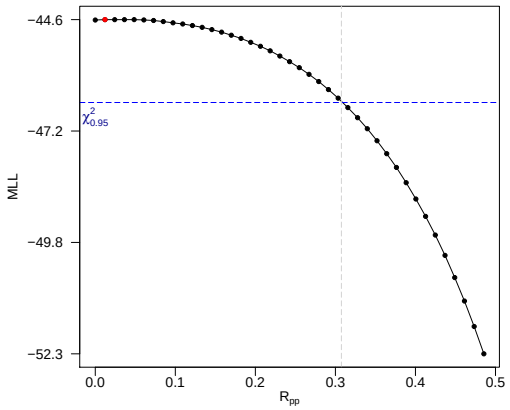


Figure 3.12: The estimation result of the basic reproduction number of pig-to-pig transmission ($\mathcal{R}_{pp}$) under new JEV strain invasion scenario with both variable λ_{vp} and ρ. The horizontal blue dashed line is the 95% confidence threshold.

skip in 1990 was in line with the cease of local rice production in Hong Kong [201] as JEV vectors principally breed in rice fields [217] (also see Fig. 3.1). Since the cease of local rice production was due to urbanization in Hong Kong, it is unlikely that the rice production came back after 2003. Therefore, cease of local rice production is reasonable to explain the 1990-2002 skip, with only one local JEV case reported in 1996. Riley *et al.* attributes the 2003-04 reemergence to the increase of public awareness and improvement of diseases surveillance system after 2002-03 SARS epidemic [201] (see Fig. 3.1), we also notice that JEV was included in the local monthly infectious disease reports of Center for Health Protection, the Hong Kong government since 2004 [16].

In this paper, we claim that the dramatical decrease of local living pigs' population

Table 3.5: Summary table of model parameters' estimates under new JEV strain invasion scenario with both variable λ_{vp} and ρ. X_{p0} denotes the initial proportion of class X_p.

Parameter	Notation	Value	Type	Initial status	Unit/Remarks
Average force of transmission	2004-10:$\langle\lambda_{vp}\rangle$	0.0043	estimated	time-dependent	before invasion
Average force of transmission	2011-16:$\langle\lambda_{vp}\rangle$	0.0071	estimated	time-dependent	after invasion
Pig latent period	σ_p^{-1}	1.5	fixed	1-2	days
Pig infection period	γ_p^{-1}	3	fixed	2-4	days
Pig convalescent period	δ_p^{-1}	2.5	fixed	1-4	days
Imported infection ratio	η	1.0%	fixed	0.43%-1.45%	Nil
Effective contact rate	β_p	0.0058	estimated	0.0-0.4	per days
Pig living period	ν_p^{-1}	234	fixed	234	days
Pig population	N_p	-	-	time-dependent	pigs
Average spill over ratio	2004-10:$\langle\rho\rangle$	0.0002	estimated	time-dependent	before invasion
Average spill over ratio	2011-16:$\langle\rho\rangle$	0.0013	estimated	time-dependent	after invasion
Average ovitrap index	$\langle\omega\rangle$	0.0564	given	time-dependent	Nil
Initial susceptible	S_{p0}	0.6470	estimated	45-75%	Nil
Initial exposed	E_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial infectious	I_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial convalescent	C_{p0}	0.001	assumed	0.0-0.25%	Nil
Initial recovered	R_{p0}	0.3400	estimated	25-55%	Nil
BIC	BIC	144.8174	estimated	-	Nil

could be responsible for the newest 2006-10 JEV skip. Furthermore, with such a low pigs' population, the newest JEV reemerge from 2011 onwards was probably due to new JEV strain invasion. Although we can not find local genetic studies to support the new strain invasion hypothesis, interestingly, following studies present positive side of our hypothesis: (i) **JEV genotype 1 (G1) strain since 2000:** As genotypes 1 and 3 occurred principally in temperate, epidemic areas [165], JEV studies of southeastern Asia reported that genotype 3 (G3) JEV was predominant during the later 20th century, then, G1 strain started replacing G3 around 2000 and became dominate thereafter [113, 165, 189]. One genetical study (Table 2 of [114]) found that G1 strain was not observed until 2008-10 in the majority of Chinese regions surrounding Hong Kong. In addition, the immune response of the existing JEV vaccine (based on G3 strain) against G1 was less pronounced [98, 245]. Thus, it is very likely that the possible newly introduced JEV strain was from these surrounding Chinese regions (specifically, it could be due to living pigs imported from mainland China and/or Taiwan [14]); (ii) **JEV genotype 5 strain (G5):** Similar JEV epidemics re-emergences were observed twice in South Korea in 1998 and 2010 (Fig A of [212]). The reemergence in 1998 is likely due to the introduction of G1 strain in the mid-90s [114, 165]. Interestingly, the first isolated local G5 strain was reported in 2010 in South Korea, matching the reemergence of JEV in 2010 [215] as the average number of annual JEV cases increased roughly 6- to

8-fold [212]. This also supports the main results of increasing the spill-over rate (Table 3.2) and the force of transmission in 3.4.3. The JEV G5 strain was also first isolated in 2009 in China [114]. The bird migration might bring the new JEV strain from the departure places (e.g., southeastern China, northern Thailand and/or Vietnam [114, 215]).

Given the almost equivalent goodness of fitting of both results under the invasion scenario (main results with partitioned ρ and other results with partitioned λ_{vp}), there exist at least three possible explanations to the JEV re-emergence since 2011: (**E1**) The newly introduced JEV strain has slightly increased the transmissibility from vectors to pigs, and largely increased the spill-over ratio from pigs to humans; (**E2**) The newly introduced JEV strain has largely increased the transmissibility from vectors to pigs, but the spill-over ratio from pigs to humans held steady; (**E3**) The newly introduced JEV strain has increased both transmissibilities from vectors to pigs and spill-over ratio from pigs to humans.

The symptomatic ratio of JEV can be employed to further refine these factors. Most JEV infections develop no symptoms, and around 30% of clinical JEV cases die (Table 3.6). We presume the JEV symptomatic ratio among pigs is in line with that of humans, and asymptomatic pigs have negligible JEV transmissibility to vectors due to low within-host viral load.

Table 3.6: Table of symptomatic JEV infection ratio and case-fatal rate (CFR) of JEV with clinical illness from different sources. The numbers in brackets (i.e., $(\cdot)$) are the geometric average of the upper and lower bounds of the ranges.

$\frac{\text{Symptomatics}}{\text{Total Infections}}$ (Symptomatic%)	CFR = $\frac{\text{Mortality}}{\text{Clinical Illness}}$	Source(s)
(0.48%) 0.4%-0.5%	30%	[24, 43]
$< 1\%$	20%-30%	[10, 53]
(0.81%) 0.33%-2%	25%	[161]
(0.63%) 0.1%-4%	25%-30%	[165, 208]
(0.63%) 0.1%-4%	-	[78, 122, 123, 125, 149]
-	35%	[176]
-	36.4% (children)	[153]
4% (occasionally)	-	[222]

Our main results are derived from **E1** (Fig 3.5). In Table 3.2, we assume the yearly force of transmission $\langle \lambda_{vp} \rangle$ is 0.0042 and fix the proportion of susceptible swine S_{p0} to be 68% at the beginning of every year [16, 150, 201], the annually average JEV infection attack

rate (IAR_p) among pigs can be roughly calculated as

$$\mathrm{IAR}_p = \frac{\langle \lambda_{vp} \rangle \cdot S_{p0}}{\text{Symptomatic\%}} \tag{3.11}$$

where Symptomatic% is the JEV symptomatic rate, by which IAR_p is estimated from 35.26% to 59.50% (with Symptomatic% $\in [0.48\%, 0.81\%]$), which is consistent with [16, 145, 201]. The results corresponding to **E2** scenario are presented in 3.4.2, where the yearly transmission rate $\langle \lambda_{vp} \rangle$ is set to be 0.0044 and 0.1763 for time instants before and after invasion respectively. The larger yearly force of transmission after invasion, for example $\langle \lambda_{vp} \rangle = 0.1763$, produces unreasonable IAR_p (3.4.2), which would be strictly larger than 100%. The mechanism **E3** implies increase in both λ_{vp} and ρ after invasion (see 3.4.3). With $S_{p0} = 64\%$, IAR_p is increased from the interval [33.98%,57.33%] to [56.10%, 94.67%] after invasion, with its range being acceptable since Symptomatic% can increase up to 4.0% occasionally [222]. Hence, **E1** and **E3** are likely to be the real situation of new JEV strain invasion in 2011 while **E2** is unlikely since it predicts unrealistic IAR_p. Comparing **E1** with **E3**, we can see that CCS is very easy to increase (could be due to transmission seasonality and/or changing of pigs' herd immunity to new strain) to local pigs' population size in scenario **E1**.

Since the scenario of **E1** would not change the value of CCS (but **E3** could), we further propose that **E3** is more likely to be the real situation that is responsible for the resurgence of JEV in Hong Kong. Further work is needed in order to identify the biological mechanisms regarding to **E3**.

Chapter 4

Modelling Zika Virus in Northeast Brazil

Between January 2015 and August 2016, two epidemic waves of Zika virus (ZIKV) disease swept Northeast region of Brazil. As a result, two waves of Guillain-Barré Syndrome (GBS), a type of severe acute paralytic neuropathology, were observed in this region. On the one hand, the mandatory reporting of ZIKV disease was started region-wide in February 2016 which means that the ZIKV cases could be highly underreported (or misdiagnosed as dengue) before that. On the other hand, the diagnosis and reporting of GBS cases was most likely reliable given the severity and the easily recognizable symptoms. It is now well-accepted that those excess GBS cases were mainly induced by ZIKV infection. Thus we propose to infer the true epidemic of ZIKV cases from the two waves of excess GBS cases. To reconstruct the epidemic of ZIKV disease is essential to assess the risk of GBS per ZIKV infection and the risk of microcephaly per ZIKV infection in this region. We modified our previous ZIKV model to allow the asymptomatic (human) cases of ZIKV infection to be infectious. We found that the estimate of the infection attack rate is dependent on the proportion and infectivity of the asymptomatic cases. We found that the overall infection attack rate of ZIKV could be lower than 50% after the two waves if the asymptomatic cases were equal or greater than 50% and their infectivity is low.

4.1 Introduction

Since 2013, Zika virus (ZIKV) disease has caused tremendous impact to global public health and the impact is still ongoing, due to its ability that may induce neurological illness (e.g. Guillain-Barré Syndrome, GBS) in adults and birth defects (e.g. microcephaly) in newborns.

French Polynesia A ZIKV wave hit French Polynesia from October 2013 to April 2014 and induced 42 GBS cases and the risk of GBS is about 0.24 per 1,000 ZIKV infections [76], when the ZIKV infection attack rate was assumed 66%. This risk is 0.32 per 1,000 ZIKV infections when the ZIKV IAR was adjusted to 49% according to a recent serological study [54]. It was found that the asymptomatic : symptomatic ratio is about 1:1 in general population and 1:2 among school children [54], and these findings are different from a previous study in the 2007 ZIKV outbreak in Yap Island during 4.4:1 [93]. The overall attack rate of the 2007 Yap Island ZIKV outbreak was 1.46% [93].

Following the ZIKV wave, a Chikungunya virus (CHIKV) wave with an estimated 66,000 cases hit French Polynesia from October 2014 to March 2015, and induced 9 GBS cases [186]. A crude risk of GBS is 0.136 per 1,000 CHIKV infections. Thus a ZIKV infection is 2.35-fold likely to induce to GBS compared to a CHIKV infections based on the studies in French Polynesia. No epidemiological cluster of GBS induced by dengue outbreaks had been reported to our knowledge.

Northeast Brazil The Northeast (NE) region of Brazil is the hardest-hit region in America during the 2015-16 ZIKV outbreak. Three mosquito-borne infections (dengue, ZIKV and CHIKV) co-circulated and weekly cases are available, so are weekly cases of GBS and microcephaly in NE Brazil [89]. In particular, two waves of ZIKV disease were accompanied by two waves of reported GBS cases, which suggests an epidemiological association. One wave of microcephaly cases with a 23-week delay to the first ZIKV wave was reported. The delay was due to that ZIKV infection in the first trimester are most likely to induce microcephaly [67, 77, 96, 185]).

A substantial CHIKV wave was also observed during the second ZIKV wave in 2016. CHIKV could induce GBS with a smaller risk ratio (1 to 2.35) as ZIKV according to previous

studies [95, 156, 186, 226, 235]. Moreover, a study conducted in Rio de Janeiro (the largest city in the East region of Brazil) among 345 pregnant women (with rash observed) [67] revealed that the infection attack rate (IAR) of CHIKV was about 16.87%, and, in contrast, IAR of ZIKV is 53% according to PCR testing results. In addition, strong cross-protection between ZIKV and CHIKV was observed, but no cross-protection was observed between ZIKV and DENV. Thus we suspected that the two wave of excess GBS in NE Brazil was mainly due to ZIKV outbreaks, due to two reasons: (i) ZIKV is 2.35-fold likely to induce GBS than CHIKV; (ii) ZIKV IAR could be three times higher than that of CHIKV if we used Rio de Janeiro study to project the situation in NE Brazil.

The ZIKV cases were almost surely underreported before February 2016 [39]. This can also be observed from the high number of microcephaly cases (with around 23-week delay) and GBS cases in contrast to a relatively small number of reported ZIKV cases in 2015 in NE Brazil. If one ignores the possible regional difference and adopts the 0.032% GBS-ZIKV risk ratio from French Polynesia 2013-14 [76], one could immediately infer the total cases of ZIKV given the reported GBS cases for 2015 and conclude that the ZIKV cases were heavily underreported by whatever stretch in 2015.

The reporting policy of ZIKV had changed in NE Brazil. The microcephaly cases were also reported with an inconstant reporting criteria [225]. Thus it is not sensible to estimate Zika infection based on reported microcephaly cases.

Alternatively, if we assume the GBS-ZIKV risk ratio is an unknown constant for NE Brazil, given the two waves of weekly GBS cases, we could fit a ZIKV model and infer this ratio. The reporting of GBS was most likely accurate due to its medical (patients are paralyzed) features [25]. ZIKV was the major source of the excess GBS reports during the most parts of the two waves [89]. Since the co-circulation and similar symptoms of dengue fever and ZIKV in the two waves, misdiagnoses of ZIKV could happen [67, 89, 185], but DENV induced GBS was not reported in these two waves.

In this chapter Mathematical modelling simulation provides a possible way to infer the epidemic waves of ZIKV (or together with a small proportion of CHIKV). Thus, it is reasonable to use weekly excess GBS to infer weekly ZIKV infection and the overall IAR of ZIKV. First, we assume a constant risk ratio between a symptomatic ZIKV case and a reported

GBS (ZIKV-GBS ratio in short), denoted as ρ. Namely ρ proportion of the symptomatic ZIKV infections will lead to reported GBS cases. Second, we simulate our ZIKV model and fit the model to observed GBS cases and we allow the transmission rate to vary. Using Iterated Filtering technique, we yield the maximum likelihood estimate of ρ and the overall IAR of ZIKV during the two waves.

In this chapter, we use the excess GBS cases which are obtained by removing the normal level of GBS from the raw GBS cases (see details in subsection: Data). As these excess GBS cases were most likely due to the new factor: ZIKV infection (and to a small extend CHIKV infection).

4.2 Data and Methods

4.2.1 Epidemic Data

The reported weekly excess (or surplus) GBS cases from NE Brazil, between Jan 2015 and Nov 2016, are kindly provided by Professor Oliveira which was used in [89] (see Fig 4.1). We observe the GBS-to-ZIKV ratio of 2016 became significantly lower than of 2015, which is likely due to the under-report of ZIKV epidemic in 2015 [39].

4.2.2 Methods

In previous works [111, 132], we have developed a ZIKV transmission model, which includes both hosts and vectors, and considers both mosquito-borne transmission and sexual (human-human) transmission, but we assumed the asymptomatic cases are not infectious. In this chapter, we allow the asymptomatic cases to be infectious and study the impact on the estimation of IAR. We applied the plug-and-play likelihood-based inference framework (see details in Model Framework below). The basic reproduction number ($\mathcal{R}_0$) of the model is derived and estimated. The Partial Rank Correlation Coefficients (PRCC) are adopted for sensitivity analyses.

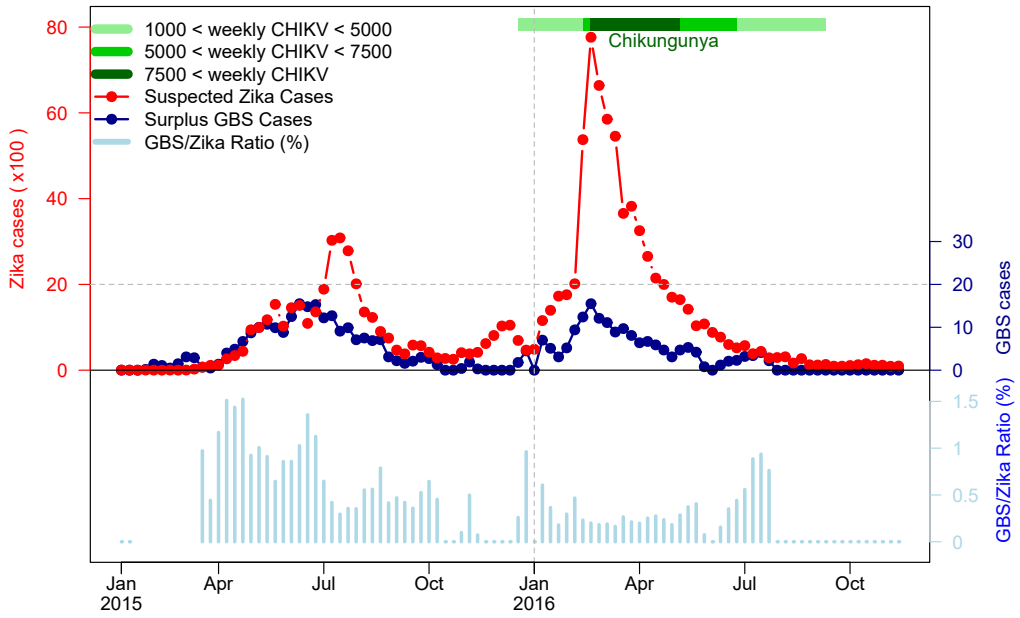


Figure 4.1: The suspected (or reported) ZIKV disease cases, excess (or surplus) GBS cases and GBS-to-ZIKV ratio in the NE region of Brazil from Jan 2015 to Nov 2016. The red dotted line represents weekly ZIKV disease cases, the dark blue dotted line represents weekly surplus GBS cases and the light blue bars are GBS-to-ZIKV ratios. The "major" (with weekly cases over 1000) chikungunya virus (CHIKV) disease outbreak of 2016 are shaded in green regarding to CHIKV disease level. The light green area denotes time periods when the weekly reported CHIKV cases were between 1000 and 5000, green denotes weekly reported CHIKV cases between 5000 and 7500 and dark green denotes weekly reported CHIKV cases over 7500. The GBS-to-ZIKV ratios are not plotted out for the start few weeks due to the scale of ZIKV data are not large enough to compute the ratio.

ZIKV-GBS Model

As in [111], we continue to assume that infected hosts are infectious during the convalescent stage to other susceptible hosts through sexual transmission but noninfectious to susceptible vectors [54, 119, 180]. Fig 4.2 shows the model diagram of the disease transmission paths. Different from our previous model, instead of assuming exposed cases (latent period) to be infectious, we assume the asymptomatic cases to be infectious to a weaker extend compared to symptomatic cases but do not develop to the convalescent stage. This is reasonable since after all there was ZIKV in their blood and sero-conversion. Therefore,

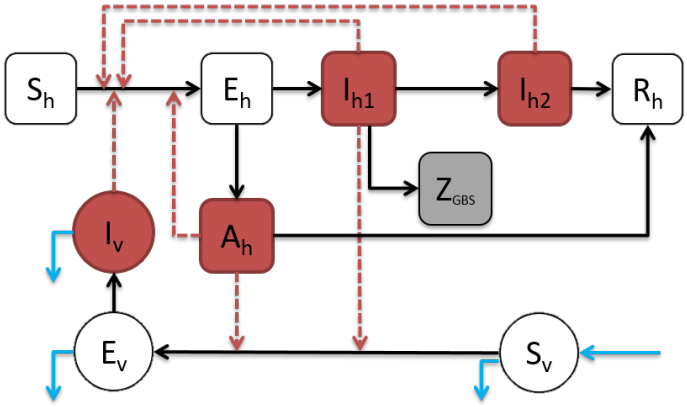


Figure 4.2: ZIKV-GBS model diagram. Black arrows represent infection status transition paths, red dashed arrows represent transmission paths and the light blue arrows represent the natural birth and death of vectors. Square compartments represent host's classes and circular compartments represent vector's classes. Red compartments represent infective classes, and gray compartment is simulated weekly excess GBS cases (Z_{GBS}).

we have following ordinary differential equation (ODE) system (see Eqn. (4.1)).

$$
\begin{cases}
S_h' = -ab \cdot \dfrac{I_v}{N_h} S_h - \beta \cdot \dfrac{\eta A_h + I_{h1} + \tau I_{h2}}{N_h} S_h \\
E_h' = \left(ab \cdot \dfrac{I_v}{N_h} + \beta \cdot \dfrac{\eta A_h + I_{h1} + \tau I_{h1}}{N_h} \right) S_h - \sigma_h E_h \\
A_h' = (1-\theta) \cdot \sigma_h E_h - \gamma_h A_h \\
I_{h1}' = \theta \cdot \sigma_h E_h - \gamma_{h1} I_{h1} \\
I_{h2}' = \gamma_{h1} I_{h1} - \gamma_{h2} I_{h2} \\
R_h' = \gamma_h A_h + \gamma_{h2} I_{h2} \\
Z_{\text{GBS}}^{(i)} = \displaystyle\int_{\text{week } i} \rho \gamma_{h1} I_{h1} \, dt \\
\\
S_v' = B_v(t) - ac \cdot \dfrac{\eta A_h + I_{h1}}{N_h} S_v - \mu_v S_v \\
E_v' = ac \cdot \dfrac{\eta A_h + I_{h1}}{N_h} S_v - (\sigma_v + \mu_v) E_v \\
I_v' = \sigma_v E_v - \mu_v I_v
\end{cases}
\tag{4.1}
$$

where, S_h is susceptible host class, E_h is the host within ZIKV infection latent period, A_h denotes the asymptomatic host class, I_{h1} denotes the infectious host class, I_{h2} denotes the convalescent host class, R_h denotes recovered class and $Z_{\text{GBS}}^{(i)}$ denotes the simulated weekly excess reported GBS cases (or surplus) for the ith week during the study period. S_v is susceptible vector class, E_v is the vector within zika infection latent period and I_v denotes

infectious vector class. Here, parameter ρ denotes the ratio of reported excess GBS cases per symptomatic ZVD infections. The model (see Eqn. (4.1)) parameters are summarized in Table 4.1.

Table 4.1: Summary table of model parameters. "H" denotes hosts and "V" denotes vectors. "X→Y" denotes ZVD infectious X infects susceptible Y.

Parameter	Notation	Value/Range	Description/Unit	Source
Mosquito biting rate	a	0.3 - 1.0	per vector·day	[51]
Transmission prob. of host	b	0.10 - 0.75	per bite	[51]
Transmission prob. of vector	c	0.30 - 0.75	per bite	[80]
Transmission rate by contact	β	0.001 - 0.10	per day	[111]
Host latent period	σ_h^{-1}	2 - 7	days	[44, 63]
Vector latent period	σ_v^{-1}	8 - 12	days	[51, 65]
Asymptomatic infectious period	γ_h^{-1}	5 - 10	days	assumed
Infectious period	γ_{h1}^{-1}	3 - 7	days	[63]
Convalescent infectious period	γ_{h2}^{-1}	14 - 30	days	[119, 180]
Proportion of symptomatic	θ	(50%) 20% - 80%	Nil	[54]
Relative infectivity of asymptomatic	η	0.0 - 0.99	H→H & H→V, Nil	to be estimated
Relative infectivity of convalescent	τ	(0.3) 0.01 - 0.99	H→H, Nil	[111]
Vector lifespan	μ_v^{-1}	4 - 35	days	[51, 80]
Ratio: $\frac{\text{reported GBS}}{\text{symptomatic ZVD}}$	ρ	0.0075% - 0.05%	Nil	[54, 76]
Initial host susceptible proportion	$S_h.0/N_h$	0.75 - 0.9999	Nil	[111]

In addition,

$$\begin{aligned} N_h &= S_h + E_h + A_h + I_{h1} + I_{h2} + R_h \\ N_v &= S_v + E_v + I_v \end{aligned}$$

where, N_h and N_v represent the total number of hosts and vectors respectively, of which N_v is time-dependent. The population of the Northeastern region of Brazil in 2011 was 53.6 million [40].

For the mosquito population, as we assume mosquitoes population ($N_v(t)$)

$$N_v = m(t) \cdot N_h \tag{4.2}$$

where $m(t)$ is the time-dependent ratio of mosquitoes population to humans population. We realize this but increasing the susceptible (through birth) when $m(t)$ is increasing, and removing all susceptible and infectious (through vector control measures) mosquitoes when $m(t)$ is decreasing. $m(t)$ is realized as an exponential cubic-spline function with certain nodes which are uniformly distributed over the study period.

Basic Reproduction Number

Following previous studies [68, 111, 221], we derive the basic reproduction number ($\mathcal{R}_0$) formula by next generation matrix method,

$$\mathcal{R}_0 = \frac{\mathcal{R}_{hh} + \sqrt{\mathcal{R}_{hh}^2 + 4\mathcal{R}_{hv}^2}}{2} \tag{4.3}$$

where,

$$\mathcal{R}_{hh} = \beta \cdot \left[\eta \cdot \frac{1-\theta}{\gamma_h} + \theta \cdot \left(\frac{1}{\gamma_{h1}} + \frac{\tau}{\gamma_{h2}} \right) \right] \quad \text{and} \quad \mathcal{R}_{hv} = a \cdot \sqrt{bcm \cdot \frac{\theta\gamma_h + (1-\theta)\eta\gamma_{h1}}{\gamma_h\gamma_{h1}} \cdot \frac{\sigma_v}{\mu_v \cdot (\mu_v + \sigma_v)}}$$

where m is the ratio of mosquitoes to humans population (see Eq 4.2).

Model Framework

The ZIKV epidemics in NE Brazil is modelled as a Partially Observed Markov process (POMP, also known as hidden Markov model), and we deploy Iterated Filtering and plug-and-play likelihood-based inference frameworks to fit reported excess GBS cases [111, 133, 147]. The Maximum Likelihood Estimates (MLE) for model parameters are adopted. `R` package "`POMP`" is available via [146].

The mosquito abundance (m) is assumed to be unknown but variable over our study period, and reconstructed. Given the mosquito abundance ($m(t)$) is time-dependent, we allow the basic reproduction number ($\mathcal{R}_0(t)$) also to be time-dependent. The parameter fitting and inference process are rigorously and exhaustively checked. Therefore, the fits of observed time-series are accurate with high confidence because of the consistency with true underlying epidemiological processes rather than because of artificial model over-fitting.

Bayesian Information Criterion (BIC) is employed as a criterion for model comparison, and quantifies the trade-off between the goodness-of-fit of a model and its complexity [205]. The simulations were conducted deploying the Euler-multinomial integration method with the time-step fixed to be one day [48, 133].

The simulated weekly reported excess GBS cases (Z_{GBS}) are modelled by Eq 4.1. The corresponding weekly observed GBS cases of the ith week, $C_{\text{GBS}}^{(i)}$, are assumed to follow a

Negative-Binomial (NB) distribution as

$$C_{\text{GBS}}^{(i)} \sim \text{NB}\left(n = \frac{1}{\tau}, p = \frac{1}{1+\tau Z_{\text{GBS}}^{(i)}}\right) \qquad \text{with} \quad \text{mean} : \mu_i = Z_{\text{GBS}}^{(i)} \tag{4.4}$$

where τ denotes an over-dispersion parameter that needs to be estimated.

Finally, the overall log-likelihood function, l, is given by

$$l\left(\Theta | C_{\text{GBS}}^{(1)}, \ldots, C_{\text{GBS}}^{(N)}\right) = \sum_{i=1}^{T} \ln L_i \tag{4.5}$$

where Θ denotes the parameter vector under estimation, and L_i is the probability measurement functions associated with $C_{\text{GBS}}^{(i)}$ vs. $Z_{\text{GBS}}^{(i)}$. T denotes the total number of weeks during the study period.

We assume that $m(t)$ is an exponential cubic spline function of time with number of nodes, n_m over the study period. Nodes are distributed evenly over the time-domain with values (m_i) that are estimated but restricted between 0 and 20. The range was selected according to that, as $m = 20$, $\mathcal{R}_0 = 8.5$ (which is higher than the reasonable range of ZIKV disease's basic reproduction number [108, 111, 152, 166]) by fixing $\theta = 0.5$, $\tau = 0.3$ and $\eta = 0.3$ (see Table 4.1).

The confidence intervals (C.I.) of parameters are estimated based on parameters' ranges in Table 4.1, using the method of profile likelihood confidence intervals [132, 133]. Parameter estimation and statistical analysis are conducted using `R` (version 3.3.3).

Sensitivity Analysis

The Partial Rank Correlation Coefficients (PRCC) are adopted for the model's sensitivity analysis [111]. Firstly, 1,000 random samples are taken for each model parameter from uniform distributions with parameter ranges as set out in Table 4.1. Secondly, for every random parameter sample set, the ZVD-GBS model was simulated to obtain the target biological quantities (e.g., $\mathcal{R}_0$ and the total number of GBS cases in this chapter). Finally, PRCCs were calculated between each parameter and target biological quantity.

4.3 Results

4.3.1 Results of Infection Attack Rate

The infection attack rate (IAR) of ZIKV are estimated regarding to total 21 pairs of (η, θ) (i.e., three different values of θ and seven different values of η). Under high proportion of symptomatic scenario (i.e., $\theta = 0.8$), the estimated range of IAR is from 35.82% to 48.19% (see Fig 4.3a). Under moderate proportion of symptomatic scenario (i.e., $\theta = 0.5$), the estimated range of IAR is from 10.26% to 49.02% (see Fig 4.3b). Under low proportion of symptomatic scenario (i.e., $\theta = 0.2$), the estimated range of IAR is from 0.01% to 15.33% (see Fig 4.3c).

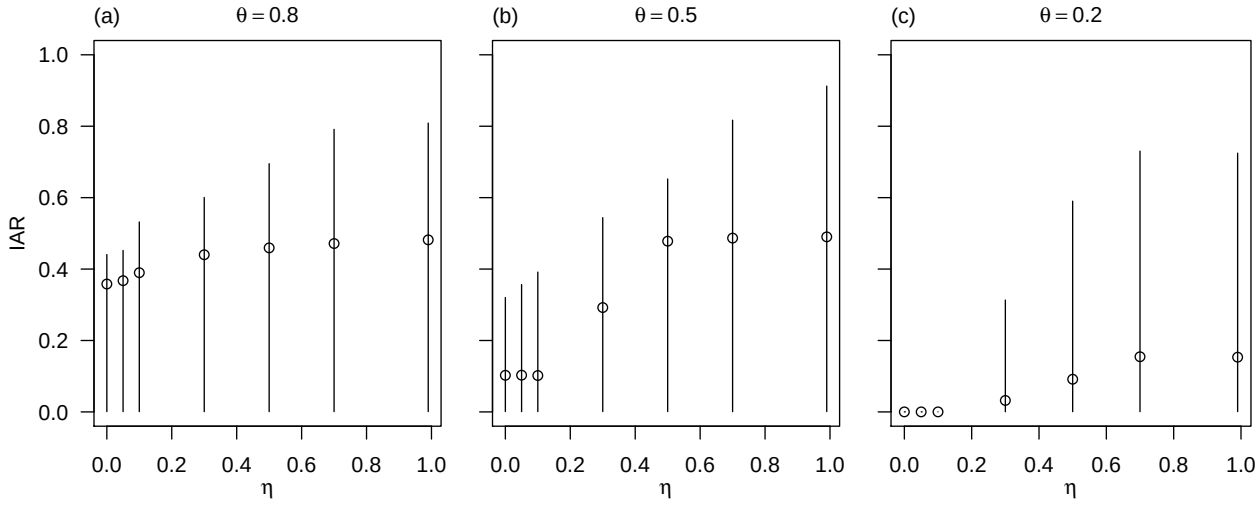


Figure 4.3: Infection attack rate (IAR) as a function of the relative infectivity of asymptomatic cases (η) under three scenarios with different symptomatic ratio (θ) fixed to be 0.8 in panel (a), 0.5 in panel (b) and 0.2 in panel (c).

With the proportion of symptomatic infections (θ) fixed, the estimated IAR is increasing (except the result of η from 0.05 to 0.1 and $\theta = 0.5$, which could be due to simulation error) as η increases. Similarly, the width of confidence intervals (CI) of IAR are also increasing as η becoming larger. Interestingly, we can observe from Fig 4.3 that, given θ fixed, there appears to be a sub-range of η such that IAR increasing faster (i.e., the increasing rate of IAR is larger) than other η values. In this chapter, we simply name this observed sub-range of η to be "η-IAR increasing interval". For $\theta = 0.8$, η-IAR increasing interval is from 0.05 to 0.3. For $\theta = 0.5$, η-IAR increasing interval is from 0.1 to 0.5. For $\theta = 0.2$, η-IAR increasing interval is from 0.3 to 0.7.

Table 4.2: Summary of parameter estimates. n_m, θ, and η are fixed. Since BIC attains minimum when $n_m = 6$ with $\eta = 0.5$, $n_m = 6$ is used for all simulation. AR stands for infection attack rate of ZVD. ρ stands for the ratio of reported GBS cases to symptomatic infections. MLE stands for maximum log likelihood.

n_m	θ	η	IAR	95% CI	ρ	95% CI	MLL
6	0.8	0.99	0.4819	(9e-04 , 0.8087)	1.9e-05	(1.1e-05 , 0.010298)	-160.653
6	0.8	0.7	0.4715	(8e-04 , 0.7913)	1.9e-05	(1.1e-05 , 0.011421)	-160.684
6	0.8	0.5	0.4594	(7e-04 , 0.6951)	1.9e-05	(1.2e-05 , 0.011421)	-160.772
6	0.8	0.3	0.4401	(6e-04 , 0.6004)	1.9e-05	(1.4e-05 , 0.012667)	-160.775
6	0.8	0.1	0.3899	(7e-04 , 0.5319)	1.9e-05	(1.4e-05 , 0.010298)	-160.777
6	0.8	0.05	0.3676	(7e-04 , 0.4521)	1.9e-05	(1.5e-05 , 0.009285)	-160.72
6	0.8	1.0e-10	0.3582	(8e-04 , 0.4406)	1.9e-05	(1.5e-05 , 0.008372)	-160.812
6	0.5	0.99	0.4902	(0.001 , 0.9124)	3.1e-05	(1.7e-05 , 0.015581)	-160.712
6	0.5	0.7	0.4868	(7e-04 , 0.8169)	3.1e-05	(1.9e-05 , 0.021257)	-160.857
6	0.5	0.5	0.4781	(0.001 , 0.6523)	3.1e-05	(2.3e-05 , 0.015581)	-160.895
6	0.5	0.3	0.2923	(9e-04 , 0.544)	5.2e-05	(2.8e-05 , 0.017281)	-161.002
6	0.5	0.1	0.1019	(8e-04 , 0.3916)	0.000148	(3.8e-05 , 0.019166)	-161.224
6	0.5	0.05	0.1029	(7e-04 , 0.3566)	0.000148	(4.3e-05 , 0.021257)	-161.601
6	0.5	1.0e-10	0.1026	(9e-04 , 0.3204)	0.000148	(4.7e-05 , 0.017281)	-162.062
6	0.2	0.99	0.1533	(8e-04 , 0.7246)	0.000248	(5.2e-05 , 0.048666)	-160.65
6	0.2	0.7	0.1545	(7e-04 , 0.7303)	0.000248	(5.2e-05 , 0.053975)	-160.808
6	0.2	0.5	0.0915	(7e-04 , 0.5902)	0.000416	(6.4e-05 , 0.053975)	-160.91
6	0.2	0.3	0.0321	(5e-04 , 0.3133)	0.001171	(0.00012 , 0.073637)	-162.159
6	0.2	0.1	1e-04	(1e-04 , 2e-04)	0.282915	(0.186978 , 0.282915)	-177.93
6	0.2	0.05	1e-04	(1e-04 , 2e-04)	0.282915	(0.229998 , 0.282915)	-188.368
6	0.2	1.0e-10	1e-04	(1e-04 , 1e-04)	0.282915	(0.282915 , 0.282915)	-197.022

The range corresponding basic reproduction number ($\mathcal{R}_0$) for all pairs (i.e., total 21 pairs) of (η, θ) is from 1.21 to 2.95, which is computed by fixing the average ratio of mosquitoes to humans as $m = 2.0$ (which is in the range according to [108]).

Range of Basic Reproduction Number The range of basic reproduction number of ZVD ($\mathcal{R}_0$ is from 1.21 to 2.95) is biologically reasonable regarding to each pair of (η, θ) (see Table 4.3). The average mosquitoes to humans ratio is fixed to be $m = 2.0$.

4.3.2 Robustness of IAR Estimation

If only focusing on the time period from Feb to Aug, 2016 (when the official ZIKV reporting program has already launched [39]), and further provided CHIKV-GBS : ZIKV-GBS = 1 : 2.35 (note that the number of CHIKV cases was roughly 2.12-multiple of the number

Table 4.3: Table of ZIKV disease $\mathcal{R}_0$ regarding to each pair of (η, θ).

	θ		
η	0.2	0.5	0.8
1e-10	1.211010	1.968412	2.539856
0.05	1.369865	2.035759	2.561801
0.1	1.512870	2.101087	2.583575
0.3	1.989813	2.345526	2.669037
0.5	2.379394	2.568616	2.752072
0.7	2.718815	2.775473	2.832893
0.99	3.153805	3.053182	2.946523

of ZIKV cases, the data of CHIKV cases can also be found in SI of [89]), one can have a intuitive calculation of symptomatic ZIKV induced GBS rate (or risk of symptomatic ZIKV infection to GBS, i.e., $\rho = \frac{G_Z}{Z}$) by solving:

$$\begin{cases} \dfrac{G_C + G_Z}{Z \cdot r_Z} = 0.3040\% \\ \dfrac{G_C}{C} : \dfrac{G_Z}{Z} = 1 : 2.35 \\ \dfrac{C \cdot r_C}{Z \cdot r_Z} \approx 2.12 \end{cases} \tag{4.6}$$

where G_C and G_Z is the number of GBS cases induced by CHIKV and ZIKV respectively, C and Z are the actual number of symptomatic CHIKV and ZIKV infections respectively and r_C and r_Z are the reporting ratio of symptomatic CHIKV and ZIKV infections (i.e., reported cases over total symptomatic infections) respectively. The 0.3040% is the approximated GBS-to-ZKIV ratio from Fig 4.1 (see light blue bars in 2016). Solving Eqn. (4.6), we have:

$$\rho = \frac{G_Z}{Z} \approx 0.3040\% \cdot r_Z \cdot \frac{2.35 r_C}{2.12 r_Z + 2.35 r_C}, \qquad r_Z, r_C \in (0, 1). \tag{4.7}$$

Fig 4.4 shows the relationship (see Eqn. (4.7)) between symptomatic ZIKV induced GBS rate ($\rho = \frac{G_Z}{Z}$) and the reporting rate of symptomatic CHIKV (r_C) and ZIKV (r_Z).

One reasonable assumption is that we consider the reporting rate of symptomatic CHIKV and ZIKV to be the same (i.e., $r_C = r_Z = r$, then the corresponding values of ρ are along the diagonal of Fig 4.4) in NE Brazil, then, we can simply derive a theoretical relationship among IAR, symptomatic (ZIKV and/or CHIKV) reporting rate (r) and weekly

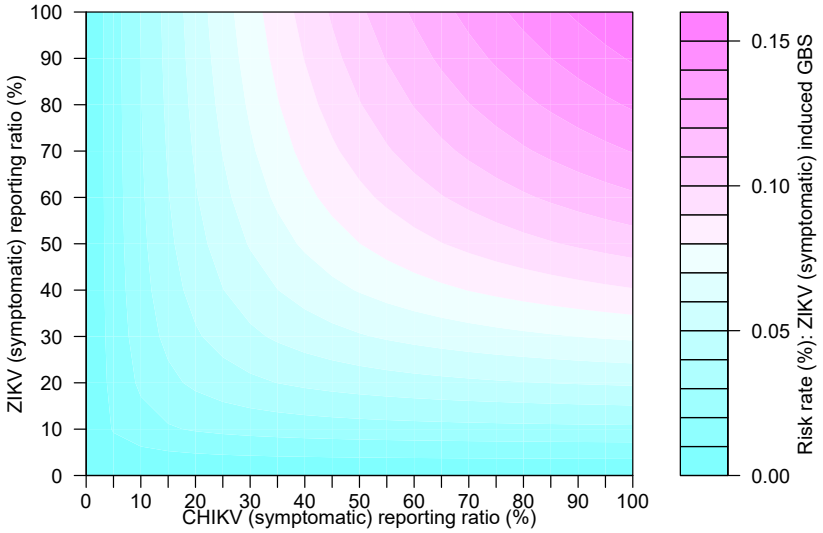


Figure 4.4: The contour plot of the relationship between symptomatic ZIKV induced GBS rate ($\rho = \frac{G_Z}{Z}$) and the reporting rate of CHIKV (r_C) and ZIKV (r_Z). The axises' labels are all presented in percentage scale (%).

reported (ZIKV+CHIKV) cases.

$$53.6 \times 10^6 \times \text{IAR} \times \theta r = 87 \times \overline{\text{cases}}$$

where 53.6×10^6 is the population of NE Brazil according to [40] and 87 is the total number of weeks from Jan, 2015 to Aug, 2016 and $\overline{\text{cases}}$ is the average number of weekly reported (ZIKV+CHIKV) cases. One can directly calculate from raw ZIKV and CHIKV reported data (see Fig 4.1 and Supplementary Information of [89]) that $\overline{\text{cases}} = 6123.45$ for time period from Feb to Aug, 2016. With ZIKV symptomatic ratio to be 50% (i.e., $\theta = 0.5$) [54], we have:

$$\begin{aligned} 53.6 \times 10^6 \times \text{IAR} \times 50\% \times r &= 87 \times 6123.45 \\ \text{IAR} \times r &= 0.0199 \end{aligned} \tag{4.8}$$

With Eqn. (4.8), one can find an explicit relation between ρ and IAR according to Eqn. (4.7) (under assumption: $r_C = r_Z = r$). If we choose the situations that the reduce infectivity of asymptomatic ZIKV infections to be mildly weaker than symptomatic infections (i.e., $\eta \in [0.1, 0.3]$), which the recent Gao *et al.*'s work deployed similar treatment with respected to the reduce infectivity (see the parameter list of [111]), the corresponding IAR is from 10.19% to 29.23%. We can further derive from Eqn. (4.8):

$$r \in [6.81\%, 19.53\%] \tag{4.9}$$

The reporting ratio results are consistent with Kucharski *et al.*'s previous modeling study of 2013-14 ZIKV outbreak in French Polynesia (8%-22%) [152]. Combining Eqn. (4.7) and 4.9 (note that we have further treat $r_C = r_Z = r$ in Eqn. (4.7)), the symptomatic ZIKV induced GBS rate (ρ) is appropriated to be $\rho \in [0.0109\%, 0.0312\%]$, which is in line with one recent metadata study [76] and one recent serological study [54]. Backwardly speaking, given the symptomatic ZIKV induced GBS rate (ρ) from 0.0075% to 0.05% according to [54, 76] (see Table 4.1), one can similarly derive the IAR to be $\rho \in [6.36\%, 42.41\%]$ which is consistent with our model estimation (see Table 4.2 for detail). Therefore, with reasonable selection of parameters' values (i.e., $\theta = 0.5$ and $\eta = 0.1$ to 0.3), we robust our IAR estimation results.

4.3.3 Sensitivity Analysis Results

Results of the sensitivity analysis are presented in Fig 4.5, which indicates how model parameters impact the basic reproduction number ($\mathcal{R}_0$) and the total reported GBS cases. $\mathcal{R}_0$ is most sensitive to vectors' biting rate (a), vectors to hosts ratio (m) and the vectors' lifespan (μ_v^{-1}, or vectors' natural death rate, μ_v), indicating the importance of the mosquitoes role in disease transmission. The total reported GBS cases are considerably sensitive to the proportion of symptomatic cases (θ), the ratio (or risk) of excess GBS cases to symptomatic ZIKV infections (ρ).

4.4 Discussion

In this chapter, we showed that the asymptomatic infections and their infectivity can be very important in the estimation of infection attack rate (IAR). The proportion of asymptomatic Zika virus disease (ZVD) infections out of all ZVD infections is chosen to be 50%, based on a recent serological study [54].

Wang *et al.*'s recent genetic study suggests that the ZIKV strains in Brazil are genetically close to the ZIKV strains of past outbreak in French Polynesia, the ZIKV strains of both regions are classified into Asian lineage (see Fig 1 of [228]).

The estimated IAR is sensitive to both the proportion of asymptomatic ZIKV infections as well as their relative infectivity. If the proportion of symptomatic ZIKV infections (θ)

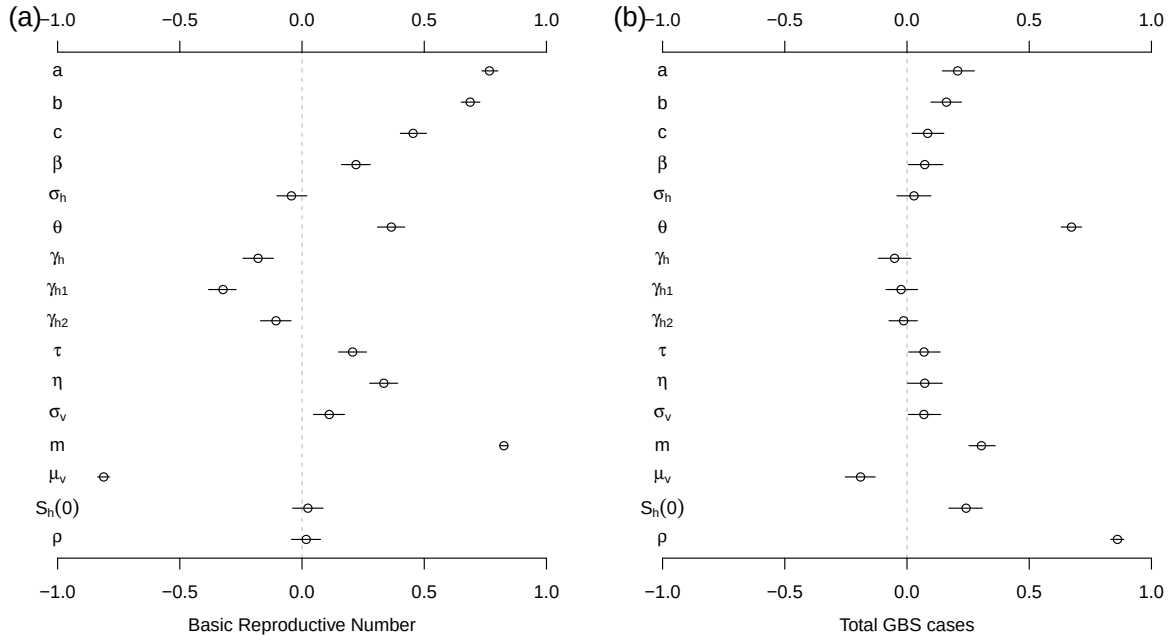


Figure 4.5: The Partial Rank Correlation Coefficients (PRCC) of basic reproduction number,$\mathcal{R}_0$, (panel (a)) and total GBS cases (panel (b)) with respect to model parameters. $S_h(0)$ denotes the initial susceptible ratio ($S_h.0/N_h$). The black circle is the estimated correlation and the bar represents 95% C.I.. The ranges of parameters are given in Table 4.1.

are presumed to be 50% according to recent serological study of French Polynesia [54], the IAR is estimated to be from 10.19% to 29.23% corresponding to the reduced infectivity of asymptomatic ZIKV infections (η) fixed at 0.05, 0.1 and 0.3, which also corresponds to a reasonable symptomatic ZIKV induced GBS rate's range.

"*Ae. aegypti* was competent for both viruses with transmission rates up to 73% (ZIKV) and 21% (CHIKV). A substantial proportion of mosquitoes became saliva-positive for both viruses (12%), suggesting that *Ae. aegypti* can transmit both CHIKV and ZIKV via a single bite. Additionally, co-infections did not influence the infection or transmission rates of either ZIKV or CHIKV." [117].

Previous experimental study [236] suggested that moderate saline water can be in favor of mosquitoes' oviposition, which could be the reason behind the relatively high ZIKV IAR in French Polynesia (since FP is surrounded by ocean). The saline water induced mosquito over-breed situation has been reported in Kaohsiung city, Taiwan, which resulted in a huge local dengue epidemic [38].

4.5 Conclusion

The proposed simple mathematical model can successfully reproduce the exceeding GBS pattern, and further infer the local IAR of ZVD of NE Brazil from 2015-16. The estimated IAR is sensitive to both the proportion of asymptomatic ZIKV infections as well as their relative infectivity. Since the estimated ZVD IAR is more likely to be less than 50%, there still exists likelihood of ZVD outbreak in coming years.

Chapter 5

Modelling Childhood Infection of Varicella in Shenzhen, China

Varicella (chickenpox) is a highly transmissible childhood disease. Between 2010 and 2015, it displayed two epidemic waves annually among school populations in Shenzhen, China. However, their transmission dynamics remain unclear and there is no school-based vaccination programme in Shenzhen to-date. In this chapter, we developed a mathematical model to compare a school-based vaccination intervention scenario with a baseline (i.e. no intervention) scenario.

Data on varicella reported cases were downloaded from the Infectious Disease Reporting Information Management System. We obtained the population size, age structure of children aged 15 or under, the class and school distribution from Shenzhen Education Bureau. We developed an Agent-Based Susceptible-Exposed-Infectious-Recovered (ABM-SEIR) Model that considered within-class, class-to-class and out-of-school transmission modes. The intervention scenario was that school-wide vaccination intervention occurred when an outbreak threshold was reached within a school. We varied this threshold level from five to ten cases. We compared the reduction of disease outbreak size and estimated the key epidemiological parameters under the intervention strategy.

Our ABM-SEIR model provided a good model fit to the two annual varicella epidemic waves from 2013 to 2015. The transmission dynamics displayed strong seasonality. Our results suggested that a school-based vaccination strategy could effectively prevent large outbreaks at different thresholds.

There was a considerable increase in reported varicella cases from 2013 to 2015 in Shenzhen. The proposed modelling framework provides important theoretical support for disease control decision making during school outbreaks and the development of a school-based vaccination programme.

5.1 Introduction

Varicella (chickenpox) is caused by the varicella zoster virus (VZV) of the Herpesviridae family. It spreads by direct contact and airborne droplets [134]. Varicella is highly transmissible during childhood, thus it has the potential to cause outbreaks at schools [134].

In China, varicella outbreaks pose serious public health threats to the school populations. The National Immunization Program does not cover vaccination against varicella, and they are only available as self-paid vaccines for children between one and 12 years of age [105]. Varicella uptake rate remains low in China, and most children only receive a single-dose vaccine [6], which, according to a recent meta-analysis, is only about 81% effective [167]. In October 2015, China introduced a two-child policy to replace its one child policy [244]. This policy change has led to increase in fertility rate and is expected to increase the future size of school populations. Thus, it is imminent to examine public health control strategies of varicella among schools in Shenzhen, China.

Transmission dynamics of infectious diseases had been investigated in previous studies. [160, 210, 211, 241]. Several studies had explored the impact of vaccination on varicella transmission [70, 71, 187, 223]. These studies applied a Who-Acquired-Infection-from-Whom (WAIFW) contact matrix, combined with age-specific transmission parameter following the methodology of Wallinga *et al.* [227], was primarily used empirical age-specific social contact data of European populations but failed to account for the class and school structure of student populations. Jackson *et al.* developed two mathematical models to study the effects of school holidays on the spread of varicella, and found that there were 22% to 31% reduction in student contacts during summer holidays, that led to a lower rate of varicella transmission [140]. A surveillance study was conducted at elementary schools, found school nurse surveillance and tracking of varicella cases are effective in lowering annual varicella incident cases [157].

In this chapter, we modelled the transmission dynamics of varicella among school children in 2013-2015 in Shenzhen, China. We considered two scenarios: (i) baseline (no intervention) scenario; (ii) school-based vaccination scenario, where all students within a school were vaccinated once the number of varicella cases were beyond a stated threshold. Here, an Agent-Based Susceptible-Exposed-Infectious-Recovered(ABM-SEIR) Model was developed, and we showed that reasonable modelling estimates could be achieved with this model by specifying individual level and group-level contact patterns [138, 171, 179]. This paper is structured as follows: First, we described the data source, model population and model structure. Next, we described the model parameters and model fitting. It is then followed by the estimation of reproduction number under different intervention scenarios. Lastly, we discussed the implications of our findings.

5.2 Data and Methods

5.2.1 Data

The varicella reported cases were obtained from Infectious Disease Reporting Information Management System in Shenzhen, China, on a weekly basis from 2010 to 2015. These reported cases included both clinically diagnosed cases and laboratory confirmed cases, which were voluntarily reported by local medical doctors. In Fig 5.1, we show the weekly reported cases per 1,000,000 population, and an increasing trend is observed from 2013 to 2015. The reported cases show a peak-to-trough pattern from school terms to school holidays. We also collected data on the monthly total number of school outbreaks from January 1, 2010 to December 31, 2015 from Shenzhen Center for Disease Control and Prevention (SZCDC). The school outbreaks were reported under a compulsory surveillance system. A school outbreak is defined as five or more varicella cases within a seven day period that occurs at a school or kindergarten.

In Fig 5.2, we used a locally weighted scatterplot smoothing model (LOESS) to show the monthly school outbreaks from 2010 to 2015. We found that both reported cases and school outbreaks displayed two epidemic waves annually. Apparently, the trend in weekly reported cases lagged behind the school outbreaks.

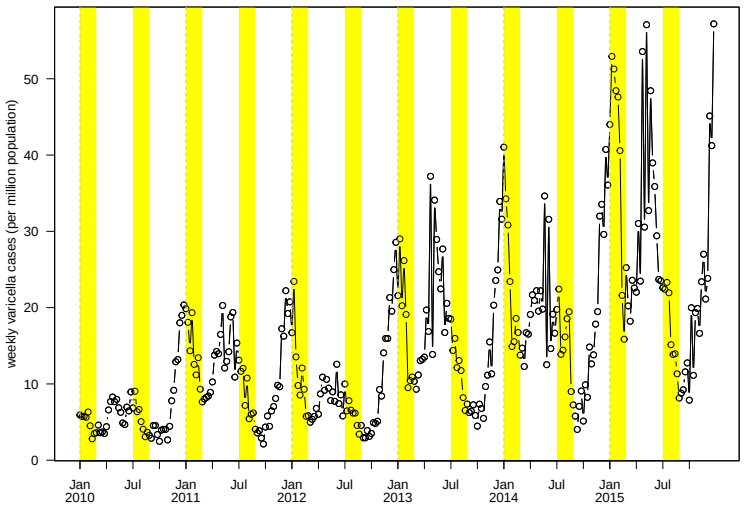


Figure 5.1: Number of weekly varicella confirmations from 2010 to 2015 per 1,000,000 population in Shenzhen from 2013 to 2015. Weekly population is computed using *loess* model. School holidays are shaded in yellow.

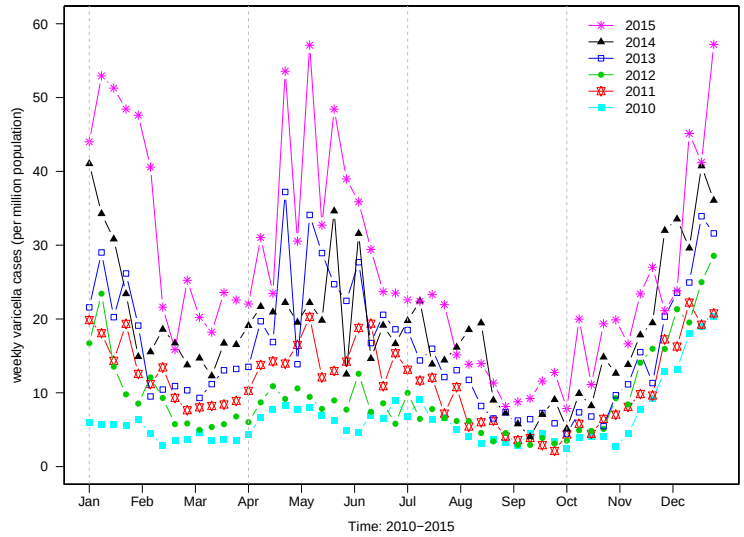


Figure 5.2: Weekly number of varicella confirmations in each year from 2010 to 2015, per 1,000,000 population. Weekly varicella cases is computed using *LOESS* model.

In Fig 5.3, we show the number of varicella school outbreaks from 2010 to 2015 in a boxplot. We could see that there are two epidemic peaks in April and November and a trough in July and August annually.

Fig 5.4 shows the distribution of varicella incidence in different districts from 2013 to 2015. We observe higher varicella incidence near Luohu which borders Hong Kong. There were substantial geographical variations. Information about the population size of those ages under 15, age structure and distribution of the number of classes and schools were downloaded from the website of Shenzhen Education Bureau [7].

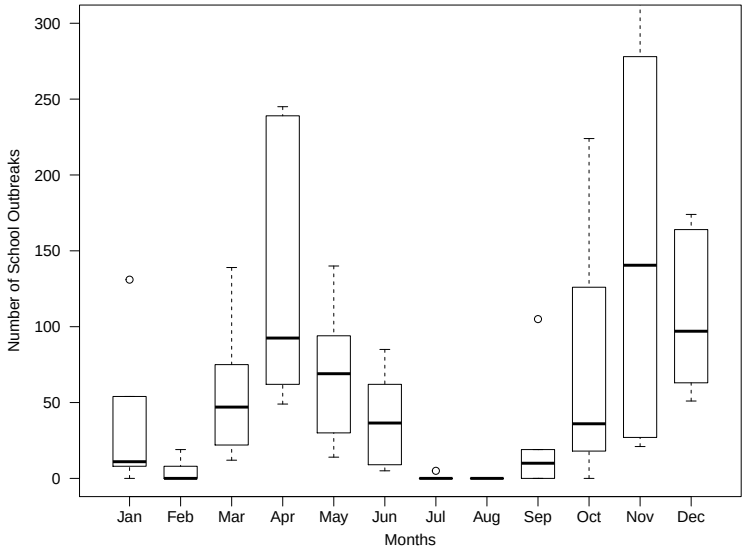


Figure 5.3: Boxplot of the number of varicella school outbreaks from 2010 to 2015, which displays similar patterns as in Fig 5.2. The number of school outbreaks per 30 days is displayed, to adjust for the variations of the number of days in each month.

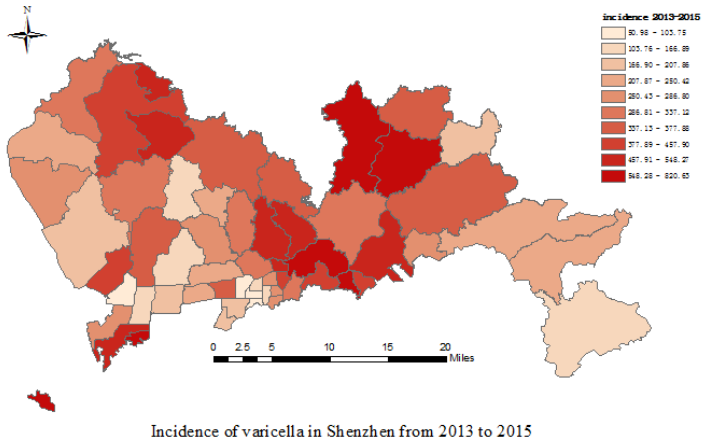


Figure 5.4: Varicella incidence distribution in Shenzhen by district from 2013 to 2015. The shade represents the levels of varicella incidence, cases are per 100,000 population within each district.

5.2.2 Target Population

Our model population consisted of individuals from 0 to 15 years old in Shenzhen, as varicella primarily affects this age range. We did not consider the effects due to Herpes Zoster caused by VZV because of its extremely low incidence within this age range. For the school populations and the students' age, class and school structure, we made adjustments including approximation and averaging to the official data [7]. Table 5.1 displays the distribution of the adjusted number of schools, classes and students by types of school.

Table 5.1: A summary table of the adjusted average number of schools, classes, and distribution of students in Shenzhen. Students per class, $N_{j,i}$, was given by Eqn. (5.2)). The information was obtained from Shenzhen Education Bureau [7].

Age groups (years)	0 - 3	4 - 6	7 - 12	13 - 15
Student status	Pre-school	Kindergarten	Primary school	Secondary school
Schools in Shenzhen	300	1500	550	250
Classes per school	25	10	24	20
Students per class: $N_{j,i}$	20	30	45	60
Proportion: p_j	0.1004	0.3012	0.3976	0.2008

Note: The age group from 0 to 3 years consist of mainly pre-school children. Therefore, "schools in Shenzhen" actually reflects the number of street blocks, "classes in school" represent the number of communities per street block and "students per class" refers to the number of children in that age group within each community [138].

5.2.3 Model Structure

We developed an ABM-SEIR model for school students in Shenzhen. The overall model structure could be conceptualized as follows: students are nested within classes, classes within schools, and schools within Shenzhen's school students population. A classical SEIR compartmental model was fitted to each class, while considering the different age structures and grade levels.

$$\text{Students} \in \text{Classes} \subset \text{Schools} \subset \text{Shenzhen}$$

Within-class Transmission

As there were frequent social contacts and interactions with other classmates during a school day, each class was treated as a group-level unit for human-to-human transmission. Thus we applied a SEIR model to each class, and the classes were expressed as the following set of non-linear ordinary differential equations (ODE):

$$\begin{aligned} \frac{dS}{dt} &= -\theta(a) \cdot \beta SI \\ \frac{dE}{dt} &= \theta(a) \cdot \beta SI - \sigma E \\ \frac{dI}{dt} &= (1-\eta) \cdot \sigma E - \gamma I \\ \frac{dR}{dt} &= \eta\sigma E + \gamma I \end{aligned} \tag{5.1}$$

Here, S, E, I and R denoted the number of Susceptible, Exposed, Infected and Recovered individuals respectively. The total number of students in each class was given by:

$$N_{j,i} = S + E + I + R \tag{5.2}$$

where, j denotes the jth school and i denoted the ith class within the jth school.

The other parameters were as follows: average transmission rate (β), average infectious rate (σ), average recovery rate (γ) and beta multiplier ($\theta(a)$), the last of which was dependent on the student's age (a). η was the average rate of losing infectiousness due to hospitalization, medical treatment or contact isolation [158].

We did not consider birth and death processes in the model, since our study period was relatively short compared with the average lifespan. The epidemiological effects of seasonal oscillations in birth rates were negligible [130]. Furthermore, once a student recovers from varicella, he or she would be immunized for 20 to 40 years, which is much longer than our study period. In effect, students entering the Recovered (R) compartment left the system.

Class-to-class Transmission

The next level of transmission would be class-to-class varicella transmission which involved social contacts and mixing of students between different classes. Such activities include assembly gathering, having meals at school canteens, taking school buses and attending extra-curricular activities. We adopted the same operational definition by SZCDC and previous studies [6, 70, 157], where an outbreak threshold was reached when there were five or more varicella cases within a class, i.e. $I_{\text{limit}} = 5$ cases.

The spread rate, δ, between classes was low under current disease control measures. Otherwise, large outbreaks could occur among school populations.

Out-of-school Transmission

The third level of transmission would be out-of-school transmission due to student contacts between different schools. Inter-school activities and private group tutorials would be examples of such. Also, Shenzhen is a popular city with many tourists, businessmen and

students visiting every year. Thus it was important to consider imported varicella cases to Shenzhen. However, the imported rate, τ, would be relatively low because out-of-school transmission was not a predominant transmission route in our model. Fig 5.2 shows obvious annual periodicity of the weekly reported varicella cases. These patterns were especially remarkable from 2013 to 2015. Previous studies attributed these to the seasonality of school terms [102, 131, 171]. We incorporated this factor in to the ABM-SEIR. Fig 5.5 shows the schematic diagram for the ABM-SEIR describing within-class transmission, class-to-class transmission and out-of-school transmission within Shenzhen's school age populations.

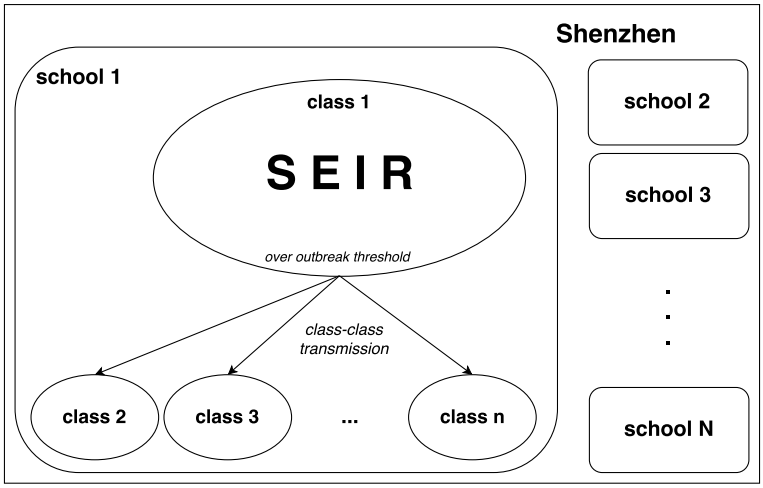


Figure 5.5: The structural diagram of the ABM-SEIR in Shenzhen. Within each classes, SEIR model structure is applied (see ODE system (5.1)). Within a school, if a class reaches the pre-defined outbreak threshold, there will be possible disease transmission to non-outbreak classes, to which we name "class-class transmission". This transmission will vanish whenever the number of cases in the outbreak classes becomes lower than the outbreak threshold.

5.2.4 Model Parameters

Del Fava *et al.* [90] found that varicella transmissibility would be the strongest within the youngest age groups. Thus we applied a beta multiplier (θ) to represent the relative transmissibility within each age group (See Table 5.2)

Table 5.2: Table of beta (or β) multiplier (θ) with respect to different age groups

Age groups (year)	0 - 3	4 - 6	7 - 12	13 - 15
beta multiplier (θ)	0.625	1.000	0.750	0.500

The transmission rate, β, is defined as the probability of a susceptible to be infected after an effective contact with one infectious individual. This is a time-dependent function. Class-to-class transmission rate, δ, is estimated according to the best-fitted varicella transmission model. Table 5.3 summarizes the list of model parameters:

Table 5.3: Summary table of parameters

Parameter	Notation	Value	Source
Latent period	σ^{-1}	14 (day)	[139]
Infectious period	γ^{-1}	7 (day)	[139]
Transmission rate	β	to be estimated	-
Beta multiplier	θ	Table 5.2	[90]
Initial immune percentage	R_0	65.00%	[223]
Initial infectious percentage	I_0	0.05%	[158]
Initial exposed percentage	E_0	0.00%	assumed
Initial susceptible percentage	S_0	34.95%	$[1-(E_0+I_0+R_0)]$
Class-to-class transmission rate	δ	to be estimated	-
Ratio of school cases to total cases	ρ	90.00%	[71, 193, 223]
Importing rate	τ	5.00%	assumed
Rate of losing infectiousness	η	30.00%	[158]

We initialize the ODE system (see Eqn. (5.1)) with the following values:

$$\{S_0, E_0, I_0, R_0\} = \{34.95\%, 0.00\%, 0.05\%, 65.00\%\}$$

According to SZCDC [6], it was common for children to be vaccinated or recovered from a varicella episode before entering schools, thus we assumed R_0 to be 65%. As in Lenne *et al.*, we assumed the rate of losing infectiousness to be $\eta = 30\%$ [158]. This rate represented the losses due to medication, contact isolation and/or hospitalization.

5.2.5 Steps for Parameter Estimation

The steps for parameter estimation are outlined as follows (See Algorithm 5.1). We evenly divide each parameter ($\Theta = [\theta_1, \cdots, \theta_p]^T$) into partitions, which is denoted as total K_i partitions for the ith parameter. At the initial step, we use the current parameter range ($\theta_i \in \left[\theta_i^{(L)}, \theta_i^{(U)}\right]$. For each of the total $\prod_{i=1}^{p} K_i$ combinations of parameters (θ_is), the simulation fitting is being run for N times, where N is a sufficiently large number. We compute the mean squared error for each run, i.e. $\widetilde{\text{mse}_j}$ for the jth combination of parameters.

After simulating all $\prod_{i=1}^{p} K_i$ combinations of parameters, we identify the jth parameters combination $\widehat{\Theta} = [\theta_{1q_1}, \cdots, \theta_{p_{q_p}}]^T$ that has the smallest $\widetilde{\text{mse}_j}$. (denoted as ($\text{MSE}_{\min} = \min\left\{\widetilde{\text{mse}_j} \middle| j \in \{1, \ldots, \prod_{i=1}^{p} K_i\}\right\}$)). Parameter ranges are then updated according to the current best-fitted parameter combination ($\widehat{\Theta}$) and the updated ranges are re-applied in the next step. When an updated range are within the acceptable error level (denoted as ε_i for the ith parameter), this parameter estimate is considered as a "acceptable". It is then output as a parameter estimate and is used to infer the rest of parameters.

Algorithm 5.1	
input:	parameter, $\Theta = [\theta_1, \cdots, \theta_p]^T$; runs for simulation, N; set of parameter index, $\mathbf{P} = \{1, \ldots, p\}$; ranges of parameters, $\theta_i \in [\theta_i^{(L)}, \theta_i^{(U)}]$; number of partitions for parameters, $K_i \geqslant 2$; error level, $0 < \varepsilon_i < (\theta_i^{(U)} - \theta_i^{(L)})$; estimator, $\widehat{\Theta} = [\widehat{\theta_1}, \cdots, \widehat{\theta_p}]^T$
Do {	
For $i \in \mathbf{P}$	
	$\{\Theta_i\} = \{\theta_{i1} = \theta_i^{(L)}, \theta_{i2}, \ldots, \theta_{iK_i} = \theta_i^{(U)}\}$, with K_i partitions evenly distributed in $[\theta_i^{(L)}, \theta_i^{(U)}]$ record $\{\Theta_i\}$
end For	
For each Θ	*(where, $\theta_i \in \{\Theta_i\}$ with $i \in \boldsymbol{P}$ and there are $\prod_{i=1}^{p} K_i$ combinations of θ_is)* given Θ_j, where $j \in \{1, \ldots, \prod_{i=1}^{p} K_i\}$, do N runs of simulation find the median of mean squared error, $\widetilde{\text{mse}_j}$, from N runs of simulation record $(\Theta_j, \widetilde{\text{mse}_j})$, with $j \in \{1, \ldots, \prod_{i=1}^{p} K_i\}$
end For	
Find $\text{MSE}_{\min}$	
	$\text{MSE}_{\min}$ = minimum of all $\widetilde{\text{mse}_j}$s, with $j \in \{1, \ldots, \prod_{i=1}^{p} K_i\}$ reset $\widehat{\Theta} = [\theta_{1q_1}, \cdots, \theta_{p_{q_p}}]^T$ for the combination of parameter that have achieved $\text{MSE}_{\min}$
end Find	
Set parameter range	
For $i \in \mathbf{P}$	
	update $\theta_i^{(L)} = \theta_{i(q_i-1)}$ and $\theta_i^{(U)} = \theta_{i(q_i+1)}$, where $q_i \in \{2, \ldots, (K_i - 1)\}$
If $\varepsilon_i \geqslant (\theta_i^{(U)} - \theta_i^{(L)})$	
	$\mathbf{P} = \mathbf{P}/\{i\}$
end If	
end For	
	record $\theta_i^{(L)}$ and $\theta_i^{(U)}$, with $i \in \mathbf{P}$
end Set	
} While $\left(\mathbf{P} \neq \varnothing\right)$	
output:	estimator: $\widehat{\Theta}$

5.2.6 Model Fitting

Fitting transmission rate

We proposed a continuous linear structure to our beta function:

$$\beta(t_{\text{week}}) = c_i^{'} + k_i \cdot t_{\text{week}}$$

where, $\beta(t_{\text{week}})$ was the transmission rate function, $c_i^{'}$ was the constant term, k_i was the slope and t_{week} was the week number of the current year. The subscript i represented the ith week segment of the school term, which was segregated by two longer school holidays. In China, summer breaks last for two months from July to August. Winter breaks are usually from mid-January to mid-February, and takes place around the lunar new year. For convenience, we converted the beta function into the following form:

$$\beta(t_{\text{week}}) = c_i + k_i(t_{\text{week}} - t_i) \qquad i \in \{1, 2, 3, ..., M\} \quad \& \quad t_{\text{week}} \in [t_i, t_{i+1}) \tag{5.3}$$

where t_i is the starting week number of the ith week segment in current year and there are total M week segments in the current year. Since the beta function was continuous within a year, our model only needed to fit the constant term (c_i) at the start of each week segment, i.e. node, such that, for the ith week segment, the estimated slope is given by:

$$\widehat{k_i} = \frac{\widehat{c}_{i+1} - \widehat{c}_i}{t_{i+1} - t_i} \tag{5.4}$$

where $\widehat{c_i}$ represents the fitted constant term for the ith week segment and the t_i is the starting week number of the ith week segment.

For each of the M nodes (c_i), we assumed they were ranged between 0.00 and 0.50. The Monte Carlo (MC) method was applied to estimate the best-fitted $\widehat{c_i}$ which had the smallest mean squared error.

School Terms

In our model, we divided each school year into M segments (Eqn. (5.3)) according to the school calendar in China: within a school year there were two semesters, each containing three segments: school vacation, beginning of semester, and end of semester, resulting in a total of six segments per year, i.e. $M = 6$.

The transmission rate (β) between a school term and vacation were different due to the differences in contact frequencies and patterns [102, 131, 138]. We separated each school term into two segments for two reasons: (i) decrease in contact between the susceptible and infected on the onset of an outbreak; and (ii) difference in seasonality due to climatic factors

such as temperature. Both reasons could lead to a change in the beta function.

Model Simulation

We ran the simulation 1000 times for each node (or c_i) combination. The Mean Squared Error (MSE) was the model fitting criteria between the weekly reported cases number and the model simulation median. A small number of cases in 2011 and 2012 were ignored, and the ABM-SEIR was only fitted to the reported cases from 2013 to 2015. The algorithm of parameter estimation were described in more details in S1 File.

The total number of infection cases, N, in Shenzhen, was given by:

$$N = \frac{\sum_j \sum_i N_{j,i}}{\rho} \tag{5.5}$$

where, $N_{j,i}$ was given by Eqn. (5.2). The c_is' combination with the smallest MSE was selected as the best-fitting model, and was adapted to the school-based vaccination scenario.

`R` software (version 3.3.1.) and `Java` (version 8) were used for modelling and computations.

5.2.7 Estimation of Basic Reproduction Number

Within our ABM, an infectious individual at model initialization could induce three levels of transmission. For the jth age group, the reproduction number for within class transmission is $\mathcal{R}_{\text{class}j}$:

$$\mathcal{R}_{\text{class}j} = \frac{(1-\eta)\cdot\theta_j\bar{\beta}S_0 n_j}{\gamma} \tag{5.6}$$

The reproduction number for class-to-class transmission, $\mathcal{R}_{\text{c-c}j}$, is:

$$\mathcal{R}_{\text{c-c}j} = \frac{\hat{\delta}\cdot(N_{\text{class}j}-1)\cdot\Pr(I_j \geqslant I_{\text{limit}})}{\gamma} \tag{5.7}$$

For out-of-school transmission, or imported cases, the reproduction number, $\mathcal{R}_{\text{import}_j}$,

is:

$$\begin{aligned}\mathcal{R}_{\text{import}_j} &= \tau(\mathcal{R}_{\text{class}j} + \mathcal{R}_{\text{c-c}j}) \\ &= \tau\frac{\theta_j\bar{\beta}S_0n_j + \hat{\delta}\cdot(N_{\text{class}j}-1)\cdot\Pr(I_j \geqslant I_{\text{limit}})}{\gamma}\end{aligned} \tag{5.8}$$

Based on Eqns. (5.6), 5.7 and 5.8, and by considering the effects from initial immunity and loss-of-infectiousness rate, the basic reproduction number of the jth age group, $\mathcal{R}_{0j}$, is derived as follows:

$$\begin{aligned}\mathcal{R}_{0j} &= \frac{\mathcal{R}_{\text{class}j} + \mathcal{R}_{\text{c-c}j} + \mathcal{R}_{\text{import}_j}}{(1-R_0)(1-\eta)} \\ &= (1+\tau)\cdot\frac{\theta_j\bar{\beta}S_0n_j + \hat{\delta}\cdot(N_{\text{class}j}-1)\cdot\Pr(I_j \geqslant I_{\text{limit}})}{\gamma\cdot(1-R_0)(1-\eta)} \qquad j\in\{1,2,...,J\}\end{aligned} \tag{5.9}$$

where, $\mathcal{R}_{0j}$ was the basic reproduction number for the jth age group, θ_j was the beta multiplier, and $\bar{\beta}$ was the average transmission rate over a one-year period. S_0 was the initial percentage of Susceptible, and we set $S_0 = 34.95\%$. $\hat{\delta}$ was the fitted class-to-class spread rate. n_i was the number of students per class of the ith age group, $N_{\text{class}j}$ was the number of classes per school for the jth age group (Table 5.1). I_j was the number of secondary infected cases within a class during the infectious period of the initial infected case, I_{limit} was the pre-defined outbreak threshold, $I_{\text{limit}} = 5$ in the ABM, which was also the trigger of class-to-class transmission. We set R_0, the initial percentage of Recovered, to 65.00%. η, the rate of losing transmissibility, to 30%; τ, the importing rate, to 5%. J was the total number of age groups, and we have $J = 4$ in ABM.

Within the probability term, $\Pr(I_j \geqslant I_{\text{limit}})$, we assumed that I_j follows a Binomial distribution, where $I_j \sim \text{Bino}(n = S_0n_j, p = \frac{\theta_j\bar{\beta}}{\gamma})$. Poisson distribution was not assumed because I_j should be a finite integer for any given class of ABM.

$\mathcal{R}_0$ is given by:

$$\mathcal{R}_0 = \sum_{j=1}^{J} p_j\mathcal{R}_{0j} \tag{5.10}$$

where, $\mathcal{R}_{0j}$ was the basic reproduction number for the jth age group. p_j was the proportion of the model population who belonged to the jth age group (Table 5.1).

5.2.8 Intervention Scenarios

In this chapter, we compared two scenarios:

- "No intervention" (baseline) scenario
- "School-based vaccination" scenario

"No intervention" scenario is the current status quo in Shenzhen. According to the Shenzhen Education Bureau, specific guidelines for handling varicella outbreaks at schools are not currently available. The "school-based vaccination" scenario is a hypothetical scenario where a school has reported varicella cases beyond the outbreak threshold level, in which case vaccination will be applied to all students within that school, except for the infected or recovered students.

We made simplifying assumptions by considering "single dose" vaccine only. We also ignored "breakthrough cases" where individuals could still get infected after vaccination [6, 70, 223], since they are negligible in numbers. Due to the short time period modelled, we ignored the effects of vaccine waning rate.

5.3 Results

5.3.1 Model Fitting Result

We fitted the reported cases on a weekly basis from 2013 to 2015 in Shenzhen, considering a summer wave and a winter wave each year. In Fig 5.6, the weekly reported cases were compared with the simulation median and their 95% CI from the ABM.

The beta function displays similar patterns in spring and fall semesters, and it appears to be lower during the school holidays than in school terms (see blue-dotted line in Fig 5.6). Thus, our simulated transmission pattern was biologically reasonable [102, 103, 131]. As described above, awareness of disease outbreaks and seasonality could explain the turning point during mid-semester [128].

The best-fitted average class-to-class transmission rate was, $\delta = 1.0$ per class·week.

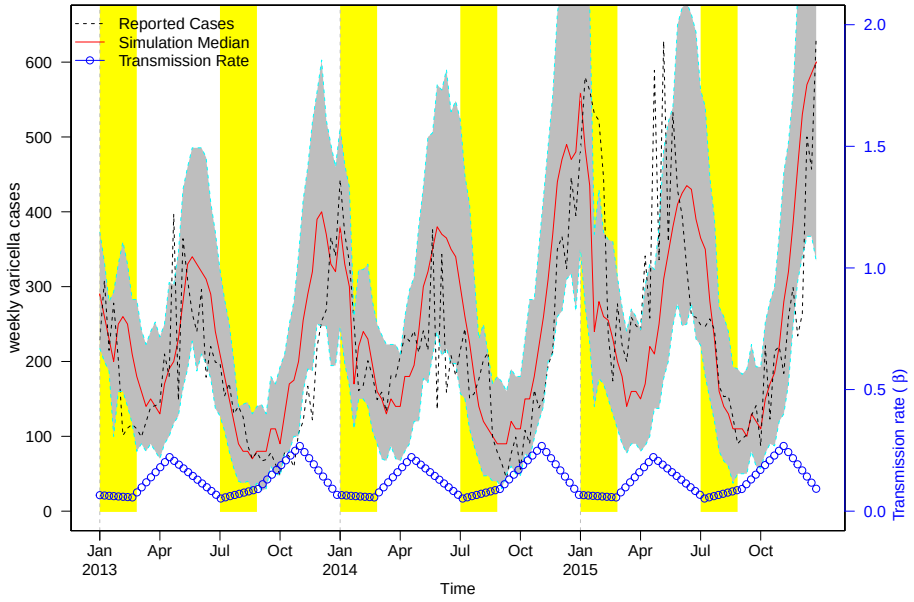


Figure 5.6: The ABM simulation results of varicella reported cases in Shenzhen from 2013 to 2015. The simulation median is plotted in red, reported cases are in black dashed line, the fitted transmission rate, $\beta(t)$, is the blue line at the bottom and the 95% Confidence Interval (C.I.) is in grey. School holidays are shaded in yellow.

5.3.2 Estimated Basic Reproduction Numbers

The basic reproduction numbers, $\mathcal{R}_{0j}$, for each age group are shown in Table 5.4. We could see that there were wide variations in $\mathcal{R}_{0j}$ among different age groups. This could be due to differences in class size ($N_{j,i}$) and school sizes (as in Table 5.1), or differences in beta multipliers, $\theta(a)$, among each age group (as in Table 5.2).

Table 5.4: Table of the basic reproduction numbers, $\mathcal{R}_{0j}$, for the jth age group in Shenzhen.

Age groups (years)	0 - 3	4 - 6	7 - 12	13 - 15
$\mathcal{R}_{0j}$	2.4597	6.4670	8.0892	6.6028

Based in Table 5.1, Table 5.4 and Eq 5.10, we estimated the overall basic reproduction number in Shenzhen as

$$\mathcal{R}_0 = 6.73.$$

Larger $\mathcal{R}_0$ was found among older age groups, which was consistent with earlier studies [243].

5.3.3 Impacts of Intervention and Varying Outbreak Thresholds on Transmission Dynamics

Fig 5.7 shows the simulation results for the intervention scenario, varying the outbreak threshold that triggered school-based vaccination from five to ten cases. We found that school-based vaccination intervention could effectively prevent large varicella outbreaks. Our results were shown in Table 5.5. By lowering the outbreak threshold, the school-based vaccination intervention could control the size of outbreaks more tightly. At an outbreak threshold of 5, varicella outbreaks could be reduced by 37% whereas a large school-level outbreak could be effectively controlled with a probability of 97%.(see Table 5.5 and panel (a) of Fig 5.7).

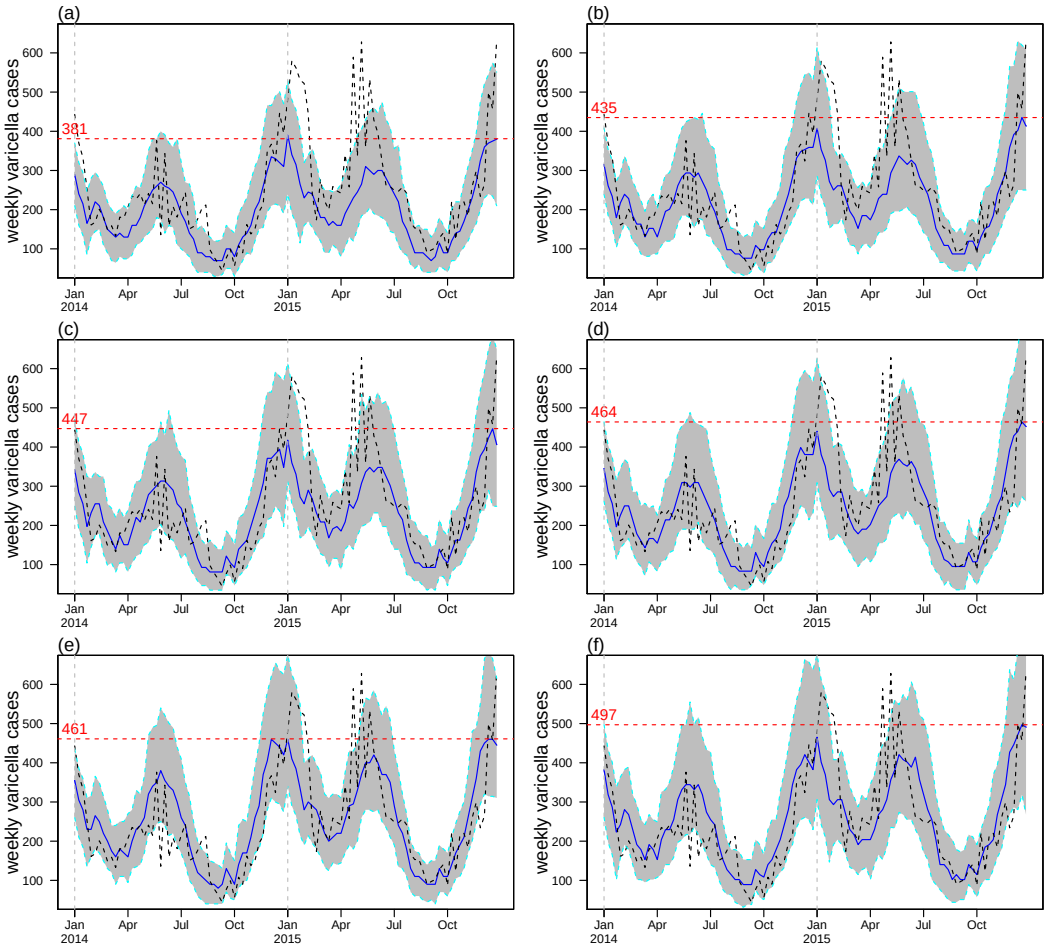


Figure 5.7: Simulation results with "vaccination" strategy from 2014-2015. The black dashed line is the confirmed cases which could be regarded as the baseline (i.e. no intervention) scenario. Simulation median is plotted in blue with 90% C.I. in grey. Panel (a), (b), (c), (d), (e) and (f) are simulation results with vaccination threshold set to be 5,6,7,8,9 and 10 (cases per week per school) respectively. The red dashed lines are the maximum weekly varicella cases during the simulation period (blue line), which represents the outbreak size under different outbreak thresholds.

Table 5.5: Summary table of the impact of intervention at various outbreak thresholds that triggered school-based vaccination. At each threshold level, we defined the "Maximum outbreak size" as the size of the largest outbreaks from 2014 to 2015, based on the simulation median. "Case reduction" was the percentage of varicella cases reduced due to the school-based vaccination strategy. "Reduction in Size of Outbreaks" was the percentage reduction in the size of the maximum outbreak compared with the baseline scenario. "Proportion of effective control" was the proportion of simulation runs that have simulated cases smaller than the reported cases, a proxy measure that the intervention could effectively bring the number of reported cases under control.

Outbreaks threshold	5	6	7	8	9	10
Maximum outbreaks size	381	435	447	464	461	497
Case reduction	27%	17%	12%	9%	3%	2%
Reduction in size of outbreaks	37%	28%	26%	23%	23%	17%
Proportion of effective control	0.97	0.93	0.89	0.80	0.88	0.75

5.4 Discussion

In this chapter, we developed a ABM-SEIR model to the reported varicella cases from 2013 to 2015 in Shenzhen. Our model adopted three transmission modes: within-class, class-to-class and out-of-school transmission. We also considered the age structure and an age-specific transmission rate. Our modelling structure is more biologically reasonable than previous studies [71, 138, 187, 223].

The key feature of our model was that the fitting of the transmission rate, $\beta(t)$, was strictly referred to as the segment of school terms in Shenzhen. The turning points of the beta function we identified when fitting transmission rate were compared to changes in school terms. Previous studies have applied more flexible time-dependent functions, such as cubic spline functions, to fit the beta function. However, cubic spline functions could possibly result in an over-fitting problem, and in some cases, the trends in transmission rates were not well-observed [49, 162]. We adopted a linear structure in our model fitting, which could offer apparent periodic dynamics in the transmission rate. Our fitted transmission rate function was the same each year, which demonstrates strong seasonality in varicella transmission (Fig 5.7). The changing dynamics of our fitted beta function were consistent with previous studies [102, 131, 171]. The annual pattern could be due to the school terms, seasonal factors (such as weather) and holiday schedules.

The estimated basic reproduction number, $\mathcal{R}_0 = 6.73$, was consistent with previous

works [138, 193, 243], suggesting that our model fitting was biologically reasonable. By varying the vaccination thresholds (Table 5.5 and Fig 5.7) and re-running the two scenarios, we show that lowering vaccination thresholds could incrementally lead to more effective varicella outbreak control. Our results were both logical and biologically reasonable. We further showed that it was not necessary to conduct school-based vaccination during non-epidemic periods.

Our results add to the varicella modelling literature in two ways. First, our use of ABM-SEIR model considered three levels of transmission that were more realistic than the WAIFW matrix used in previous studies [70, 71, 187, 223]. Second, our transmission rate function accounted for major school holidays and provided reasonable model fitness.

Our results of the impact of school-based vaccination (Table 5.5) were biological reasonable and logical, which provided important theoretical support of disease control decision-making among school population and development of school-based vaccination program. Our model was subject to some limitations. Household transmission, such as those between siblings, as well as reactive behavioral responses during a varicella outbreak, such as contact avoidance, taking medications or seeking clinical treatment, were not considered in this chapter. These factors could have altered the transmission rate function and the modelling parameters in our ABM-SEIR, and should be a focus in future studies.

5.5 Conclusions

There was a considerable increase in reported varicella cases from 2013 to 2015 in Shenzhen. Our ABM-SEIR was able to fit the two varicella confirmation waves from 2013 to 2015. The results showed that implementing a school-based vaccination intervention could effectively prevent large outbreaks at various vaccination thresholds. This chapter provides important theoretical support for disease control decision making during school outbreaks and the development of a school-based vaccination programme.

Chapter 6

An Epidemic Model with Theory of Game on Travelling

The conflict between what is best for the group and what populations do when individuals act in self-interest is a common research topic in socio-epidemiology. Most previous research studies how this conflict plays out in a closed population, but visitors can also play an important role in socio-epidemiology. Here, we study an "epidemiological travelling game". Individuals must decide whether or not to travel to an area affected by an infectious disease outbreak. To balance the benefits of travelling against potential disease risks, we regard the decision to travel as a game played against other visitors. The game theoretical framework is combined with an epidemic model to investigate the effects of travel strategies on local infection control. In contrast to many game theoretical analysis in a closed population, we find perfect agreement between individual and group optimal strategies for a broad range of epidemiologically and economically plausible values. However, extreme conflicts between individual and group optimal strategies emerge suddenly in other conditions, even under slight changes in epidemiological or economic conditions. Hence, game theoretical frameworks may be useful for anticipating when governments need to act during outbreaks to maximize population utility and when this may conflict with population behaviour.

6.1 Introduction

Visitors may play important roles in infectious disease transmission and spread. They can serve as susceptible hosts and be infected while staying in one place, and act as mobile sources of case imports to other populations [69, 86, 231]. On the one hand, more passengers or visitors could lead to substantial benefits of local economy and business. On the other hand, some infectious diseases spread aggressively in main tourism cities (e.g., Beijing, Hong Kong, Singapore, New York, Toronto), and a large number of passengers could bring unexpected impacts on public health [86, 97, 126], for example, Severe Acute Respiratory Syndromes (SARS) outbreaks in 2003 [62, 81, 124, 163, 190, 204], pandemic influenza [84, 120, 192], recent Ebola fever [107] and Middle East respiratory syndrome coronavirus (MERS-CoV) outbreaks. For example, SARS was introduce to population in Beijing, China by few infected visitors around early March, 2003 [190], which caused a large SARS endemic. Enforced travelling restriction could, in contrast, be an effective way to implement local disease control [52, 62, 127, 135].

Game theory attempts to analyse situations where individuals must make decisions in a group environment and where individual decisions influence the payoff received by the others in the group [183]. Many interventions such as vaccination and social distancing create positive externalities-benefits to those who did not take up the intervention because of herd immunity. Hence, many previous models have illustrated a conflict or discrepancy between the optimal individual strategy that maximizes personal interest, and the strategy that is best for the group as a whole [60, 61, 79, 83, 100, 196]. Although any number of factors may complicate this picture and have been explored in successive work, these models often illustrate this conflict occurring across a very broad region of parameter space, covering most epidemiologically and economically relevant regimes [60, 61, 79, 83]. However, previous works mostly concerned with individuals making decisions in the same population when the disease is spreading [57, 60, 61, 109, 196, 232], and did not consider multi-population interactions, or the strategic considerations faced by a visitor deciding whether to travel to an affected area during an outbreak.

In the context of travel decisions, game theory can be used to answer questions such as "whether or not to travel a place during an epidemic", "what is the optimal decision (to travel or not) based on individual interests", and "what should be the optimal control strategy for

government in response to minimize epidemic risks". In this work, we incorporate a game theoretical framework in an epidemic model (based on the classic SIR model) to investigate the effects of travelling strategies on local disease control. We found the positive relationship between incoming passengers rate (rate of visitors entering the border of the target place) and the possibility of passengers being infected during their staying.

Many previous game theoretical analyses of decision-making in epidemiological systems in a closed population find a significant discrepancy between individual and group optima, across a broad range of parameter values [57, 60, 61, 79, 83, 100, 109, 232, 233]. In contrast, for this visitor's game, we find perfect agreement between individual and group optimal strategies for a range of epidemiologically and economically plausible values. This agreement can be observed in two forms, namely: both individual and group optimal strategies completely reject travelling when real or perceived disease risk level are sufficiently high, or both strategies completely accept travelling when real or perceived disease risk level are sufficiently low (across epidemiologically plausible parameter ranges). Disagreement (or conflict) between individual visitor strategies and the group optimal strategy are found in two forms, namely: overload and deficit of visitors, compared to the group optimum. In regions where disagreement occurs, the disagreement between the individual optimum (corresponding to a "voluntary entrance" scheme) and the group optimum (corresponding to a "restricted entrance" scheme) is significant. During an outbreak, this conflict is likely to appear at any real or perceived disease risk level. More importantly, in this region, the model outcomes are highly sensitive to small changes in the infection transmissibility and visitor costs/benefits.

Uncontrolled visitors' inflow could bring unexpected large-scale outbreaks when disease risk level suddenly rises even a small amount, and local government's travel restrictions could effectively control the disease outbreaks when visitors' inflow is considered as "overload" during epidemics. Interestingly, the rate (or efficiency, in term of λ) of the disease risk information transmission to travelling population is find related to the conflict of interest between individual visitors and the whole visitors' population. The faster the disease risk information is updated, the more likely the conflict of interest could occur. Moreover, disease risk information being updated faster could effectively avoid visitors' inflow "overload", and thereby stop the outbreak.

The remaining parts of this work are organized as follows. In the next two sections, we establish the game theory framework to model the individual decision-making process

that incorporates an epidemic model including both travelling and local populations. In the subsequent section, the results are presented with detailed discussions. A summary on the major findings are provided at the end.

6.2 Travelling Game

Individuals in a travelling population (i.e., members of the population making a travelling decision and having the possibility to travel) can move through the following states:

$$\text{potential visitor} \rightarrow \text{visitor outside} \rightarrow \text{visitor inside} \rightarrow \text{removed visitor}. \tag{6.1}$$

The "potential visitor" corresponds to N_1 in Eqn. (6.8), "visitor outside" (i.e., visitor outside border, ρN_1 in Eqns. (6.10)), "visitor inside" (i.e., visitor inside border) correspond to $(S_1 + I_1 + R_1)$ from Eqns. (6.10), and "removed visitor" means that visitors leave or are removed from the system. Fig. 6.1 also presents the procedures of a "travelling" individual joining the epidemic system (i.e., from "potential visitor" to "removed visitor").

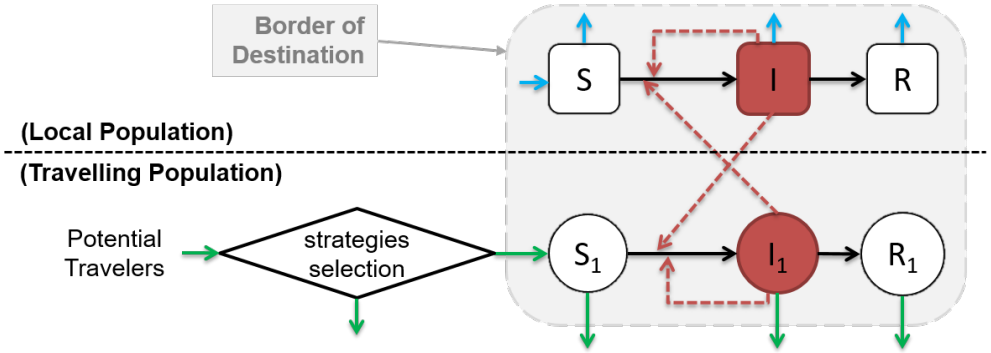


Figure 6.1: The epidemic model diagram. Black arrows represent infection status transition paths and red dashed arrows represent transmission paths. The light blue arrows represent the natural birth and death, and green arrows represent the visitors entry and leaving. Square compartments represent local classes, circular compartments represent visitors (travellers) classes, and the diamond denotes the "decision" procedure of potential visitors. Red compartments represent infective (or infectious) classes. The light grey area (rounded by grey dashed line) represents "inside border". The horizontal black dashed line separated the total population as "local population" (or local residence) and "travelling population" (as in Path (6.1)).

6.2.1 Description of the Game

In this travelling game, for simplicity, we suppose every individual receives the same information and makes decisions (or strategies) in the same way (i.e., equivalent payoff for the same strategy). An individual can make a decision whether to travel (i.e., "travelling" strategy) or not to travel (i.e., "non-travelling" strategy) the target place. We use r_1 to denote the perceived payoff of morbidity and/or mortality risk (i.e., the risk of disease, or as a term of "health cost") from infection and use r_0 to denote the perceived payoff of the risk of utility loss (as a measurement of "unsatisfactory") due to the failure of travel. Therefore, the payoff to an individual playing travelling strategy is presented as

$$E_1 = -\alpha \cdot \phi(\rho; P) \cdot r_1, \tag{6.2}$$

where α is the probability of an attack occurs ($\alpha = 1$, if the epidemic is ongoing), $\phi(\rho; P)$ is the probability that a visitor becoming infected after an attack if (given) the pre-existing immunity level in the population is P, and ρ is the overall proportion of visitors of all travelling-players (i.e., all game participants, the total number of individuals playing travelling or non-travelling strategies). Hence, for a disease with basic reproduction number: $\mathcal{R}_0 > 1$ (defined as: the expected number of secondary cases generated by a typical primary case during his/her infectious period in an otherwise susceptible population.), we have $\phi(\rho; P) = 0$ if parameter $P \geqslant \left(1 - \frac{1}{\mathcal{R}_0}\right)$ (see Fig. 6.2), because an outbreak can not persist when the disease-protecting level reaches $\left(1 - \frac{1}{\mathcal{R}_0}\right)$ (thus, no visitor will be infected) [48, 143]. We also have the payoff to an individual playing non-travelling strategy as

$$E_0 = -r_0, \tag{6.3}$$

and note that E_0 may, but not necessarily equal to E_1.

There exist a mixed strategy (namely, "p strategy") with a probability p to play travelling strategy (i.e., becoming a visitor), and a probability $(1 - p)$ to play non-travelling strategy (i.e., becoming a non-visitor, who will not travel after making this decision). The payoff function can be obtained by mixing the payoff functions of two pure strategies (see Eqns. (6.2)-(6.3)),

$$\begin{aligned} E(p,\rho;P) &= pE_1 + (1-p)E_0 \\ &= -p\alpha r_1 \cdot \phi(\rho;P) - (1-p)r_0, \end{aligned} \tag{6.4}$$

and the game will hold unchanged if we scale the payoff function by a constant, thus, we eliminate one parameter of Eqn. (6.4) by leaving only a relative risk, $r = \frac{r_0}{r_1}$ (normally, $0 < r_0 \ll r_1$ due to the payoff of utility loss, r_0 in Eqn. (6.3), should be less than that of health loss, r_1 in Eqn. (6.2), hence $0 < r \ll 1$), and we have

$$E(p,\rho;P) = p \cdot [r - \alpha\phi(\rho;P)] - r. \tag{6.5}$$

For convenience, we denote $\phi(\rho;P)$ as $\phi(\rho)$ and $E(p,\rho;P)$ as $E(p,\rho)$ because parameter P will be fixed for the rest of this work.

6.2.2 Individual Equilibrium

Suppose that a proportion ε $(0 < \varepsilon < 1)$ of travelling population would travel with a probability p (i.e., playing p strategy) and the rest of population $(1-\varepsilon)$ would travel with a probability $q \neq p$, then, the overall proportion of visitors $(\bar{\rho})$ of all travelling-players is

$$\bar{\rho} = \varepsilon p + (1-\varepsilon)q. \tag{6.6}$$

Therefore, the payoff to individuals playing p strategy (formulated in Eqn. (6.5)) and q strategy are $E(p,\bar{\rho})$ and $E(q,\bar{\rho})$ respectively. The payoff gain (or loss if negative) to an individual playing p strategy against q strategy is the difference of two payoff functions (see Eqn. (6.5)),

$$\Delta E = E(p,\bar{\rho}) - E(q,\bar{\rho}) = (p-q)\left[r - \alpha\phi(\bar{\rho})\right]. \tag{6.7}$$

Existence of Nash Equilibria The probability of a visitor becomes infected after an attack $(0 < \phi(\rho) < 1)$ must increase strictly (which is in line with [124], please also refer to Epidemic Model section for details) with a proportion (ρ) of travelling-players choosing

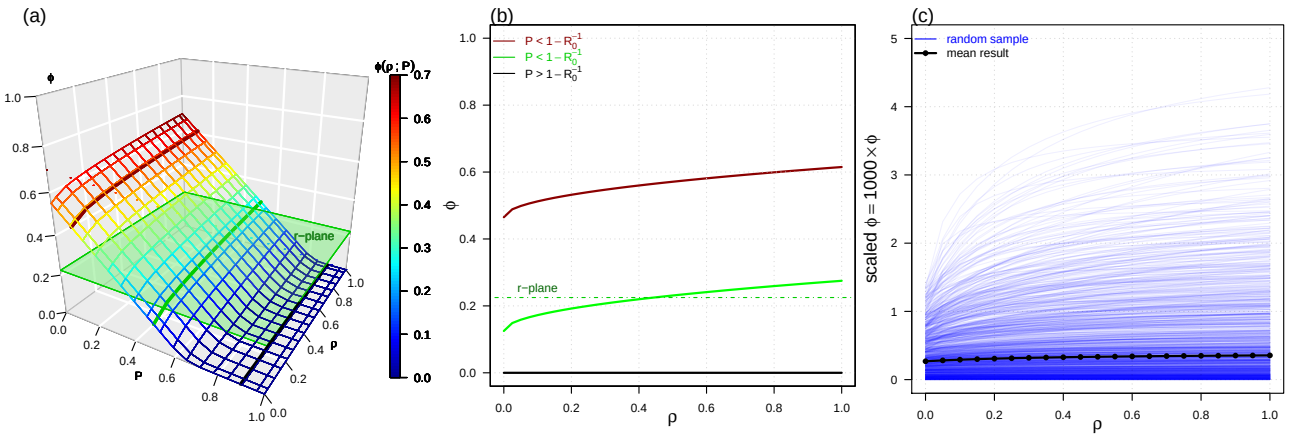


Figure 6.2: Schematic diagram of Nash equilibria under three situations (panel (a) and (b)) and numerical results of the relation between $\phi(\rho)$ and ρ (panel (c)). In order to have a clear illustration of three kinds of Nash equilibria, panel **(a)** and **(b)** show the rough (relations) trends among $\phi(\rho; P)$, ρ and P. Panel **(c)** shows the relation between scaled $\phi(\rho)$ and ρ under state $\mathcal{E}^{(3)}$ (see Eqn. (6.13)). The scaled $\phi(\rho) = 1000 \times \phi(\rho)$. In panel **(c)**, the transparently blue line are from 1,000 random samples with parameter sets, and the black dotted line is the result with fixed parameter values. The parameters' values and ranges can be found in Table 6.1.

travelling strategy (see Fig. 6.2). Hence, as P fixed, the minimum of $\phi(\rho)$ occurs at $\rho = 0$ and the maximum of $\phi(\rho)$ occurs at $\rho = 1$. Here, we show the existence of the unique Nash equilibria (by achieving $\Delta E > 0$ of Eqn. (6.7)) under three situations:

- If $\alpha \cdot \min\{\phi(\rho)\} = \alpha\phi(\rho = 0) \geqslant r$, $\alpha\phi(\rho) > r$ for all $0 < \rho < 1$, so for any $0 < \varepsilon < 1$ of Eqn. (6.6), $\Delta E > 0$ for any $q \neq p$ if and only if $p = 0$ (such that $p - q < 0$ for all $0 < q < 1$), thus, $p^* = 0$ is the unique Nash equilibrium.

- If $\alpha \cdot \max\{\phi(\rho)\} = \alpha\phi(\rho = 1) \leqslant r$, $\alpha\phi(\rho) < r$ for all $0 < \rho < 1$, so for any $0 < \varepsilon < 1$ of Eqn. (6.6), $\Delta E > 0$ for any $q \neq p$ if and only if $p = 1$ (such that $p - q > 0$ for all $0 < q < 1$), thus, $p^* = 1$ is the unique Nash equilibrium.

- If $\alpha \cdot \max\{\phi(\rho)\} = \alpha\phi(\rho = 1) > r > \alpha\phi(\rho = 0) = \alpha \cdot \min\{\phi(\rho)\}$, there exist one and only one p^* such that $\alpha\phi(\rho = p^*) = r$. For all $q < p$, we have $\bar{\rho} < p$ (according to Eqn. (6.6)) for any $0 < \varepsilon < 1$ and, similarly, for all $q > p$, we have $\bar{\rho} > p$ for any $0 < \varepsilon < 1$. Hence, for $\alpha\phi(\rho = 1) > r > \alpha\phi(\rho = 0)$, we always have $\Delta E > 0$ for all $q \neq p$ if and only if $p = p^*$, so p^* is the unique Nash equilibrium such that $\alpha\phi(p^*) = r$.

The various situations of the relationship between $\alpha\phi(\rho)$ and r is due to different values of the pre-existing immunity level in population (i.e., P, also see Fig. 6.2(a)-(b)) and different values of model parameters (see Fig. 6.2(c) and Table. 6.1).

Convergent Stability Follow previous work [60], let p be closer than q to p^* (the unique Nash equilibrium of Eqn. (6.7)), which means $q < p \leqslant p^*$ or $q > p \geqslant p^*$ (note that p may not necessarily equal to p^*). Given $\phi(\rho)$ increases with respect to ρ, if $q < p \leqslant p^*$, $(r - \alpha\phi(\bar{\rho})) > 0$ for all ε in Eqn. (6.6), thus we have $\Delta E > 0$, and, similarly, we can also have $\Delta E > 0$ if $q > p \geqslant p^*$ as desired. Therefore, the existed Nash equilibria for all three scenarios are convergently stable.

6.2.3 Travelling Optimum

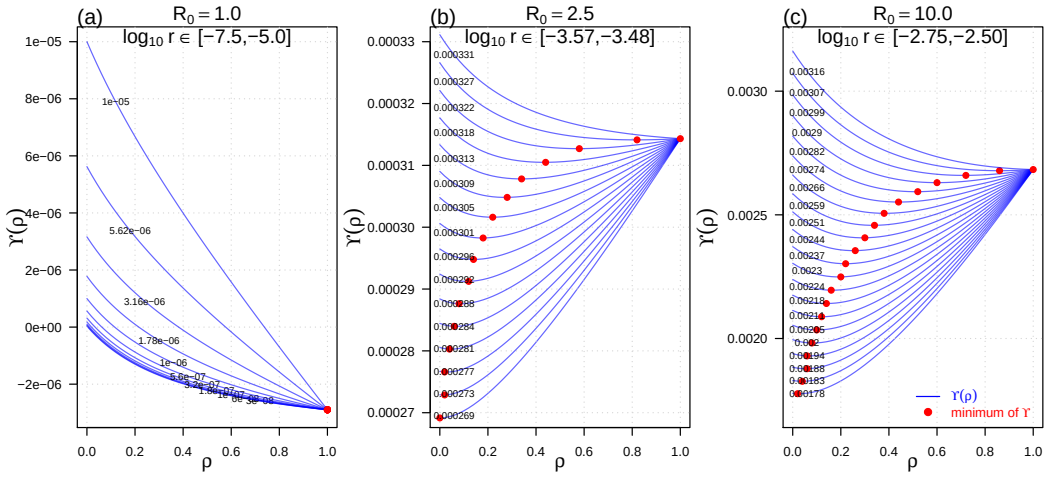


Figure 6.3: The optimal proportion of travelling-players becoming visitors (i.e., ρ^* corresponding to Eqn. (6.9)) during epidemic (i.e., $\alpha = 1$). Panel (a)-(c) corresponds to $\mathcal{R}_0 = 1.0, 2.5$ and 10.0 respectively. Blue lines are $\Upsilon(\rho)$ in Eqn. (6.9) with respect to different values of r and red dots are the minima (when $\rho = \rho^*$) of $\Upsilon(\rho)$, of which $\rho \in [0, 1]$. The values of r are shown on each blue line. The range of $\mathcal{R}_0$ and values of other parameters are in Table 6.1.

On the standpoint of the travelling population, we aim to minimize the overall risk level of all travelling-players as a whole, which also appears to be the goal of government control. We further assume that infected visitors do not pass the disease in the origin region, where the visitors come from (see more details in Discussion section). We can express the expected risk level due to the possible epidemic and the utility loss from non-travelling strategy in term of ρ (i.e., the overall proportion of visitors of all travelling-players),

$$\Upsilon(\rho) = N_1 \cdot [\rho\alpha \cdot \phi(\rho) \cdot r_1 + (1 - \rho)r_0] \,, \tag{6.8}$$

where N_1 is the ratio of total number of travelling-players to the total local population capacity (i.e., sum of maximal visitors capacity and the number of local population) and

other terms have same meaning as in Eqns. (6.2) and (6.3). We can further scale $\Upsilon(\rho)$ by eliminating N_1 (because N_1 can be fixed as a constant) and one risk term (replacing r_0 and r_1 by $r = \frac{r_0}{r_1}$, similar to the approach from Eqn. (6.4) to Eqn. (6.5)) [61], thus, the expected overall risk of visitors population is

$$\Upsilon(\rho) = \rho\alpha \cdot \phi(\rho) + (1 - \rho)r, \tag{6.9}$$

where all terms have same meaning as in Eqn. (6.5). The optimal travelling proportion (i.e., the ratio of successful entered visitors against the total travelling-players, if the government restriction on border entry is implemented), ρ^*, can be obtained by minimizing $\Upsilon(\rho)$ (see Fig 6.3 as numerical examples with the basic reproduction number $\mathcal{R}_0 = 1.0, 2.5, 10.0$) on the interval $\rho \in [0, 1]$, which should correspond to the minimal expected risk based on population level.

6.3 Epidemic Model

6.3.1 Formulation of Epidemic Model

To achieve a better understanding of infection probability $\phi(\rho)$, we develop an epidemiological model based on the standard susceptible-infectious-removed (SIR) model. Individuals are either susceptible (S, can be infected) to the disease, infectious (I, i.e., capable to transmit disease to susceptible individuals) and removed (R, either recovered or died, and cannot be infected). Additional susceptible (S_1), infectious (I_1) and removed (R_1) classes are included as the visitors (this patchy environmental mechanism was proposed previously in [69, 112, 229, 231]) who are considered to be totally susceptible when joining in the model system. Fig. 6.1 presents the modelling mechanism of this "local-and-travelling population" interactive epidemic system. The changing rates of the population in each compartment correspond to the transmission pattern of the disease (see Eqns. (6.10)). The susceptible visitors (S_1) are further assumed to be subject to the logistic growth.

$$\begin{cases} S' = \mu \cdot (1 - K_1 - S) - \beta S \cdot (I + I_1) \\ S_1' = f_\rho \cdot \left(1 - \dfrac{S_1 + I_1 + R_1}{K_1}\right) - \beta S_1 \cdot (I + I_1) - \nu S_1 \\ I' = \beta S \cdot (I + I_1) - (\gamma + \mu) I \\ I_1' = \beta S_1 \cdot (I + I_1) - (\gamma + \nu) I_1 \\ R' = \gamma I - \mu R \\ R_1' = \gamma I_1 - \nu R_1 \end{cases} \tag{6.10}$$

where $f_\rho = f(\rho) = \rho\lambda N_1$ represents the rate of travelling entry (i.e., input visitors to the target place), K_1 is the ratio of maximal capacity of visitors to the total population capacity (i.e., the maximal number of individuals can be contained in the target place) and N_1 has the same meaning as in Eqn. (6.8). K_1 controls the upper bound of the magnitude of visitors in the model system (thus, generally, K_1 could be fixed) and N_1 is responsible to the magnitude of potential visitors, for simplicity, we fix N_1 in this work. Model parameters are summarized in Table 6.1. Most of visitors are staying inside border (i.e., in the target place) for a considerably short period (3 days, see ν^{-1} in Table 6.1). In model (6.10), since $(S + I + R) + K_1 \equiv 1$ (i.e., the population threshold, or the total population capacity, is scaled to unity, 1) and $S_1 + I_1 + R_1 \leqslant K_1 < 1$, we have $(S + I + R) + (S_1 + I_1 + R_1) \leqslant 1$.

Basic reproduction number Using the next generation matrix method [221], the basic reproduction number (defined as: the average number of secondary cases generated by one case joining a completely susceptible population over its infectious period) of the epidemic model (Eqns. (6.10)) can be derived as

$$\mathcal{R}_0 = \beta \cdot \left[\frac{(1 - K_1)}{\gamma + \mu} + \frac{K_1}{\gamma + \nu}\right], \tag{6.11}$$

and thus, $\beta \propto \mathcal{R}_0$ when other parameters are fixed.

Elimination of Eqn. R_1' Under quasi-steady-state assumption (which is widely adopted in within-host modelling studies [64, 82]), we replace the term $\frac{S_1+I_1+R_1}{K_1}$ (in model (6.10)) by $\frac{S_1+\left(1+\frac{\gamma}{\nu}\right)I_1}{K_1}$ (by forcing $R_1' = 0$) in order to eliminate equation of R_1. This approximation can be interpreted as that all R_1 come from I_1 and only $\frac{\gamma}{\gamma+\nu}$ of I_1 could transit to R_1 at any time (other part of I_1 simply leaving the system at rate ν). Thus, $R_1 \leqslant \frac{\gamma}{\gamma+\nu} I_1 \leqslant \frac{\gamma}{\nu} I_1$

Table 6.1: Summary table of model parameters. The ranges of parameters are used for sensitivity analysis.

Parameter	Notation	Value	Range	Source(s)
Basic reproduction number	$\mathcal{R}_0$	$2.5^{\dagger}$	$[1.0, 10.0]$	[116, 159, 170, 234]
Mean period: visitors outside border	λ^{-1}	3 days	$[0.1, 10]$	[4]
Ratio: $\frac{\text{travelling-players}}{\text{population threshold}}$	N_1	7.5%	$[5.0\%, 15.0\%]$	assumed
Ratio: $\frac{\text{visitors capacity}}{\text{population threshold}}$	K_1	7.0%	$[5.0\%, 15.0\%]$	[1]
Mean infectious period	γ^{-1}	5 days	$[2.0, 10.0]$	[11]
Mean human lifespan	μ^{-1}	70 years	fixed	-
Mean period: visitors inside border	ν^{-1}	3 days	$[0.5, 15.0]$	[3]
Relative risk (as in Eqn. (6.5))	$r = \frac{r_0}{r_1}$	10^{-3}	$[10^{-4}, 10^{-2}]$	[2]
Probability to travel (as in Eqn. (6.4))	p	-	$[0.0, 1.0]$	defined
Proportion of visitors (as in Eqn. (6.2))	ρ	-	$[0.0, 1.0]$	defined
Probability: disease outbreaks occur	α	$0.01^{\ddagger}$	$[0.001, 0.02]$	assumed

The values of disease's parameters refer to influenza, and ranges of parameters refer to the majority of infectious diseases.
The values and ranges of parameters related to travel (i.e., K_1, r, ν^{-1} and λ^{-1}) refer to Hong Kong as the default destination.
† One can determine the function of β against $\mathcal{R}_0$ (i.e., $\beta(\mathcal{R}_0)$) explicitly from Eqn. (6.11), and $\mathcal{R}_0 = 2.5$ is also applied to 2003 SARS epidemic according to [62, 81, 124, 190, 200, 204]
‡ $\alpha = 1.0$ during epidemics.

(both γ and ν are positive), and then, $S_1 + I_1 + R_1 \leqslant S_1 + \left(1 + \frac{\gamma}{\nu}\right) I_1$. Since infected ($I_1$) visitors will quickly join R_1 class at the rate γ and the proportion of recovered visitors are relatively small, term $S_1 + I_1 + R_1$ is very close to $S_1 + \left(1 + \frac{\gamma}{\nu}\right) I_1$. Note that $\frac{\gamma}{\nu} I_1$ is simply the upper bound of R_1, and, after all, the effects of both I_1 and R_1 are small (compared with S_1) regarding to the visitors input.

Since equations of R and R_1 can be eliminated, we reformulate the epidemic model as,

$$\begin{cases} S' = \mu \cdot (1 - K_1 - S) - \beta S \cdot (I + I_1) \\ I' = \beta S \cdot (I + I_1) - (\gamma + \mu) I \\ S_1' = f_\rho \cdot \left[1 - \dfrac{S_1 + \left(1 + \frac{\gamma}{\nu}\right) I_1}{K_1}\right] - \beta S_1 \cdot (I + I_1) - \nu S_1 \\ I_1' = \beta S_1 \cdot (I + I_1) - (\gamma + \nu) I_1 \end{cases} \tag{6.12}$$

where β is a function of $\mathcal{R}_0$ due to Eqn. (6.11).

6.3.2 Model Equilibria

The disease-free equilibrium (DFE, with $I = I_1 = 0$) is $\mathcal{E}^{(1)} = \left(S^{(1)}, I^{(1)}, S_1^{(1)}, I_1^{(1)}\right) = \left((1-K_1), 0, \frac{f_\rho K_1}{f_\rho + \nu K_1}, 0\right)$, and in particular, $S_1^{(1)} = \frac{f_\rho K_1}{f_\rho + \nu K_1} < K_1$. The DFE ($\mathcal{E}^{(1)}$) is globally stable when $\mathcal{R}_0 < 1$, but unstable when $\mathcal{R}_0 > 1$.

When $\mathcal{R}_0 > 1$, there is an endemic, pure non-visitor equilibrium (with $S_1 = I_1 = 0$), $\mathcal{E}^{(2)} = \left(S^{(2)}, I^{(2)}, S_1^{(2)}, I_1^{(2)}\right) = \left(\frac{\gamma+\mu}{\beta}, \mu \cdot \left(\frac{1-K_1}{\gamma+\mu} - \frac{1}{\beta}\right), 0, 0\right)$. Specifically, $S^{(1)} = \frac{\gamma+\mu}{\beta}$ is the reciprocal of $\mathcal{R}_0$ of the standard SIR model [143]. $\mathcal{E}^{(2)}$ can be realized when f_ρ in S_1' (see Eqns. (6.12)) becomes 0 and it is locally stable.

When $\mathcal{R}_0 > 1$, there also exist an endemic equilibrium corresponding to a mixed state of local and visitor's infections (i.e., infected visitors), denoted as $\mathcal{E}^{(3)} = \left(S^{(3)}, I^{(3)}, S_1^{(3)}, I_1^{(3)}\right)$. The solution of $\mathcal{E}^{(3)}$ can be obtained explicitly by taking the non-negative root of $[S', I', S_1', I_1']^T = \mathbf{0}$ ($\mathbf{0}$ represents the zero vector) with both $I, I_1 \neq 0$.

There also exist two special kinds of equilibria: one with only susceptible local population (i.e., $S \neq 0$ and $I = S_1 = I_1 = 0$), and another one with only susceptible travelling population (i.e., $S_1 \neq 0$ and $S = I = I_1 = 0$). Since this work aims to study the pattern of interaction between local and travelling population during epidemics, we ignore these two types of equilibria.

6.3.3 Infected Probability of Visitors

During an epidemic (i.e., $\alpha = 1$), for $\mathcal{E}^{(3)}$, we have $S_1 = S_1^{(3)}$ and $I = I^{(3)} \neq 0$ to be constant. Thus the probability that a visitor becomes infected after an attack if (given) the immunity level in the population equals P is, according to Eqns. (6.12), the proportion of S_1 becoming infected versus leaving the place at any time as susceptibles [60],

$$\begin{aligned} \phi(\rho) &= \frac{\beta S_1^{(3)}(I^{(3)} + I_1^{(3)})}{\beta S_1^{(3)}(I^{(3)} + I_1^{(3)}) + \nu S_1^{(3)}} = 1 - \frac{\nu}{\beta(I^{(3)} + I_1^{(3)}) + \nu}, \\ \text{thus,} \quad \alpha\phi(\rho) &= \alpha - \frac{\nu\alpha}{\beta(I^{(3)} + I_1^{(3)}) + \nu}, \end{aligned} \tag{6.13}$$

we also show numerical results of the relation between $\phi(\rho)$ and ρ in Fig. 6.2(c). Provided the relation between β (the effective contact rate) and $\mathcal{R}_0$ (the basic reproduction number,

see Eqn. (6.11)), one can also derive the relation between $\mathcal{R}_0$ and $\phi(\rho)$ explicitly. Note that during an epidemic (outbreak, i.e., $\alpha = 1$), $\alpha\phi(\rho) = \phi(\rho)$, and in this case, $\phi(\rho)$ can directly reflect the probability of infection after travel.

6.4 Results and Discussion

6.4.1 Results of Individual Equilibrium and Travelling Optimum

During an epidemic (i.e., $\alpha = 1$), Fig. 6.4(a) shows the relationships between relative risk (r, see Eqn. (6.5)), basic reproduction number ($\mathcal{R}_0$) and individual equilibrium (p^*, see Eqn. (6.7)), and Fig. 6.4(b) shows the relationships between r (Eqn. (6.5)), $\mathcal{R}_0$ and travelling optimum (or population optimum, ρ^*, Eqn. (6.9)). Other parameters are fixed according to Table 6.1, and the variations of other parameters' values do not change the trend of these relationships.

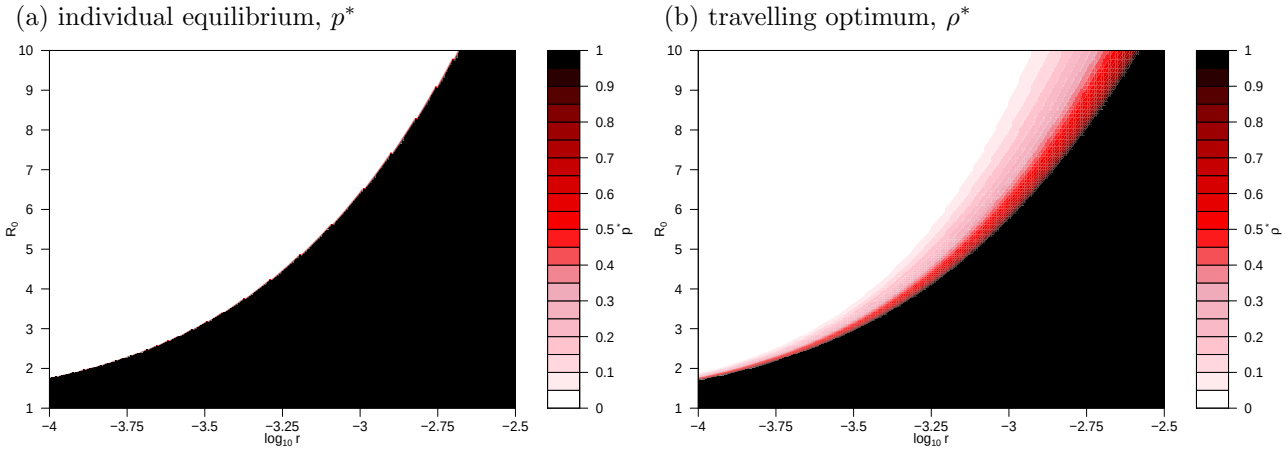


Figure 6.4: Individual and population travelling optima, as functions of the basic reproduction number and the relative risk. Panel **(a)** shows the relationships between r (Eqn. (6.5)), $\mathcal{R}_0$ and p^* (Eqn. (6.7)); panel **(b)** shows the relationships between r, $\mathcal{R}_0$ and ρ^* (Eqn. (6.9)) during epidemic (i.e., $\alpha = 1$). The value of r is in "$\log_{10}$" form. The color code of the optimal individuals travelling probability, p^*, and the optimal travelling proportion, ρ^*, are shown on the color key of each panel. The range of $\mathcal{R}_0$ and values of other parameters are in Table 6.1.

Both individual and population optima have qualitatively the same relationships with $\mathcal{R}_0$ and r, where optima are a monotonically decreasing function of $\mathcal{R}_0$, but a monotonically increasing function r. Besides, the transition from 0 to 1 of individual optimum is sharper (abrupt) in panel (a) than the transition of population optimum in panel (b).

For the individual equilibrium (see Fig. 6.4(a)), if the relative risk (r) is fixed, decreasing the basic reproduction number ($\mathcal{R}_0$) increases the individual travelling probability (p^*), which could be due to some seasonal factors that decline $\mathcal{R}_0$. However, the decline of visitors' perceived knowledge about disease risk (leads to the decline of r_1), which could be due to lack of media coverage on outbreak and relevant education programs [58, 59, 85, 87, 164, 216], will rise the relative risk (r), thus, rise the optimal individual travelling probability (p^*).

For the travelling optimum (see Fig. 6.4(b)), if the relative risk (r) is fixed, increasing the basic reproduction number ($\mathcal{R}_0$) decreases the optimal travelling proportion (ρ^*), which could be due to the unnoticed evolution of disease factors such that rising $\mathcal{R}_0$. In contrast, the decline of disease risk (leads to the decline of r_1, because disease risk should be positively related to r_1), which could be due to increasing of effective vaccination uptake of potential visitors (i.e., travelling-players), will rise the relative risk (r), thus, rise the optimal travelling proportion (ρ^*). Fig. 6.3 also shows the relation of ρ^* and r with the basic reproduction number fixed to be 1.0, 2.5 and 10 respectively.

6.4.2 Conflict between Individual Equilibrium and Travelling Optimum

In this subsection, local travelling optimum and individual equilibrium are compared by taking the difference, denoted as $\Delta\rho$, between the optimal local travelling proportion (ρ^*, based on population level, see Eqn. (6.9)) and probability (p^*, based on individual interest, see Eqn. (6.7)),

$$\Delta\rho = \rho^* - p^*. \tag{6.14}$$

During an epidemic (i.e., $\alpha = 1$), Fig. 6.5 shows the relationships between relative risk (r, see Eqn. (6.5)), basic reproduction number ($\mathcal{R}_0$) and $\Delta\rho$ (Eqn. (6.14)). When $\Delta\rho = 0$, both population optimum (ρ^*) and individual equilibrium (p^*) have reached a perfect agreement such that $\rho^* = p^* = 0$ or 1 (see white area in Fig. 6.5). These two situations could be attributed to both disease risk (reflected by $\mathcal{R}_0$) and perceived disease risk to be considerably high (corresponding to $\rho^* = p^* = 0$, in which no one intends to travel and complete restriction of border-entrance should be implemented) or considerably low

(corresponding to $\rho^* = p^* = 1$, in which everyone would like to travel and there is totally unrestricted border-entrance). The change in other parameters' values will not change the trend of these relationships.

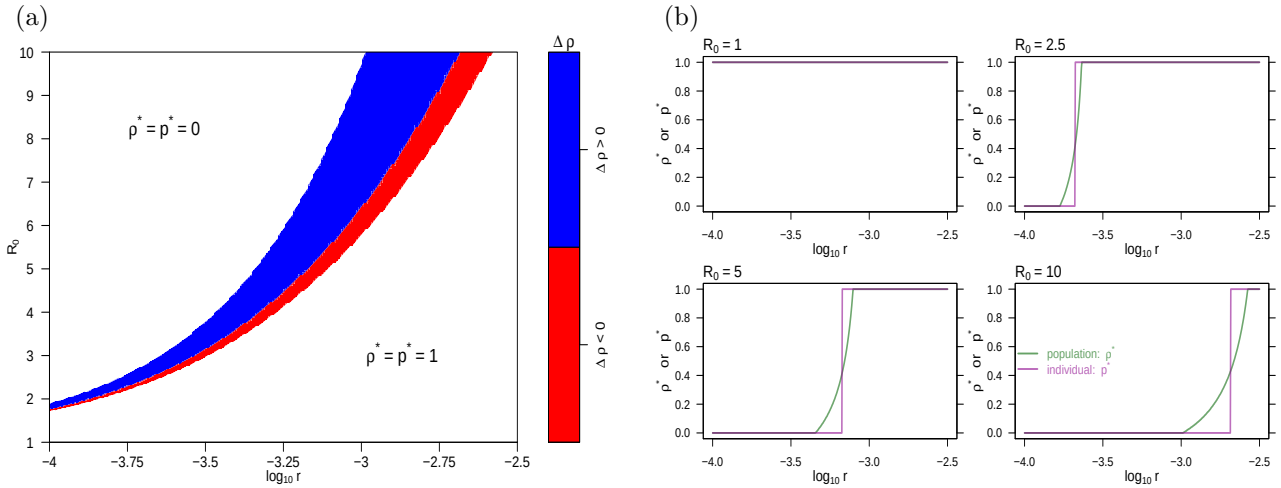


Figure 6.5: Conflict of interest between individual and population travel optimum as a function of the basic reproduction number and the relative risk. Panel **(a)** shows the relationships between r (Eqn. (6.5)), $\mathcal{R}_0$ and $\Delta\rho$ (Eqn. (6.14)); and panel **(b)** shows the relationships between r and $\Delta\rho$ for $\mathcal{R}_0 = 1.0, 2.5, 5.0, 10.0$ during epidemic (i.e., $\alpha = 1$). The value of r is in "$\log_{10}$" form. In panel **(a)**, the color code of the difference of individual and population strategy, $\Delta\rho$, is shown on the color key. The white area represents $\Delta\rho = 0$ under two situations that $\rho^* = p^* = 0$ or 1. In panel **(b)**, ρ^* is in green, and p^* is in purple. For both of panels, the range of $\mathcal{R}_0$ and values of other parameters are in Table 6.1.

With large ongoing epidemics, most of the places are expected to be at a deficit of visitors state in which limited visitor entrance (i.e., only a small proportion of potential visitors are intend to travel) would be unrestricted (i.e., $\Delta\rho > 0$ and $0 < \rho^* < 1$, see blue area in Fig. 6.5). If either disease risk (reflected by $\mathcal{R}_0$) or perceived payoff to disease risk (r_1) **slightly** declines (could be due to seasonal factors and/or lack of relevant media coverage), it could make the conflict of interest ($\Delta\rho$) change from $\Delta\rho > 0$ to $\Delta\rho < 0$ (i.e., $\rho^* < p^*$). In this case, local governments are suggested to enhance travelling entrance restriction (with only $\frac{\rho^*}{p^*}$ proportion of visitors should be allowed to enter the destination) to achieve minimal population payoff, see Eqn. (6.9). Otherwise, this could result in an overload of visitors state (i.e., $\Delta\rho < 0$, see red area in Fig. 6.5).

6.4.3 Numerical Examples

Fig. 6.6(a) shows a numerical example of epidemic worsening as the basic reproduction number ($\mathcal{R}_0$) declines from 2.5 to 2.4 while the enhancement of local travelling entrance

restriction fails (see the red line), and the outbreak can be effectively controlled if the previous entrance restriction maintains (i.e., hold $\rho = 0.1$ unchange, see the green line). The occurrence of this conflict between local travelling optimum (ρ^*) and individual equilibrium (which is also the individuals' optimum, p^*) is due to p^* being more sensitive than ρ^* with respect to $\mathcal{R}_0$ and r.

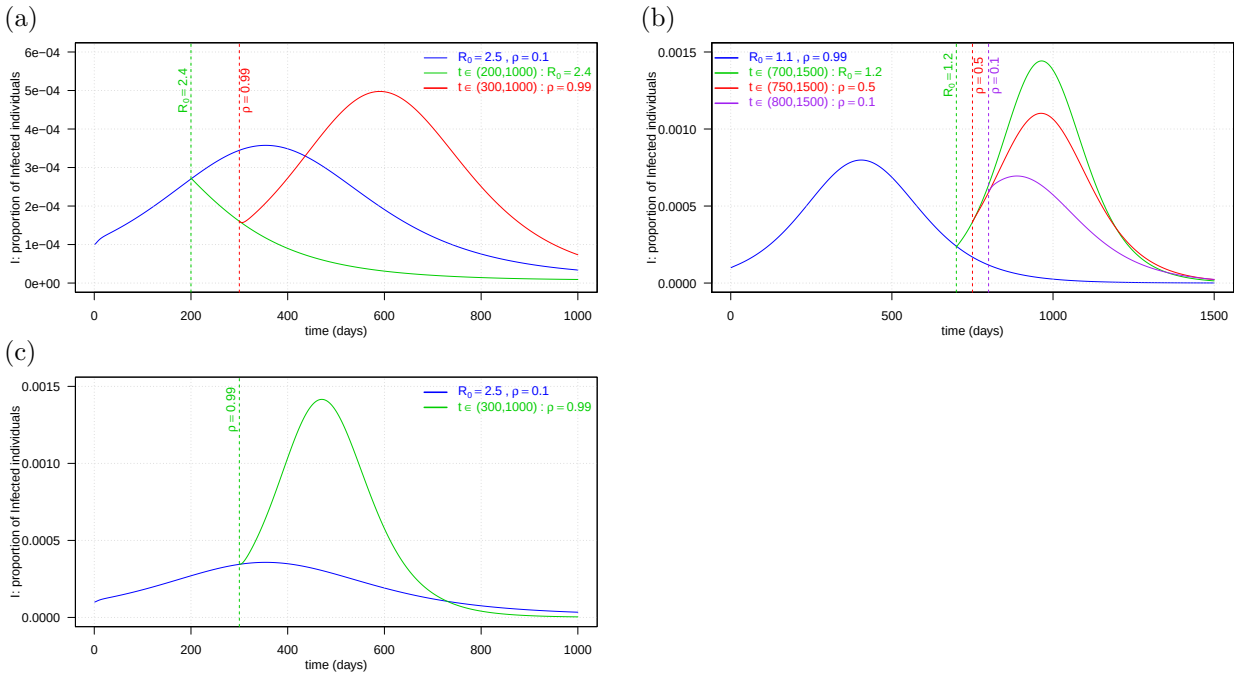


Figure 6.6: The simulation results of local infections (I) of epidemic model (panel (a) and (b), see Eqns. (6.12)) and the SARS epidemic in China in 2002-03 (panel (c)). The baseline scenario contains that initial states are set as $[S(0), I(0), S_1(0), I_1(0)] = \left[\frac{1}{\mathcal{R}_0}, 1 \times 10^{-4}, \left(K_1 - 5 \times 10^{-6}\right), 5 \times 10^{-6}\right]$; with $\mathcal{R}_0 = 2.5$ and $\rho = 0.1$ for panel **(a)** and **(c)**, and $\mathcal{R}_0 = 1.1$ and $\rho = 0.99$ for panel **(b)**. Values of other parameters are in Table 6.1. In panel **(a)**, the blue line is the simulation results under baseline scenario of panel **(a)**; the green line is of basic reproduction number ($\mathcal{R}_0$) decreasing to 2.4 since the 201-st day (vertical green dashed line); based on the change of green line, the red line is of travelling proportion (ρ) increasing to 0.99 since the 301-st day (vertical red dashed line). In panel **(b)**, the blue line is the simulation results under baseline scenario of panel **(b)**; the green line is of basic reproduction number ($\mathcal{R}_0$) increasing to 1.2 since the 701-st day (vertical green dashed line); based on the change of green line, the red line is of travelling proportion (ρ) decreasing to 0.50 since the 751-st day (vertical red dashed line); based on the change of red line, the purple line is of travelling proportion (ρ) continually decreasing to 0.10 since the 801-st day (vertical purple dashed line). In panel **(c)**, the blue line is the simulation results under baseline scenario of panel **(c)**; the green line is of travelling proportion (ρ) increasing to 0.99 since the 301-st day (vertical green dashed line).

It is plausible that the optimal individual traveling probability, ρ^* (equivalent to ρ in Fig. 6.4(a)), could change from 0.1 to 0.99 provided $\mathcal{R}_0$ varies from 2.5 to 2.4. This is due to ρ^* is considerably sensitive to both $\mathcal{R}_0$ and r (see Fig. 6.4(a)), and notice that ρ^* changes extremely fast from 0 to 1 within a small range of either $\mathcal{R}_0$ or r. However, practically, only the decline of disease risk ($\mathcal{R}_0$) might not cause this conflict immediately (or simultaneously) because potential visitors (or travelling-players, N_1) could be informed with delay (due to

untimely news or notices, see the gap between the vertical green and red dashed lines in Fig. 6.6(a)), thus, holding $\mathcal{R}_0$ unchanged (also r fixed as default) would not yield decrease of p^*.

When a place is currently at a state with relatively low epidemic level (i.e., the white area with $\rho^* = p^* = 1$, in which disease risk, $\mathcal{R}_0$, is relatively lower than the perceived disease risk), local government would consider to adopt border-entrance control measures in response to the slight rise of epidemic risk (reflected by $\mathcal{R}_0$). Fig. 6.6(b) shows a numerical example of epidemic worsening as the basic reproduction number ($\mathcal{R}_0$) slightly rises from 1.1 to 1.2 without any local travelling entrance restriction (see the green line), and the outbreaks situation can be controlled by even implementing a "non-strict" entrance restriction (see the red and purple lines). Increasing of either $\mathcal{R}_0$ or r_1 that leads to $\Delta\rho > 0$, implies individual equilibrium (p^*) decreases much faster (see Fig. 6.4) than population (i.e., group) optimum (ρ^*).

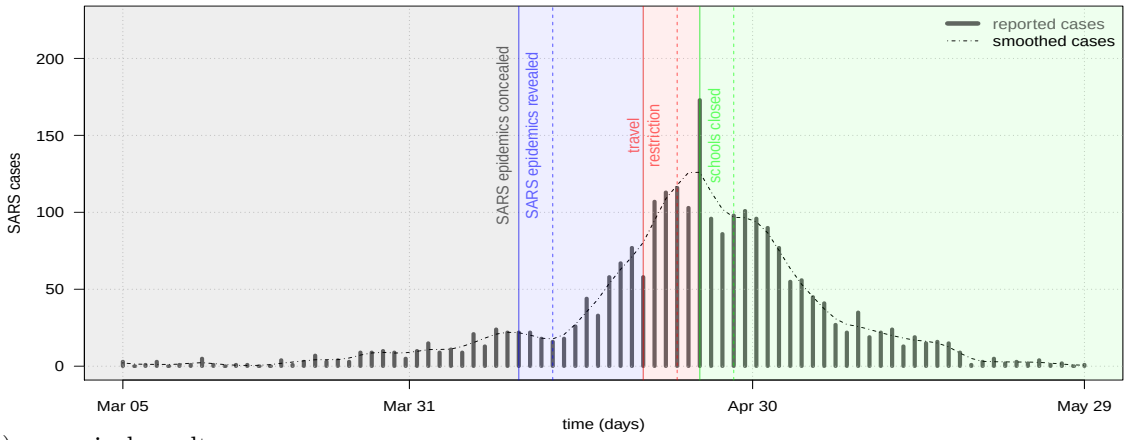
Fig. 6.6(c) shows the similar trend as the early stage of SARS epidemic (in Jan - Feb, 2003), which could be mainly due to the increased travelling flow during Chinese new year (see Fig. 2(a) of Ref. [99]). This indicates the increase of visitors could lead to disease outbreak.

6.4.4 Example of the 2003 SARS Outbreaks in Beijing

Fig. 6.7(a) shows the reported cases during the 2003 SARS outbreak in Beijing, China (this figure is revised from Ref. [190]). The time point where knowledge of the epidemic was first made public (see event "*SARS made reportable (Apr 10)*" in Fig. 1 of Ref. [190]) is referring to news press [33]. The time point of the official start of restrictions on travel is referring to the events "*outbreak announced publicly by government (Apr 20)*" and "*fever check at airport begin (Apr 22)*" in Fig. 1 of Ref. [190] (these two events resulted in almost no one travelling to Beijing, $\rho = 0$, until the end of SARS [32]).

Fig. 6.7(b) shows an epidemic curve from the model that qualitatively matches the SARS epidemic curve in Beijing. Ignoring the latent period, we adjust the values of two parameters (i.e., $\mathcal{R}_0$ and ρ, where $\mathcal{R}_0$ quantifies disease transmissibility and ρ quantifies the travelling proportion) qualitatively according to the recorded government policies. The

(a) The 2003 SARS in Beijing, China



(b) numerical results

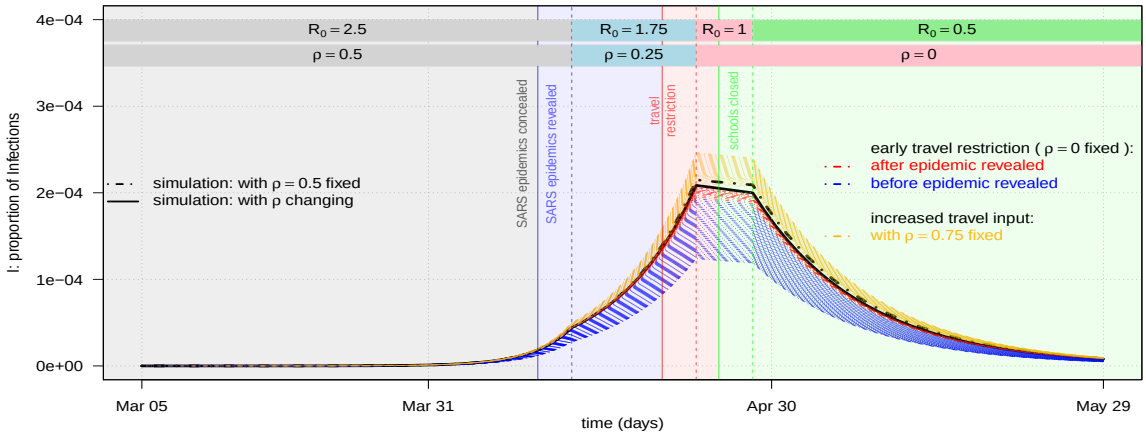


Figure 6.7: The 2003 SARS outbreak in Beijing, China. Panel **(a)** shows the reported cases during 2003 SARS outbreak in Beijing, China (this panel is revised from Ref. [190]); and panel **(b)** shows the numerical results of the epidemic model (see Eqns. (6.12)). In both panels, the vertical lines represent the starting points of events, and the vertical dashed lines represent the time-points with lag to be 3 days. In panel **(a)**, the epidemic of SARS and intervention of government are given on timeline from Mar 05 to May 29, 2003. The back dashed line is the smoothed time series by using *loess* function. In panel **(b)**, the initial states are set as $[S(0), I(0), S_1(0), I_1(0)] = [(1-K_1), 0, (K_1 - 1\times10^{-8}), 1\times10^{-8}]$, with $\mathcal{R}_0 = 2.5$, $N_1 = 15\%$ and $\rho = 0.5$ (see grey parts of the bars on the top). The blue and red dashed lines are the simulations under "what if" scenarios that travel restriction policies were implemented earlier. The black and gold dashed lines are under "what if" scenario that travel restriction (or reduction) was failed and travel input suddenly increased respectively. The values of other parameters are assumed to be same as in Table 6.1, and the changes of parameters are marked on the top of panel. Note that the timelines are the same and consistent in panel (a) and (b).

decrease of ρ from 0.5 to 0.25 (i.e., decrease of travel input, see the blue dashed vertical line in Fig. 6.7(b)) could be mainly due to public awareness of the SARS risk in Beijing after it was revealed [33]. Similarly, the decrease of $\mathcal{R}_0$ from 2.5 to 1.75 (i.e., decrease of infectivity, see the blue dashed vertical line in Fig. 6.7(b)) could also be mainly due to the reduction of effective contacts (i.e., the product of the contact rate and transmission probability per contact, which is believed to be (thus modelled) non-positively related to

the disease incidence [85, 87, 112, 164, 242]) among the population as a consequence of the increase of public awareness of SARS risk after it was revealed [33]. The trends of numerical results hold under changing magnitudes of $\mathcal{R}_0$ and ρ. The time lag (see the gaps of vertical lines and vertical dashed lines of the same color in Fig. 6.7) is fixed to be 3 days under due to the mixed effects of incubation period (or exposed period) of SARS infections and delay of human reaction to the events. Similar patterns can be observed between the numerical results and SARS cases time series in Mar - May, 2003 (see Fig. 6.7(a)-(b) and Fig. 8 of [230]).

The results of earlier implementation of travel restriction (see blue and red dashed lines in Fig. 6.7(b)) are obtained under "what if" scenario by fixing the structures of $\mathcal{R}_0$ and N_1, and setting $\rho = 0$ (i.e., nobody is able or willing to cross the border due to travel restriction or caution of SARS risk). We found that the earlier the implementation of travel restrictions is, the more effectively the disease outbreak level is reduced. By contrast, uncontrolled sudden increase of the proportion of visitors (by increasing ρ from 0.5 to 0.75) could lead to larger outbreaks (see gold dashed lines in Fig. 6.7(b)).

6.4.5 Sensitivity Analysis of Payoffs

Partial rank correlation coefficient (PRCC) are used to assess the dependence of the model results on the parameters [82, 111, 239]. The range of model parameters used for sensitivity analysis are summarized in Table 6.1.

Fig. 6.8 shows the PRCCs between model parameters and individual payoff (E, see Eqn. (6.5)), and population risk level (Υ, see Eqn. (6.9)) respectively. Since "payoff" (the term in Fig. 6.8(a)) is the defined as the opposite number of "risk level" (the term in Fig. 6.8(b)), some model parameters have symmetric PRCC results with respect to level "0" (see the vertical grey dashed line in Fig. 6.8) on both panels. The PRCCs show that the results are most sensitive to the group of the relative risk (r), the basic reproduction number ($\mathcal{R}_0$), and the rate at which individuals leave the destination country (ν). Hence, these parameters should be the focus of data collection efforts during outbreaks when a travel policy must be decided. In Fig. 6.8(b), the basic reproduction number ($\mathcal{R}_0$) and relative risk (r) is strongly positively related to the population risk level (Υ), and the visitors leaving rate (ν) is negatively related to Υ. Opposite results can be seen in Fig. 6.8(a) for individual

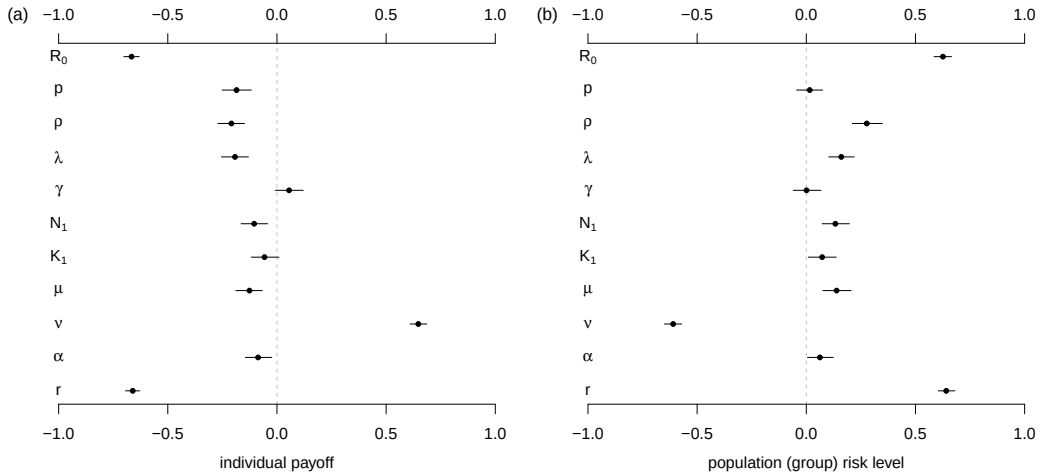


Figure 6.8: Sensitivity analysis results of (PRCCs) between model parameters and individual payoff (panel (a), see Eqn. (6.5)), and population risk level (panel (b), see Eqn. (6.9)). The black dots are the estimated correlations and the bars represent 95% C.I.s.

payoff.

6.4.6 Results and Discussion on Model Parameters

Relative risk $r = \frac{r_0}{r_1}$ (see Eqns. (6.2)-(6.5) and Table 6.1) is the ratio of the "non-travelling" payoff ($E_0 = -r_0$, see Eqn. (6.3)) to the maximum (or upper bound) of the "travelling" payoff (i.e., $E_1 = -r_1$ on the condition that a visitor is known to be infected, see Eqn. (6.2)). Its (r) range could be obtain via referring to the claim-settlement-odds of the travel insurance with regard to the travelling destination (normally, $r \approx 10^{-3}$, e.g., see note [2]).

Number of visitors N_1 is the ratio of total number of potential visitors (i.e., travelling-players, see Path (6.1)) to the total population capacity (i.e., sum of maximal visitors capacity and the number of local population, see Eqn. (6.8) and corresponding to ($S+I+R+K_1$) in model (6.10)). Provided total population capacity can be fixed in short term, the magnitude of N_1 is proportional to the number of potential visitors. We fixed N_1 in this work, however, the number of potential visitors could be affected by seasonal factors (such as weather, school terms, holidays) and economic and political factors (such as traffic expenditures, hotel fees, travelling policies [86]), thus N_1 could be time-dependent in reality.

Agreement and conflict between ρ and p In Eqns. (6.10), $\frac{f_\rho}{\lambda} = \rho N_1$ is the proportion of visitors (outside border and about to be inside border shortly) to the total population capacity. ρ (see Eqns. (6.5), (6.9) and (6.10)) is the proportion of potential visitors eventually becoming visitors correspond to the optimal travelling strategy selection. Therefore, we have $\rho = p^*$ (where p^* is individual's optimal travelling probability, see Eqn. (6.7), e.g., white area of $\rho^* = p^* = 1$ in Fig. 6.5) under normal scenario (i.e., no serious disease outbreak, in which there is no restriction on visitor entry). However, during a serious disease outbreak, the local government will consider restricting travelling entry (in order to lower the number of visitors inside border) according to population's optimal travelling proportion (i.e., ρ^*, see Eqn. (6.9)), and this would change $\rho = \min\{p^*, \rho^*\}$. Fig. 6.6(b) shows a numerical example of local governmental intervention on travelling entry (i.e., ρ). Note that, under governmental intervention scenario, ρ should only equal to ρ^* if $\rho^* < p^*$ (otherwise $\rho^* \geqslant p^*$, $\rho = p^*$ is equivalent to normal scenario), which is the red area in Fig. 6.5.

Period of visitors staying outside the border λ^{-1} is defined as the mean period of staying outside border for a visitor (see Table 6.1). Since we can divide the "travelling" populations as in Path (6.1), λ^{-1} represents the mean period for a visitor evolving from a "visitor outside" border to a "visitor inside" border. Note that a "potential visitor" can only become a "visitor outside" if he has made his last decision (i.e., travelling strategies selection), which means only a "potential visitor" who confirms to play "travelling" strategy (i.e., will travel eventually) can then, be regarded as a "visitor outside". One can have knowledge of the range of λ^{-1} by referring to the "deadline" of withdrawal of various travelling "services" (e.g., see note [4]). Therefore, the efficiency of disease risk information transmission could be related to λ^{-1} because that updating of relevant information can "update" individual's last decision (i.e., force individual to re-choose strategy), then higher efficiency of information transmission is corresponding to lower value of λ^{-1}.

Fig. 6.9 shows the relationships between relative risk (r, see Eqn. (6.5)), rate of visitors pass border (λ, and λ^{-1} is the mean period of a visitor staying outside border, see Table 6.1) and $\Delta\rho$ (Eqn. (6.14)) during an epidemic (i.e., $\alpha = 1$). When λ increases (i.e., visitors' pass-border rate rises), the discrepancy ($\Delta\rho$) of individual and group optimum (p^* and ρ^* respectively) appears under a wider range of relative risk (r). The discrepancy ($\Delta\rho$) shifts towards left (the direction r increases) as $\mathcal{R}_0$ increasing. Particularly, p^* and ρ^* meet

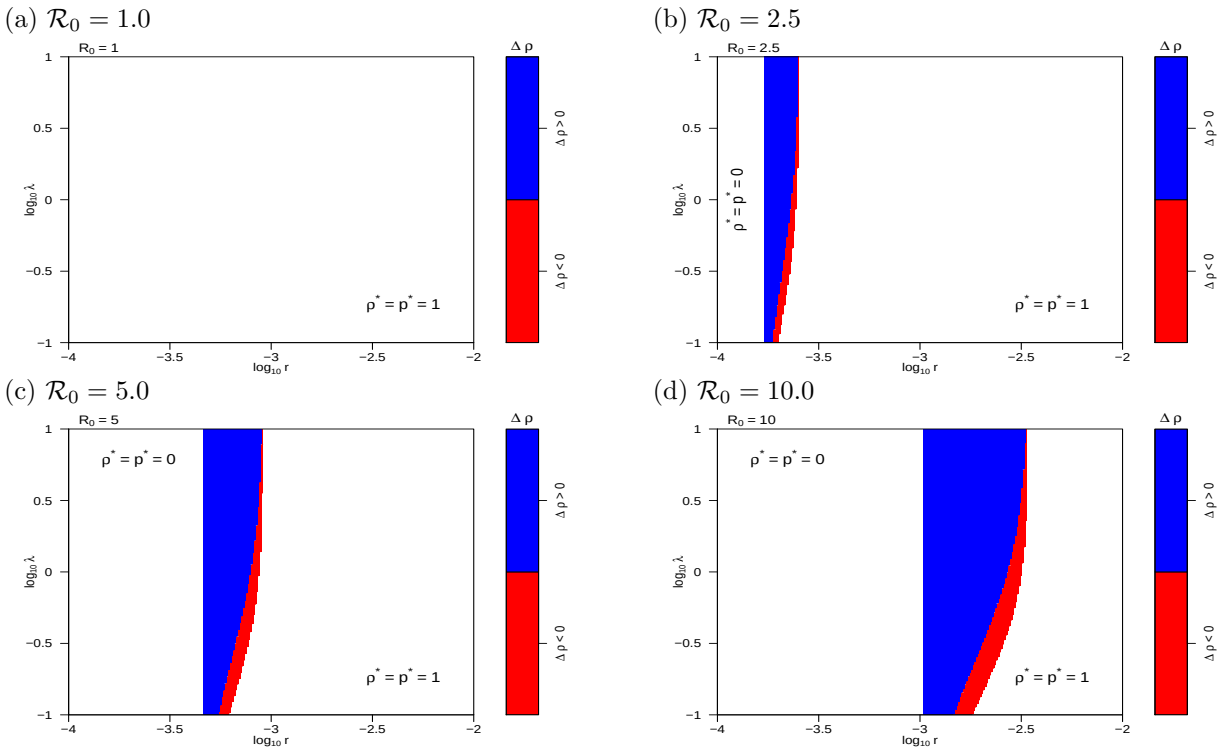


Figure 6.9: The relationships between r (Eqn. (6.5)), λ and $\Delta\rho$ (Eqn. (6.14)) during epidemic (i.e., $\alpha = 1$) with $\mathcal{R}_0 = 1.0, 2.5, 5.0, 10.0$ for panel (a)-(d) respectively. The values of r and λ are in "$\log_{10}$" form. The color code of the difference of individual and population strategy, $\Delta\rho$, is shown in the color key. The white area (in each panel) represents $\Delta\rho = 0$ under two situations that $\rho^* = p^* = 0$ or 1. The values of other parameters are in Table 6.1.

agreement (i.e., no discrepancy as $\rho^* = p^* = 1$) when $\mathcal{R}_0 = 1.0$ (which means disease cannot spread). The change of other parameters' values will not change the trend of this relationship.

6.4.7 Model Limitations and Further Discussion

Risk of disease spread at origin We have assume that "infected visitors do not pass the disease in the origin" (in Eqn. (6.8)). Nevertheless, there exist some chances that the disease could be "passed" from travelling destination to the origin via visitors. To address this point, one additional probabilistic factor of risk level are needed in Eqn. (6.8), hence, the improved travelling risk function $\Upsilon = \Upsilon(\rho, \pi) = N_1 \cdot \Big[\rho \cdot \alpha\phi(\rho) \cdot (1 + \pi \cdot \frac{\varrho}{r_1}) \cdot r_1 + (1 - \rho) r_0\Big]$. Here, π is the average probability of the disease is "passed" to one of the original places, and ϱ is the average payoff of disease spreading in one (i.e., a random-selected one) of the origin. Generally, $\varrho > r_1$ because, speaking from utilitarianism, the consequence of disease spreading in regional level are presumed to be more serious than one individual infected.

One can fix the ratio of $\frac{\varrho}{r_1}$ (the similar idea as $r = \frac{r_0}{r_1}$), and term $(1 + \pi \cdot \frac{\varrho}{r_1})$ is considered as a scalar (by further assigning a value to π), so that the results of original framework still hold. In this work, the simplified version of epidemic risk level of travelling population (as of Eqn. (6.8)) can be interpreted as $\pi = 0$ as assumed.

Pre-existing immunity among visitors In model (6.10), we assumed that visitors are totally susceptible when entering the population. In practice, there is probably pre-existing immunity among visitors during an on-going epidemic, and the protection level of visitors' population could be obtained from previous outbreaks and/or vaccination programs. We denote P_T as the proportion of the disease protections out of visitors population, P_O as the non-visitors population of the original places (i.e., where the visitors come from), and $P_D = P$ (as the same as in Eqn. (6.2)) as for the local population of travelling destination. Then, it is unnecessary that $P_D = P_T$ or $P_D = P_O$ (i.e., the disease protection level of travelling destination are probably not as the same as that of the origin). Moreover, it is also unnecessary that $P_T = P_O$ because, considering the effects of human behavior, individuals with disease protection (mostly obtained by vaccination) are relatively more likely to travel than those without protection (or do not know they are under protection), thus we could have $P_T \geqslant P_O$ for the most of situations.

Provided the knowledge of P_T, the revised epidemic model (6.10) is

$$\begin{cases} S' = \mu \cdot (1 - K_1 - S) - \beta S \cdot (I + I_1) \\ S_1' = (1 - P_T) f_\rho \cdot \left[1 - \dfrac{S_1 + I_1 + R_1}{(1 - P_T) K_1}\right] - \beta S_1 \cdot (I + I_1) - \nu S_1 \\ I' = \beta S \cdot (I + I_1) - (\gamma + \mu) I \\ I_1' = \beta S_1 \cdot (I + I_1) - (\gamma + \nu) I_1 \\ R' = \gamma I - \mu R \\ R_1' = \gamma I_1 - \nu R_1 \end{cases}$$

with all terms unchanged except for including two factors $(1-P_T)$ in $(1-P_T) f_\rho \cdot \left[1 - \frac{S_1+I_1+R_1}{(1-P_T)K_1}\right]$. We note that, in principle, there is supposed to be one more equation: $X_1' = P_T f_\rho \cdot \left[1 - \frac{S_1+I_1+R_1+X_1}{(1-P_T)K_1}\right] - \nu X_1$, where the additional state X_1 denotes the visitors being protected from disease, and the term $\frac{S_1+I_1+R_1}{(1-P_T)K_1}$ (in the revised model) should originally be written as $\frac{S_1+I_1+R_1+X_1}{K_1}$ (same as in Eqn. X_1'). Since the magnitude of both I_1 and R_1 are relatively

small with respect to S_1 and X_1, we ignore the effect of I_1 and R_1 on the visitors incoming rate, thus we have $S_1' \approx (1-P_T)f_\rho \cdot \left[1 - \frac{S_1+I_1+R_1+X_1}{K_1}\right] - \nu S_1$. We can easily see that P_T of f_ρ joins in X_1, $(1-P_T)$ of f_ρ joins in S_1, and the leaving rates of X_1 and S_1 are same as ν. To eliminate term X_1, we have $X_1 \approx \frac{P_T S_1}{(1-P_T)}$, and therefore, $\frac{S_1+I_1+R_1+X_1}{K_1} \approx \frac{S_1+I_1+R_1}{(1-P_T)K_1}$ as appeared in the above revised model.

The term $(1-P_T)$ could be interpreted as that the protected visitors (P_T) are directly removed from the system (not by joining R_1, but "completely" removed from the model system), and the effect on visitors incoming rate is partially reflected by "reducing" the local visitors' capacity (i.e., replacing K_1 by $(1-P_T)K_1$). In this work, P_T is fixed to be 0 as assumed. Then, the revised model (6.12) can also be derived following the same way from model (6.10) to (6.12). Since we regard P_T as a fixed non-zero constant (i.e., $P_T \neq 0$) during a short time period, and in mathematical terms, the effect of P_T can be transform into the reduction of the magnitude of f_ρ and K_1 [94], thus the main results in this work will hold for the revised epidemic model.

Difference and delay between real situation and human perspective We treated $\mathcal{R}_0$ as the real basic reproduction number (or the real risk level) of a disease of the target epidemic. Due to human perspective is probably different from the real situations (i.e., an imbalance between perception and reality) [56, 207, 209], we denote $\widetilde{\mathcal{R}_0}$ as the perceived "reproduction number" (i.e., the perceived risk level) of a disease. Normally, $\widetilde{\mathcal{R}_0}$ is positively correlated with $\mathcal{R}_0$ for most of cases because human's perspective is based on facts, thus we have $\widetilde{\mathcal{R}_0}(\mathcal{R}_0)$ is a non-decreasing function of $\mathcal{R}_0$. Provided the perceived disease risk ($\widetilde{\mathcal{R}_0}$), the payoff of disease risk ($r_1\left(\widetilde{\mathcal{R}_0}\right)$, i.e., r_1 as a function of $\widetilde{\mathcal{R}_0}$, as of Eqn. (6.2)) is expected to be a non-decreasing function of $\widetilde{\mathcal{R}_0}$, and thus also non-decreasing with respect to $\mathcal{R}_0$. One simple form of $r_1\left(\widetilde{\mathcal{R}_0}\right)$ could be modelled as $r_1 \propto \widetilde{\mathcal{R}_0}$ with a positive scalar.

In addition, there could be time delay between $\mathcal{R}_0$ and $\widetilde{\mathcal{R}_0}$ due that travelling-players may not always be informed timely with disease events, thus we wrote $\widetilde{\mathcal{R}_0}(t;\tau) = \widetilde{\mathcal{R}_0}(\mathcal{R}_0(t-\tau))$ where $\tau \geqslant 0$ is the time lag between occurrence of facts and human perceived knowledge of facts. If we set $\tau = 0$ for all t by assuming human receiving the correct knowledge of a disease event while it occurs, we have $\lim_{\tau\to 0^+} \widetilde{\mathcal{R}_0}(t;\tau) = \widetilde{\mathcal{R}_0}(\mathcal{R}_0(t))$. In this work, we considered a limiting case of $\tau = 0$, in reality it is likely to be nonzero and it depends on many factors. The value of τ depends on the impacts and efficiency of the relevant

information transmission (e.g., news press coverage [60, 85, 87, 112], educational programs [59, 60, 169], communication effectiveness in social networks [58, 59, 106, 121, 191, 206] and public health awareness [58, 84, 106], etc.).

Various patchy risk preference In this work, we assumed the same information set as well as the same strategic response of the whole visitor's population (see Eqn. (6.2) and (6.3)). There could be different perspective (i.e., risk preference) of different groups of people, and various perspective on costs and benefits (i.e., payoff) in the context of game theory which may lead to different results of equilibria and optima regarding to human response to epidemics [195, 196]. Consider the situation that $E_1 = E_0$ (see Eqns. (6.2) and (6.3)), individuals may prefer "travelling" strategy (i.e., risk seeking preference), however, others may prefer "non-travelling" strategy (i.e., risk averse preference). Future studies with more detailed patchy preferential environment could improve the rationale of the initial settings in this work.

6.4.8 Conclusions

Many game theoretical studies of closed socio-epidemiological systems find a significant discrepancy between individual and group optima across a broad range of economic and epidemiological parameters values. However, in this work, we studied an open socio-ecological system where visitors decide whether or not to travel to a location with an on-going outbreak, and the local government of the outbreak decides whether or not let more visitors in, and surprisingly we found perfect agreement between individual and group optimal strategies across similarly broad ranges of parameters values. When a discrepancy exists between the individual and group optimal strategies, their conflict is likely to be very large and highly sensitive to small changes of disease transmissibility and visitor costs/benefits. For instance, if disease transmissibility rises even a small amount, uncontrolled visitor inflow is capable of causing an unexpected outbreak. This suggests that a conflict between individual and group optima could emerge suddenly in real-world settings, provided slight changes in economic and epidemiological parameters. However, timely implementation of travel restriction by governments may effectively prevent large-scale outbreaks. The earlier the restrictions are implemented, the better the outcomes will be.

Chapter 7

Summary and Future Works

This chapter will summarize the major findings of this thesis and suggest some directions for further research.

7.1 Summary

Infectious diseases pose a serious health risk to human society due to their potential for rapid spread between countries and continents and the increasing population density. This thesis shed lights on the understanding of the transmission patterns of yellow fever, Japanese encephalitis virus, Zika virus disease and varicella focusing on explore and explain the risk level of infectious diseases, how virus transmission will be affected the potential non-epidemical factors, how to evaluate the effects of public health control measures and the potential cause of the complex epidemic dynamics. The proposed model framework in this thesis will be applicable to many other vector-borne diseases (and many infectious diseases) that are affected by other non-epidemical risk factors, and the framework will be of value to other works in mathematical epidemiology field. Overall, this work contributes to the understanding of the vector-borne diseases dynamics and childhood infections, and provides theoretical backup for the design of effective prevention and control strategies.

7.2 Future works

Current modelling works on yellow fever epidemics focuses on the regional mosquitos' dynamics as will as the effects of vaccination programs. However, the factors including "human behaviors" proposed in this thesis (see Chapter 2 section 2.3.3) implies the possibility of short-term human reactions to vector-borne diseases, which reflects on the population perspective to yellow fever deaths data. Future works could look into details on this sides.

The JEV modelling framework avoids including mosquitos' population explicitly, instead, the mosquitos' dynamics are modelled based on regional ovitrap index. The human cases are simulated according to pig's infection with a spill-over rate. The framework simplified the biological mechanism of JEV transmission and fitted the observed data well. Further studies focusing on the possible newly invade JEV strain are needed.

The model of Zika virus disease contains both vector-borne transmission path and the vector-free (i.e., human sesxual) transmission path, as well as the possible infectivity of asymptomatic infections. The studies in the future may consider the effects (i.e., proportion and infectivity) of asymptomatic infections and the biological supports on this side are necessary.

Since school-based varicella vaccination program can prevent large-scale outbreak among schools children, further works focusing on the associated cost-effective evaluation are worthwhile. Network framework can be further included into the agent-based model to explore the potential spatial pattern of the varicella epidemics.

The proposed simplified travelling game theory associated with simple epidemic model framework is well-discussed in this thesis (please see Chapter 6). Furthermore, an interesting idea (proposed by Prof. Chris T. Bauch) for future work is that there are two games going on at the same time. Then, we have two patches where individuals in each patch can adopt a traveller or non-traveller strategy, for travel to the other patch. At the same time, each government plays a game of restrict/do not restrict travel. The Nash equilibrium government strategy depends on what is optimal for the individuals, and vice versa. To the best of my knowledge, this approach has not been done before so far, but it is for sure to leave for the future.

One of the major and common challenges about the modelling framework adopted in

this thesis is to balance the model complexity and the its fitting performance. Although several criteria were implemented to measure the tradeoff between the the model complexity and performance, numbers of approaches can be considered to further avoid overfitting. In addition to the methods in this thesis, one also could:

- determine the values of model parameters from biological or medical studies,
- simplify the models based on soundable reasons,
- cross validate the fitting performance,
- include various sensitivity analysis,
- consider more modelling scenarios based on different model parameters.

Bibliography

[1] According to the monthly travelling statistics (`http://partnernet.hktb.com/en/research_statistics/latest_statistics/index.html`) and travelling summary sheet (`http://partnernet.hktb.com/en/research_statistics/index.html`) from partnernet—hong kong tourism website for travel trade partners, there were approximate 58,000,000 travelers in hong kong of 2015 or 2016, and the local hotel room occupancy is roughly 87% over the whole period of time. provided the information in [3], the local tavelling capacity of hong kong can be estimated as $\mathcal{N}k_1 = \frac{(58000000/87\%)\times 3}{365} \approx 550000$, here $\mathcal{N}$ denotes the number of total population capacity in hong kong (i.e., the sumation of upper bound of the number of travelers and local population, $\mathcal{N} = \mathcal{N}k_1 + \mathcal{N}\cdot(s+i+r)$). given the popualtion statistics from world bank (`https://data.worldbank.org/indicator/SP.POP.TOTL?locations=HK`), 7,300,000 is the number of local population in hong kong in 2015-16, thus $\mathcal{N} = \mathcal{N}k_1 + \mathcal{N}\cdot(s+i+r) = 550000+7300000 = 7850000$, and $k_1 = \frac{550000}{7850000} \approx 7.0\%$.

[2] The range of relative risk (r) can be approximated by simply checking the claim settlement odds of the travel insurance corresponding to the target place. for an example, according to travel insurance premium and coverage websites of hang seng bank (`https://bank.hangseng.com/1/2/personal/insurance/travel-leisure/travel-insurance/travel-premium` and `https://bank.hangseng.com/1/2/personal/insurance/travel-leisure/travel-insurance/travel-coverage`), $r \approx 10^{-3}$.

[3] Referring to immigration department of the government of hong kong (`http://www.immd.gov.hk/eng/services/visas/visit-transit/visit-visa-entry-permit.html`), **Chinese** citizens can stay in hong kong for at maximal 7 days, and the majority of **non-Chinese** citizens can stay for roughly at maximal 15 days. according to the monthly travelling statistics from partnernet - hong kong tourism website for travel trade partners (`http://partnernet.hktb.com/en/research_statistics/latest_statistics/index.html`), averagely, 75% of travelers are from mainland china and 25% are from other regions; for **Chinese** travelers, 50% of them are overnight passengers (expected to stay for $\frac{1}{2}\times(7+1) = 4$ days) and 50% them are sameday visitors (expected to stay for $\frac{1}{2}\times(0+1) = 0.5$ day); for **non-Chinese** travelers, 66.67% of them are overnight passengers (expected to stay for $\frac{1}{2}\times(15+1) = 8$ days) and 33.33% them are sameday visitors (expected to stay for $\frac{1}{2}\times(0+1) = 0.5$ day). therefore, on average, one random-selected traveler would be expected to stay

in hong kong for $\nu^{-1} = 75\% \times (50\% \times 4 + 50\% \times 0.5) + 25\% \times (\frac{2}{3} \times 8 + \frac{1}{3} \times 0.5) \approx 3$ days (thus, ν^{-1} are set to be 3 days).

[4] The value of the mean period of a traveler stay outside border (λ^{-1}) can be estimated by referring to the "deadline" of cancellation of hotel room, flight or even car-rent for travelling usage. for example, according to cancellation policies of airbnb (`https://www.airbnb.com/home/cancellation_policies`), the waiver of refund charges can be considered for room-cancellation at least 1 day (with "flexible" policy) or 5 days (with "moderate" policy) in advance, thus we can make a rough estimation that $\lambda^{-1} \approx \frac{1}{2} \times (1 + 5) = 3$ days. according to hong kong airline refund policies (`http://www.hongkongairlines.com/en_HK/flight/refund`), $\lambda^{-1} > 2$ days.

[5] European centre for disease prevention and control. mission report: Yellow fever outbreak in angola., 2016.

[6] Shenzhen center for diseases control and prevention. available from: `http://www.szcdc.net/SZCDC/`, Jun 2016.

[7] Shenzhen education bureau. available from: `http://www.szeb.edu.cn/enportal/`, Jun 2016.

[8] African health observatory (aho). publications website, health topic: Yellow fever. yellow fever outbreak in angola situation reports. `http://www.afro.who.int/publications`, July 2017.

[9] African health observatory (aho). yellow fever outbreak in angola situation report 26 feb 2016. `http://www.afro.who.int/sites/default/files/2017-06/sitrep_ang_feb_26_2016.pdf`, July 2017.

[10] Centers for disease control and prevention (cdc) japanese encephalitis fact sheet website. `https://www.cdc.gov/japaneseencephalitis/index.html`, Jul 2017.

[11] Centers for disease control and prevention (cdc), webpage of "key facts about influenza (flu)". `https://www.cdc.gov/flu/keyfacts.htm`., Nov 2017.

[12] Centers for disease control and prevention. yellow fever. `https://www.cdc.gov/yellowfever/`, March 2017.

[13] The government of hong kong, agriculture, fisheries and conservation department (afcd), annual report official website. `http://www.afcd.gov.hk/misc/download/annualreport2012/agriculture.html`, Aug 2017.

[14] The government of hong kong, agriculture, fisheries and conservation department (afcd), import live pigs official website. `https://www.afcd.gov.hk/English/quarantine/qua_ie/qua_ie_ipab/qua_ie_ipab_ilpbp/qua_ie_ipab_ilpbp.html`, Jul 2017.

[15] The government of hong kong, agriculture, fisheries and conservation department (afcd) official website. `http://www.afcd.gov.hk/eindex.html`, Jul 2017.

[16] The government of hong kong, center for health protection (chp), japanese encephalitis fact sheet website. http://www.chp.gov.hk/en/content/9/24/28.html, Jul 2017.

[17] The government of hong kong, center for health protection (chp), statistics on communicable diseases official website. http://www.chp.gov.hk/en/notifiable1/10/26/43.html, Jul 2017.

[18] The government of hong kong, food and environmental hygiene department (fehd), monthly average daily supply and auction prices of live pigs official website. http://www.fehd.gov.hk/tc_chi/sh/data/supply_avg_tw.html, Aug 2017.

[19] The government of hong kong, food and environmental hygiene department (fehd), press release official website. http://sc.isd.gov.hk/gb/www.info.gov.hk/gia/general/201011/04/P201011040289.htm and http://www.fehd.gov.hk/english/index.html, Jul 2017.

[20] The government of hong kong, food and environmental hygiene department, pest control website. http://www.fehd.gov.hk/english/safefood/dengue_fever/index.html, Jul 2017.

[21] The government of hong kong, scientific committee on vector-borne diseases official website. http://www.chp.gov.hk/en/sas7/101/110/107.html, Jul 2017.

[22] The government of people's republic of china, national health and family planning commission (nhfpc), control of communicable diseases official website. http://www.nhfpc.gov.cn/jkj/s3578/new_list.shtml, Jul 2017.

[23] The government of r.o.c. (taiwan), taiwan national infectious disease statistics system official website. https://nidss.cdc.gov.tw/en/Default.aspx, Jul 2017.

[24] The government of taiwan, centers for disease control, r.o.c. (taiwan), japanese encephalitis fact sheet website. http://www.cdc.gov.tw/english/info.aspx?treeid=E79C7A9E1E9B1CDF&nowtreeid=E02C24F0DACDD729&tid=784CF7F674BD5DFA, Jul 2017.

[25] Guillain-barre syndrome fact sheets from world health organization official website. http://www.who.int/mediacentre/factsheets/guillain-barre-syndrome/en/, Jun 2017.

[26] Hong kong 01, press release. https://www.hk01.com/%E6%B8%AF%E8%81%9E/35799/-%E6%AF%92%E8%B1%AC%E8%82%89-%E6%83%B3%E9%A3%9F%E6%96%B0%E9%AE%AE-%E6%B8%AF%E8%B1%AC-%E6%A5%AD%E7%95%8C%E6%95%99%E8%B7%AF%E7%9D%87%E8%B1%AC%E7%9A%AE%E8%97%8D%E5%8D%B0%E5%88%86%E8%BE%A8%E4%BE%%86%E6%BA%90%E5%9C%B0, Aug 2017.

[27] Imperial college, london. yellow fever immunization coverage across africa. https://polici.shinyapps.io/yellow_fever_africa/, March 2017.

[28] International society for infectious diseases. program for monitoring emerging diseases (promed). http://www.promedmail.org/, March 2017.

[29] Legislative council of hong kong, official report (cb(2) 1663/05-06(05)). https://www.google.com.hk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&ved=0ahUKEwipgdWfn9HVAhVDKpQKHQrIAScQFghJMAU&url=http%3A%2F%2Fwww.legco.gov.hk%%2Fyr05-06%2Fchinese%2Fpanels%2Ffseh%2Fpapers%2Ffe0411cb2-1663-05-c.pdf&usg=AFQjCNHEbGyXdsLbJAjID3RDg-TdYLkWJQ., Aug 2017.

[30] Legislative council of hong kong, official report (cb(2)2445/06-07(01)). https://www.google.com.hk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=0ahUKEwjPh7flndHVAhWBxJQKHRJeBvoQFgg-MAM&url=http%3A%2F%2Fwww.legco.gov.hk%%2Fyr06-07%2Fchinese%2Fpanels%2Ffseh%2Fpapers%2Ffe0710cb2-2445-1-c.pdf&usg=AFQjCNEip77zuVMBUTktaWQNOr1LtgE9vQ., Aug 2017.

[31] Mulesoft. city population.https://www.citypopulation.de/, March 2017.

[32] The news press of national tourism administration (see (in chinese) http://www.people.com.cn/GB/jingji/1038/1970533.html) and who travel advice (see http://www.who.int/csr/sars/archive/2003_04_23/en/)., Nov 2017.

[33] News press of the reveal of 2003 sars epidemic in beijing, china: http://www.lifeweek.com.cn/2003/0729/5582.shtml (in chinese) and http://news.eastday.com/epublish/gb/paper148/20030530/class014800003/hwz953568.htm., Nov 2017.

[34] President's malaria initiative operational plan, angola. https://www.pmi.gov/docs/default-source/default-document-library/malaria-operational-plans/fy17/fy-2017-angola-malaria-operational-plan.pdf?sfvrsn=7, July 2017.

[35] South china morning post website. http://www.scmp.com/news/hong-kong/health-environment/article/2103667/hong-kong-tackles-worlds-first-case-patient, Aug 2017.

[36] Ta kung pao in hong kong, press release. https://www.google.com.hk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=10&cad=rja&uact=8&ved=0ahUKEwi3kcGul83VAhXJJJQKHe_HDBsQFghtMAk&url=http%3A%2F%2Fpaper.takungpao.com%2Fresfile%2F2013-01-14%2FA02%2FA2_Screen.pdf&usg=AFQjCNFaZruvkl3gCzaPB1o46so7IyMpbg, Jul 2017.

[37] United states department of agriculture (usda), global agriculture information network (gain) reports. http://www.usfoods-hongkong.net/res/mns/00321/HK1330.pdf and http://www.usfoods-hongkong.net/res/mns/00433/HK1524.pdf, Sep 2017.

[38] http://www.huaxia.com/thpl/wyps/4146328.html, Jun 2017.

[39] http://www.reuters.com/article/us-health-zika-brazil-exclusive-idUSKCN0VA331, Jun 2017.

[40] Wikipedia of northeast region, brazil. `https://en.wikipedia.org/wiki/Northeast_Region,_Brazil`, Jun 2017.

[41] Winnipeg regional health authority (wrha). travel health and tropical medicines services. `http://www.wrha.mb.ca/community/travel/`, March 2017.

[42] World health organization. emergencies preparedness, response. yellow fever case definitions for public health surveillance. `http://www.who.int/csr/disease/yellowfev/case-definition/en/`, May 2017.

[43] World health organization (who) japanese encephalitis fact sheet website. `http://www.who.int/mediacentre/factsheets/fs386/en/`, Jul 2017.

[44] World health organization (who) official website, zika virus fact sheet. `http://www.who.int/mediacentre/factsheets/zika/en/`, Jun 2017.

[45] World health organization. yellow fever fact sheet. `http://www.who.int/mediacentre/factsheets/fs100/en/`, March 2017.

[46] World health organization. yellow fever knowledge pack. `http://www.who.int/risk-communication/yellow-fever/YF-Knowledge-pack.pdf?ua=1`, April 2017.

[47] Rama S Akondy, Philip LF Johnson, Helder I Nakaya, Srilatha Edupuganti, Mark J Mulligan, Benton Lawson, Joseph D Miller, Bali Pulendran, Rustom Antia, and Rafi Ahmed. Initial viral load determines the magnitude of the human cd8 t cell response to yellow fever vaccination. *Proceedings of the National Academy of Sciences*, 112(10):3050–3055, 2015.

[48] Linda JS Allen, Fred Brauer, Pauline Van den Driessche, and Jianhong Wu. *Mathematical epidemiology*, volume 1945. Springer, 2008.

[49] Linda JS Allen and David B Thrasher. The effects of vaccination in an age-dependent model for varicella and herpes zoster. *IEEE transactions on automatic control*, 43(6):779–789, 1998.

[50] Roy M Anderson, Robert M May, and B Anderson. *Infectious diseases of humans: dynamics and control*, volume 28. Wiley Online Library, 1992.

[51] Mathieu Andraud, Niel Hens, Christiaan Marais, and Philippe Beutels. Dynamic epidemiological models for dengue transmission: a systematic review of structural approaches. *PloS one*, 7(11):e49085, 2012.

[52] Andrea Apolloni, Chiara Poletto, José J Ramasco, Pablo Jensen, and Vittoria Colizza. Metapopulation epidemic models with heterogeneous mixing and travel behaviour. *Theoretical Biology and Medical Modelling*, 11(1):3, 2014.

[53] Satoru Arai, Yasuko Matsunaga, Tomohiko Takasaki, Keiko Tanaka-Taya, Kiyosu Taniguchi, Nobuhiko Okabe, Ichiro Kurane, Vaccine Preventable Diseases Surveillance Program of Japan, et al. Japanese encephalitis: surveillance and elimination effort in japan from 1982 to 2004. *Jpn J Infect Dis*, 61(5):333–338, 2008.

[54] Maite Aubry, Anita Teissier, Michael Huart, Sébastien Merceron, Jessica Vanhomwegen, Claudine Roche, Anne-Laure Vial, Sylvianne Teururai, Sébastien Sicard, Sylvie Paulous, et al. Zika virus seroprevalence, french polynesia, 2014–2015. *Emerging infectious diseases*, 23(4):669, 2017.

[55] Alan DT Barrett. Yellow fever in angola and beyondxthe problem of vaccine supply and demand. *New England Journal of Medicine*, 375(4):301–303, 2016.

[56] Sanjay Basu, Gretchen B Chapman, and Alison P Galvani. Integrating epidemiology, psychology, and economics to achieve hpv vaccination targets. *Proceedings of the National Academy of Sciences*, 105(48):19018–19023, 2008.

[57] Chris Bauch, Alberto dOnofrio, and Piero Manfredi. Behavioral epidemiology of infectious diseases: an overview. In *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases*, pages 1–19. Springer, 2013.

[58] Chris T Bauch. Imitation dynamics predict vaccinating behaviour. *Proceedings of the Royal Society of London B: Biological Sciences*, 272(1573):1669–1675, 2005.

[59] Chris T Bauch and Samit Bhattacharyya. Evolutionary game theory and social learning can determine how vaccine scares unfold. *PLoS computational biology*, 8(4):e1002452, 2012.

[60] Chris T Bauch and David JD Earn. Vaccination and the theory of games. *Proceedings of the National Academy of Sciences of the United States of America*, 101(36):13391–13394, 2004.

[61] Chris T Bauch, Alison P Galvani, and David JD Earn. Group interest versus self-interest in smallpox vaccination policy. *Proceedings of the National Academy of Sciences*, 100(18):10564–10567, 2003.

[62] Chris T Bauch, James O Lloyd-Smith, Megan P Coffee, and Alison P Galvani. Dynamically modeling sars and other newly emerging respiratory illnesses: past, present, and future. *Epidemiology*, 16(6):791–801, 2005.

[63] WGC Bearcroft. Zika virus infection experimentally induced in a human volunteer. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 50(5):442–448, 1956.

[64] Rotem Ben-Shachar and Katia Koelle. Minimal within-host dengue models highlight the specific roles of the immune response in primary and secondary dengue infections. *Journal of the Royal Society Interface*, 12(103):20140886, 2015.

[65] JPT Boorman and JS Porterfield. A simple technique for infection of mosquitoes with viruses transmission of zika virus. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 50(3):238–242, 1956.

[66] Martin CJ Bootsma and Neil M Ferguson. The effect of public health measures on the 1918 influenza pandemic in us cities. *Proceedings of the National Academy of Sciences*, 104(18):7588–7593, 2007.

[67] Patrícia Brasil, Jose P Pereira Jr, M Elisabeth Moreira, Rita M Ribeiro Nogueira, Luana Damasceno, Mayumi Wakimoto, Renata S Rabello, Stephanie G Valderramos, Umme-Aiman Halai, Tania S Salles, et al. Zika virus infection in pregnant women in rio de janeiro. *N Engl J Med*, 2016(375):2321–2334, 2016.

[68] Fred Brauer, Carlos Castillo-Chavez, Anuj Mubayi, and Sherry Towers. Some models for epidemics of vector-transmitted diseases. *Infectious Disease Modelling*, 1(1):79–87, 2016.

[69] Fred Brauer and Pauline van den Driessche. Models for transmission of disease with immigration of infectives. *Mathematical Biosciences*, 171(2):143–154, 2001.

[70] M Brisson, WJ Edmunds, NJ Gay, B Law, and G De Serres. Modelling the impact of immunization on the epidemiology of varicella zoster virus. *Epidemiology & Infection*, 125(3):651–669, 2000.

[71] Marc Brisson, Gurgen Melkonyan, Melanie Drolet, Gaston De Serres, Roseline Thibeault, and Philippe De Wals. Modeling the impact of one-and two-dose varicella vaccination on the epidemiology of varicella and zoster. *Vaccine*, 28(19):3385–3397, 2010.

[72] MN Burattini, M Chen, A Chow, FAB Coutinho, KT Goh, LF Lopez, S Ma, and E Massad. Modelling the control strategies against dengue in singapore. *Epidemiology & Infection*, 136(3):309–319, 2008.

[73] Kenneth P Burnham and David R Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media, 2003.

[74] A Camacho, AJ Kucharski, S Funk, J Breman, P Piot, and WJ Edmunds. Potential for large outbreaks of ebola virus disease. *Epidemics*, 9:70–78, 2014.

[75] Lei Cao, Shihong Fu, Xiaoyan Gao, Minghua Li, Shiheng Cui, Xiaolong Li, Yuxi Cao, Wenwen Lei, Zhi Lu, Ying He, et al. Low protective efficacy of the current japanese encephalitis vaccine against the emerging genotype 5 japanese encephalitis virus. *PLoS neglected tropical diseases*, 10(5):e0004686, 2016.

[76] Van-Mai Cao-Lormeau, Alexandre Blake, Sandrine Mons, Stéphane Lastère, Claudine Roche, Jessica Vanhomwegen, Timothée Dub, Laure Baudouin, Anita Teissier, Philippe Larre, et al. Guillain-barré syndrome outbreak associated with zika virus infection in french polynesia: a case-control study. *The Lancet*, 387(10027):1531–1539, 2016.

[77] Simon Cauchemez, Marianne Besnard, Priscillia Bompard, Timothée Dub, Prisca Guillemette-Artur, Dominique Eyrolle-Guignot, Henrik Salje, Maria D Van Kerkhove, Véronique Abadie, Catherine Garel, et al. Association between zika virus and microcephaly in french polynesia, 2013–15: a retrospective study. *The Lancet*, 387(10033):2125–2132, 2016.

[78] MS Chakraborty, SK Chakravarti, KK Mukherjee, AC Mitra, et al. Inapparent infection by japanese encephalitis (je) virus in west bengal. *Indian journal of public health*, 24(3):121–127, 1980.

[79] Frederick H Chen. A susceptible-infected epidemic model with voluntary vaccinations. *Journal of mathematical biology*, 53(2):253–272, 2006.

[80] Eriko Chikaki and Hirofumi Ishikawa. A dengue transmission model in thailand considering sequential infections with all four serotypes. *The Journal of Infection in Developing Countries*, 3(09):711–722, 2009.

[81] Gerardo Chowell, Paul W Fenimore, Melissa A Castillo-Garsow, and Carlos Castillo-Chavez. Sars outbreaks in ontario, hong kong and singapore: the role of diagnosis and isolation as a control mechanism. *Journal of theoretical biology*, 224(1):1–8, 2003.

[82] Stanca M Ciupe and Jane M Heffernan. In-host modeling. *Infectious Disease Modelling*, 2017.

[83] Claudia T Codeço, Paula M Luz, Flavio Coelho, Alison P Galvani, and Claudio Struchiner. Vaccinating in disease-free regions: a vaccine model with application to yellow fever. *Journal of The Royal Society Interface*, 4(17):1119–1125, 2007.

[84] Daniel M Cornforth, Timothy C Reluga, Eunha Shim, Chris T Bauch, Alison P Galvani, and Lauren Ancel Meyers. Erratic flu vaccination emerges from short-sighted behavior in contact networks. *PLoS Computational Biology*, 7(1):e1001062, 2011.

[85] Jingan Cui, Yonghong Sun, and Huaiping Zhu. The impact of media on the control of infectious diseases. *Journal of dynamics and differential equations*, 20(1):31–53, 2008.

[86] Jingan Cui, Yasuhiro Takeuchi, and Yasuhisa Saito. Spreading disease with transport-related infection. *Journal of theoretical biology*, 239(3):376–390, 2006.

[87] Jingan Cui, Xin Tao, Huaiping Zhu, et al. An sis infection model incorporating media coverage. *Rocky Mountain J. Math*, 38(5):1323–1334, 2008.

[88] KM De Cock, A Nasidi, J Enriquez, RB Craven, BC Okafor, TP Monath, PM Tukei, P Lichfield, A Fabiyi, C Ravaonjanahary, et al. Epidemic yellow fever in eastern nigeria, 1986. *The Lancet*, 331(8586):630–633, 1988.

[89] Wanderson K de Oliveira, Eduardo H Carmo, Claudio M Henriques, Giovanini Coelho, Enrique Vazquez, Juan Cortez-Escalante, Joaquin Molina, Sylvain Aldighieri, Marcos A Espinal, and Christopher Dye. Zika virus infection and associated neurologic disorders in brazil. *New England Journal of Medicine*, 376(16):1591–1593, 2017.

[90] Emanuele Del Fava, Grazina Rimseliene, Elmira Flem, Birgitte Freiesleben De Blasio, Gianpaolo Scalia Tomba, and Piero Manfredi. Estimating age-specific immunity and force of infection of varicella zoster virus in norway using mixture models. *PLoS one*, 11(9):e0163636, 2016.

[91] Kamal Desai, L Coudeville, and F Bailleux. Modelling the long-term persistence of neutralizing antibody in adults after one dose of live attenuated japanese encephalitis chimeric virus vaccine. *Vaccine*, 30(15):2510–2515, 2012.

[92] Mawlouth Diallo, Jocelyn Thonnon, and Didier Fontenille. Vertical transmission of the yellow fever virus by aedes aegypti (diptera, culicidae): dynamics of infection in f1 adult progeny of orally infected females. *The American journal of tropical medicine and hygiene*, 62(1):151–156, 2000.

[93] Mark R Duffy, Tai-Ho Chen, W Thane Hancock, Ann M Powers, Jacob L Kool, Robert S Lanciotti, Moses Pretrick, Maria Marfel, Stacey Holzbauer, Christine Dubray, et al. Zika virus outbreak on yap island, federated states of micronesia. *New England Journal of Medicine*, 360(24):2536–2543, 2009.

[94] David JD Earn, Pejman Rohani, Benjamin M Bolker, and Bryan T Grenfell. A simple model for complex dynamical transitions in epidemics. *Science*, 287(5453):667–670, 2000.

[95] A Economopoulou, M Dominguez, B Helynck, D Sissoko, Ole Wichmann, P Quenel, P Germonneau, and I Quatresous. Atypical chikungunya virus infections: clinical manifestations, mortality and risk factors for severe disease during the 2005–2006 outbreak on reunion. *Epidemiology & Infection*, 137(4):534–541, 2009.

[96] Sascha R Ellington, Owen Devine, Jeanne Bertolli, Alma Martinez Quiñones, Carrie K Shapiro-Mendoza, Janice Perez-Padilla, Brenda Rivera-Garcia, Regina M Simeone, Denise J Jamieson, Miguel Valencia-Prado, et al. Estimating the number of pregnant women infected with zika virus and expected infants with microcephaly following the zika virus outbreak in puerto rico, 2016. *JAMA pediatrics*, 170(10):940–945, 2016.

[97] Joshua M Epstein, D Michael Goedecke, Feng Yu, Robert J Morris, Diane K Wagener, and Georgiy V Bobashev. Controlling pandemic flu: the value of international air travel restrictions. *PloS one*, 2(5):e401, 2007.

[98] Elina O Erra, Helena Hervius Askling, Sutee Yoksan, Lars Rombo, Jukka Riutta, Sirkka Vene, Lars Lindquist, Olli Vapalahti, and Anu Kantele. Cross-protective capacity of japanese encephalitis (je) vaccines against circulating heterologous je virus genotypes. *Clinical Infectious Diseases*, 56(2):267–270, 2012.

[99] Dan Feng, Sake J De Vlas, Li-Qun Fang, Xiao-Na Han, Wen-Juan Zhao, Shen Sheng, Hong Yang, Zhong-Wei Jia, Jan Hendrik Richardus, and Wu-Chun Cao. The sars epidemic in mainland china: bringing together all epidemiological data. *Tropical Medicine & International Health*, 14(s1):4–13, 2009.

[100] Paul E M Fine and Jacqueline A Clarkson. Individual versus public priorities in the determination of optimal vaccination policies. *American journal of epidemiology*, 124(6):1012–1020, 1986.

[101] Paul EM Fine and Jacqueline A Clarkson. Measles in england and wales – i: an analysis of factors underlying seasonal patterns. *International journal of Epidemiology*, 11(1):5–14, 1982.

[102] Paul EM Fine and Jacqueline A Clarkson. Measles in england and wales – i: an analysis of factors underlying seasonal patterns. *International Journal of Epidemiology*, 11(1):5–14, 1982.

[103] Paul EM Fine and Jacqueline A Clarkson. Measles in england and wales – ii: the impact of the measles vaccination programme on the distribution of immunity in the population. *International Journal of Epidemiology*, 11(1):15–25, 1982.

[104] David Fisman, Edwin Khoo, and Ashleigh Tuite. Early epidemic dynamics of the west african 2014 ebola outbreak: estimates derived with a simple two-parameter model. *PLoS currents*, 6, 2014.

[105] Chuanxi Fu, Ming Wang, Jianhua Liang, Jianxiong Xu, Chengbin Wang, and Stephanie Bialek. The effectiveness of varicella vaccine in china. *The Pediatric infectious disease journal*, 29(8):690–693, 2010.

[106] Sebastian Funk, Erez Gilad, Chris Watkins, and Vincent AA Jansen. The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences*, 106(16):6872–6877, 2009.

[107] Sebastian Funk, Gwenan M Knight, and Vincent AA Jansen. Ebola: the power of behaviour change. *Nature*, 515(7528):492–492, 2014.

[108] Sebastian Funk, Adam J Kucharski, Anton Camacho, Rosalind M Eggo, Laith Yakob, Lawrence M Murray, and W John Edmunds. Comparative analysis of dengue and zika outbreaks reveals differences by setting and virus. *PLoS neglected tropical diseases*, 10(12):e0005173, 2016.

[109] Sebastian Funk, Marcel Salathé, and Vincent AA Jansen. Modelling the influence of human behaviour on the spread of infectious diseases: a review. *Journal of the Royal Society Interface*, page rsif20100142, 2010.

[110] Daozhou Gao, Abdou Amza, Baidou Nassirou, Boubacar Kadri, Nicholas Sippl-Swezey, Fengchen Liu, Sarah F Ackley, Thomas M Lietman, and Travis C Porco. Optimal seasonal timing of oral azithromycin for malaria. *The American journal of tropical medicine and hygiene*, 91(5):936–942, 2014.

[111] Daozhou Gao, Yijun Lou, Daihai He, Travis C Porco, Yang Kuang, Gerardo Chowell, and Shigui Ruan. Prevention and control of zika as a mosquito-borne and sexually transmitted disease: a mathematical modeling analysis. *Scientific reports*, 6:28070, 2016.

[112] Daozhou Gao and Shigui Ruan. An sis patch model with variable transmission coefficients. *Mathematical biosciences*, 232(2):110–115, 2011.

[113] Xiaoyan Gao, Hong Liu, Minghua Li, Shihong Fu, and Guodong Liang. Insights into the evolutionary history of japanese encephalitis virus (jev) based on whole-genome sequences comprising the five genotypes. *Virology journal*, 12(1):43, 2015.

[114] Xiaoyan Gao, Hong Liu, Huanyu Wang, Shihong Fu, Zhenyang Guo, and Guodong Liang. Southernmost asia is the source of japanese encephalitis virus (genotype 1) diversity from which the viruses disperse and evolve throughout asia. *PLoS neglected tropical diseases*, 7(9):e2459, 2013.

[115] Tini Garske, Maria D Van Kerkhove, Sergio Yactayo, Olivier Ronveaux, Rosamund F Lewis, J Erin Staples, William Perea, Neil M Ferguson, Yellow Fever Expert Committee, et al. Yellow fever in africa: estimating the burden of disease and impact of mass vaccination from outbreak and serological data. *PLoS medicine*, 11(5):e1001638, 2014.

[116] Timothy C Germann, Kai Kadau, Ira M Longini, and Catherine A Macken. Mitigation strategies for pandemic influenza in the united states. *Proceedings of the National Academy of Sciences*, 103(15):5935–5940, 2006.

[117] Giel P Göertz, Chantal BF Vogels, Corinne Geertsema, Constantianus JM Koenraadt, and Gorben P Pijlman. Mosquito co-infection with zika and chikungunya virus allows simultaneous transmission without affecting vector competence of aedes aegypti. *PLOS Neglected Tropical Diseases*, 11(6):e0005654, 2017.

[118] Marcelo FC Gomes, Ana Pastore y Piontti, Luca Rossi, Dennis Chao, Ira Longini, M Elizabeth Halloran, and Alessandro Vespignani. Assessing the international spreading risk associated with the 2014 west african ebola outbreak. *PLoS currents*, 6, 2014.

[119] Ann-Claire Gourinat, Olivia OConnor, Elodie Calvez, Cyrille Goarant, and Myrielle Dupont-Rouzeyrol. Detection of zika virus in urine. *Emerging infectious diseases*, 21(1):84, 2015.

[120] Rebecca F Grais, J Hugh Ellis, and Gregory E Glass. Assessing the impact of airline travel on the geographic spread of pandemic influenza. *European journal of epidemiology*, 18(11):1065–1072, 2003.

[121] Thilo Gross, Carlos J Dommar DLima, and Bernd Blasius. Epidemic dynamics on an adaptive network. *Physical review letters*, 96(20):208701, 2006.

[122] Richard A Grossman, Robert Edelman, P Chiewanich, P Voodhikul, and C Siriwan. Study of japanese encephalitis virus in chiangmai valley, thailand. ii. human clinical infections. *American journal of epidemiology*, 98(2):121–132, 1973.

[123] Richard A Grossman, Robert Edelman, Milton Willhight, Suntaree Pantuwatana, and Suchinda Udomsakdi. Study of japanese encephalitis virus in chiangmai valley, thailand: Iii. human seroepidemiology and inapparent infections. *American journal of epidemiology*, 98(2):133–149, 1973.

[124] Abba B Gumel, Shigui Ruan, Troy Day, James Watmough, Fred Brauer, P Van den Driessche, Dave Gabrielson, Chris Bowman, Murray E Alexander, Sten Ardal, et al. Modelling strategies for controlling sars outbreaks. *Proceedings of the Royal Society of London B: Biological Sciences*, 271(1554):2223–2232, 2004.

[125] Scott B Halstead, Carl R Grosz, et al. Subclinical japanese encephalitis i. infection of americans with limited residence in korea. *American journal of hygiene*, 75(2):190–201, 1962.

[126] Daihai He, Alice PY Chiu, Qianying Lin, and Benjamin J Cowling. Differences in the seasonality of mers-cov and influenza in the middle east. *International journal of infectious diseases: IJID: official publication of the International Society for Infectious Diseases*, 40:15, 2015.

[127] Daihai He, Alice PY Chiu, Qianying Lin, and Duo Yu. Spatio-temporal patterns of proportions of influenza b cases. *Scientific reports*, 7:40085, 2017.

[128] Daihai He, Jonathan Dushoff, Troy Day, Junling Ma, and David JD Earn. Inferring the causes of the three waves of the 1918 influenza pandemic in england and wales. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1766):20131345, 2013.

[129] Daihai He, Jonathan Dushoff, Raluca Eftimie, and David JD Earn. Patterns of spread of influenza a in canada. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1770):20131174, 2013.

[130] Daihai He and David JD Earn. Epidemiological effects of seasonal oscillations in birth rates. *Theoretical population biology*, 72(2):274–291, 2007.

[131] Daihai He and David JD Earn. The cohort effect in childhood disease dynamics. *Journal of The Royal Society Interface*, 13(120):20160156, 2016.

[132] Daihai He, Daozhou Gao, Yijun Lou, Shi Zhao, and Shigui Ruan. A comparison study of zika virus outbreaks in french polynesia, colombia and the state of bahia in brazil. *Scientific Reports*, 7(1):273, 2017.

[133] Daihai He, Edward L Ionides, and Aaron A King. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *Journal of the Royal Society Interface*, 2009.

[134] Ulrich Heininger and Jane F Seward. Varicella. *The Lancet*, 368(9544):1365–1376, 2006.

[135] T Déirdre Hollingsworth, Neil M Ferguson, and Roy M Anderson. Will travel restrictions control the international spread of pandemic influenza? *Nature medicine*, 12(5):497–499, 2006.

[136] Yan-Jang S Huang, Stephen Higgs, Kate McElroy Horne, and Dana L Vanlandingham. Flavivirus-mosquito interactions. *Viruses*, 6(11):4703–4730, 2014.

[137] Daniel E Impoinvil, Matthew Baylis, and Tom Solomon. Japanese encephalitis: on the one health agenda. In *One Health: The Human-Animal-Environment Interfaces in Emerging Infectious Diseases*, pages 205–247. Springer, 2012.

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

[139] Hector S Izurieta, Peter M Strebel, and Paul A Blake. Postlicensure effectiveness of varicella vaccine during an outbreak in a child care center. *JAMA*, 278(18):1495–1499, 1997.

[140] Charlotte Jackson, Punam Mangtani, Paul Fine, and Emilia Vynnycky. The effects of school holidays on transmission of varicella zoster virus, england and wales, 1967–2008. *PloS one*, 9(6):e99762, 2014.

[141] Michael A Johansson, Neysarí Arana-Vizcarrondo, Brad J Biggerstaff, and J Erin Staples. Incubation periods of yellow fever virus. *The American journal of tropical medicine and hygiene*, 83(1):183–188, 2010.

[142] Michael A Johansson, Pedro FC Vasconcelos, and J Erin Staples. The whole iceberg: estimating the incidence of yellow fever virus infection from the number of severe cases. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, 108(8):482–487, 2014.

[143] Matt J Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, 2008.

[144] William O Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. In *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, volume 115, pages 700–721. The Royal Society, 1927.

[145] Salah Uddin Khan, Henrik Salje, A Hannan, Md Atiqul Islam, AA Mamun Bhuyan, Md Ariful Islam, M Ziaur Rahman, Nazmun Nahar, M Jahangir Hossain, Stephen P Luby, et al. Dynamics of japanese encephalitis virus transmission among pigs in northwest bangladesh and the potential impact of pig vaccination. *PLoS neglected tropical diseases*, 8(9):e3166, 2014.

[146] Aaron A King, Edward L Ionides, CM Bretó, Steve Ellner, Bruce Kendall, Helen Wearing, Matthew J Ferrari, Michael Lavine, and Daniel C Reuman. pomp: Statistical inference for partially observed markov processes (r package). *URL http://pomp. r-forge. r-rproject. org*, 2010.

[147] Aaron A King, Dao Nguyen, and Edward L Ionides. Statistical inference for partially observed markov processes via the r package pomp. *arXiv preprint arXiv:1509.00503*, 2015.

[148] K Kodama, N Sasaki, and Y Kanda Inoue. Studies of live attenuated japanese encephalitis vaccine in swine. *The Journal of Immunology*, 100(1):194–200, 1968.

[149] Eiji Konishi and Tomoyuki Suzuki. Ratios of subclinical to clinical japanese encephalitis (je) virus infections in vaccinated populations: evaluation of an inactivated je vaccine by comparing the ratios with those in unvaccinated populations. *Vaccine*, 21(1):98–107, 2002.

[150] Jiro Konno. Cyclic outbreak of japanese encephalitis among pigs and humans. *Japanese Journal of Tropical Medicine*, 10(2):132–133, 1969.

[151] Moritz UG Kraemer, Nuno R Faria, Robert C Reiner, Nick Golding, Birgit Nikolay, Stephanie Stasse, Michael A Johansson, Henrik Salje, Ousmane Faye, GR William Wint, et al. Spread of yellow fever virus outbreak in angola and the democratic republic of the congo 2015–16: a modelling study. *The Lancet Infectious Diseases*, 17(3):330–338, 2017.

[152] Adam J Kucharski, Sebastian Funk, Rosalind M Eggo, Henri-Pierre Mallet, W John Edmunds, and Eric J Nilles. Transmission dynamics of zika virus in island populations: a modelling analysis of the 2013–14 french polynesia outbreak. *PLoS neglected tropical diseases*, 10(5):e0004726, 2016.

[153] R Kumar, A Mathur, A Kumar, S Sharma, S Chakraborty, and UC Chaturvedi. Clinical features & prognostic indicators of japanese encephalitis in children in lucknow (india). *The Indian journal of medical research*, 91:321–327, 1990.

[154] Kevin D Lafferty. The ecology of climate change and infectious diseases. *Ecology*, 90(4):888–900, 2009.

[155] K Lam, OT Tsang, RW Yung, and KK Lau. Japanese encephalitis in hong kong. *Hong Kong Med J*, 11(3):182–8, 2005.

[156] Gaëtan Lebrun, Karim Chadda, Anne-Hélène Reboux, Olivier Martinet, and Bernard-Alex Gaüzère. Guillain-barré syndrome after chikungunya infection. *Emerging infectious diseases*, 15(3):495, 2009.

[157] Lore E Lee, Hon Ho, Eileen Lorber, Jeanne Fratto, Sara Perkins, and Paul R Cieslak. Vaccine-era varicella epidemiology and vaccine effectiveness in a public elementary school population, 2002–2007. *Pediatrics*, 121(6):e1548–e1554, 2008.

[158] X Lenne, J Diez Domingo, A Gil, M Ridao, JA Lluch, and B Dervaux. Economic evaluation of varicella vaccination in spainxresults from a dynamic model. *Vaccine*, 24(47):6980–6989, 2006.

[159] Justin Lessler, Nicholas G Reich, Derek AT Cummings, New York City Department of Health, and Mental Hygiene Swine Influenza Investigation Team. Outbreak of 2009 pandemic influenza a (h1n1) at a new york city school. *New England Journal of Medicine*, 361(27):2628–2636, 2009.

[160] Ming-Tao Li, Gui-Quan Sun, Yan-Fang Wu, Juan Zhang, and Zhen Jin. Transmission dynamics of a multi-group brucellosis model with mixed cross infection in public farm. *Applied Mathematics and Computation*, 237:582–594, 2014.

[161] Daniel H Libraty, Ananda Nisalak, Timothy P Endy, Saroj Suntayakorn, David W Vaughn, and Bruce L Innis. Clinical and immunological risk factors for severe disease in japanese encephalitis. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 96(2):173–178, 2002.

[162] Qianying Lin, Zhigui Lin, Alice PY Chiu, and Daihai He. Seasonality of influenza a (h7n9) virus in chinaxfitting simple epidemic models to human cases. *PloS one*, 11(3):e0151333, 2016.

[163] Marc Lipsitch, Ted Cohen, Ben Cooper, James M Robins, Stefan Ma, Lyn James, Gowri Gopalakrishna, Suok Kai Chew, Chorh Chuan Tan, Matthew H Samore, et al. Transmission dynamics and control of severe acute respiratory syndrome. *Science*, 300(5627):1966–1970, 2003.

[164] Rongsong Liu, Jianhong Wu, and Huaiping Zhu. Media/psychological impact on multiple outbreaks of emerging infectious diseases. *Computational and Mathematical Methods in Medicine*, 8(3):153–164, 2007.

[165] John S Mackenzie, Duane J Gubler, and Lyle R Petersen. Emerging flaviviruses: the spread and resurgence of japanese encephalitis, west nile and dengue viruses. *Nature medicine*, 10(12s):S98, 2004.

[166] Maimuna S Majumder, Mauricio Santillana, Sumiko R Mekaru, Denise P McGinnis, Kamran Khan, and John S Brownstein. Utilizing nontraditional data sources for near real-time estimation of transmission dynamics during the 2015-2016 colombian zika virus disease outbreak. *JMIR public health and surveillance*, 2(1), 2016.

[167] Mona Marin, Melanie Marti, Anita Kambhampati, Stanley M Jeram, and Jane F Seward. Global varicella vaccine effectiveness: a meta-analysis. *Pediatrics*, pages peds–2015, 2016.

[168] Eduardo Massad, Marcelo Nascimento Burattini, Francisco Antonio Bezerra Coutinho, and Luiz Fernandes Lopez. Dengue and the risk of urban yellow fever reintroduction in sao paulo state, brazil. *Revista de Saúde Pública*, 37(4):477–484, 2003.

[169] Martial L Ndeffo Mbah, Jingzhou Liu, Chris T Bauch, Yonas I Tekel, Jan Medlock, Lauren Ancel Meyers, and Alison P Galvani. The impact of imitation on vaccination behavior in social contact networks. *PLoS computational biology*, 8(4):e1002469, 2012.

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

[171] George J Milne, Joel K Kelso, Heath A Kelly, Simon T Huband, and Jodie McVernon. A small community model for the transmission of infectious diseases: comparison of school closure as an intervention in individual-based models of an influenza pandemic. *PloS one*, 3(12):e4005, 2008.

[172] Thomas P Monath. Treatment of yellow fever. *Antiviral research*, 78(1):116–124, 2008.

[173] Thomas P Monath and Abdulsalami Nasidi. Should yellow fever vaccine be included in the expanded program of immunization in africa? a cost-effectiveness analysis for nigeria. *The American journal of tropical medicine and hygiene*, 48(2):274–299, 1993.

[174] Thomas P. Monath and Pedro F.C. Vasconcelos. Yellow fever. *Journal of Clinical Virology*, 64(Supplement C):160 – 173, 2015.

[175] TP Monath, RB Craven, A Adjukiewicz, Max Germain, DB Francy, Léo Ferrara, E Mm Samba, H N'jie, K Cham, SA Fitzgerald, et al. Yellow fever in the gambia, 1978–1979: epidemiologic aspects with observations on the occurrence of orungo virus infections. *The American journal of tropical medicine and hygiene*, 29(5):912–928, 1980.

[176] TP Monath, I Levenbook, K Soike, Z-X Zhang, M Ratterree, K Draper, ADT Barrett, R Nichols, R Weltzin, J Arroyo, et al. Chimeric yellow fever virus 17d-japanese encephalitis virus vaccine: dose-response effectiveness and extended safety testing in rhesus monkeys. *Journal of virology*, 74(4):1742–1751, 2000.

[177] TP Monath, DC Wilson, VH Lee, G Stroh, K Kuteyi, and EA Smith. The 1970 yellow fever epidemic in okwoga district, benue plateau state, nigeria: 1. epide miological observations. *Bulletin of the World Health Organization*, 49(2):113, 1973.

[178] Rosa Moreira. Reining in angola's yellow fever outbreak. *Bulletin of the World Health Organization*, 94:716–717, 2016.

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

[180] Didier Musso, Claudine Roche, Emilie Robin, Tuxuan Nhan, Anita Teissier, and Van-Mai Cao-Lormeau. Potential sexual transmission of zika virus. *Emerging infectious diseases*, 21(2):359, 2015.

[181] Ingemar Nasell. A new look at the critical community size for childhood infections. *Theoretical population biology*, 67(3):203–216, 2005.

[182] A Nasidi, TP Monath, K DeCock, O Tomori, R Cordellier, OD Olaleye, TO Harry, JA Adeniyi, AO Sorungbe, AO Ajose-Coker, et al. Urban yellow fever epidemic in western nigeria, 1987. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 83(3):401–406, 1989.

[183] J.v. Neumann, Oskar Morgenstern, and HW Kuhn. *Theory of games and economic behavior*. Princeton University Press, Commemorative Ed edition (1 May 2007), 1953.

[184] Hiroshi Nishiura and Gerardo Chowell. Early transmission dynamics of ebola virus disease (evd), west africa, march to august 2014. 2004.

[185] Hiroshi Nishiura, Kenji Mizumoto, Kat S Rock, Yohei Yasuda, Ryo Kinoshita, and Yuichiro Miyamatsu. A theoretical estimate of the risk of microcephaly during pregnancy with zika virus infection. *Epidemics*, 15:66–70, 2016.

[186] Erwan Oehler, Emmanuel Fournier, Isabelle Leparc-Goffart, Philippe Larre, Stéphanie Cubizolle, Chantal Sookhareea, Stéphane Lastère, and Frédéric Ghawche. Increase in cases of guillain-barré syndrome during a chikungunya outbreak, french polynesia, 2014 to 2015. *Euro Surveill*, 20(48):30079, 2015.

[187] Benson Ogunjimi, Niel Hens, Nele Goeyvaerts, Marc Aerts, Pierre Van Damme, and Philippe Beutels. Using empirical social contact data to model person to person infectious disease transmission: an illustration for varicella. *Mathematical biosciences*, 218(2):80–87, 2009.

[188] World Health Organization et al. Fourth meeting of the emergency committee under the international health regulations (2005) regarding microcephaly, other neurological disorders and zika virus. *Saudi medical journal*, 37(10):1166–1167, 2016.

[189] Xiao-Ling Pan, Hong Liu, Huan-Yu Wang, Shi-Hong Fu, Hai-Zhou Liu, Hai-Lin Zhang, Ming-Hua Li, Xiao-Yan Gao, Jing-Lin Wang, Xiao-Hong Sun, et al. Emergence of genotype i of japanese encephalitis virus as the dominant genotype in asia. *Journal of virology*, 85(19):9847–9853, 2011.

[190] Xinghuo Pang, Zonghan Zhu, Fujie Xu, Jiyong Guo, Xiaohong Gong, Donglei Liu, Zejun Liu, Daniel P Chin, and Daniel R Feikin. Evaluation of control measures implemented in the severe acute respiratory syndrome outbreak in beijing, 2003. *Jama*, 290(24):3215–3221, 2003.

[191] Ana Perisic and Chris T Bauch. Social contact networks and disease eradicability under voluntary vaccination. *PLoS computational biology*, 5(2):e1000280, 2009.

[192] Piero Poletti, Marco Ajelli, and Stefano Merler. The effect of risk perception on the 2009 h1n1 pandemic influenza dynamics. *PloS one*, 6(2):e16460, 2011.

[193] Piero Poletti, Alessia Melegaro, Marco Ajelli, Emanuele del Fava, Giorgio Guzzetta, Luca Faustini, Giampaolo Scalia Tomba, Pierluigi Lopalco, Caterina Rizzo, Stefano Merler, et al. Perspectives on the impact of varicella immunization on herpes zoster. a model-based evaluation from three european countries. *PloS one*, 8(4):e60732, 2013.

[194] Paul Reiter. Climate change and mosquito-borne disease. *Environmental health perspectives*, 109(Suppl 1):141, 2001.

[195] Timothy C Reluga. An sis epidemiology game with two subpopulations. *Journal of Biological Dynamics*, 3(5):515–531, 2009.

[196] Timothy C Reluga, Chris T Bauch, and Alison P Galvani. Evolving public perceptions and stability in vaccine uptake. *Mathematical biosciences*, 204(2):185–198, 2006.

[197] Meret E Ricklin, Obdulio García-Nicolás, Daniel Brechbühl, Sylvie Python, Beatrice Zumkehr, Antoine Nougairede, Remi N Charrel, Horst Posthaus, Anna Oevermann, and Artur Summerfield. Vector-free transmission and persistence of japanese encephalitis virus in pigs. *Nature communications*, 7, 2016.

[198] Meret E Ricklin, Obdulio Garcìa-Nicolàs, Daniel Brechbühl, Sylvie Python, Beatrice Zumkehr, Horst Posthaus, Anna Oevermann, and Artur Summerfield. Japanese encephalitis virus tropism in experimentally infected pigs. *Veterinary research*, 47(1):34, 2016.

[199] S Riley, GM Leung, LM Ho, and BJ Cowling. Transmission of japanese encephalitis virus in hong kong. *Hong Kong medical journal = Xianggang yi xue za zhi*, 18 Suppl 2:4546, February 2012.

[200] Steven Riley, Christophe Fraser, Christl A Donnelly, Azra C Ghani, Laith J Abu-Raddad, Anthony J Hedley, Gabriel M Leung, Lai-Ming Ho, Tai-Hing Lam, Thuan Q Thach, et al. Transmission dynamics of the etiological agent of sars in hong kong: impact of public health interventions. *Science*, 300(5627):1961–1966, 2003.

[201] Steven Riley, Gabriel Matthew Leung, Lai-Ming Ho, and Benjamin John Cowling. Identification of the key elements of transmission of japanese encephalitis virus in hong kong.

[202] Leon Rosen, Jih-Ching Lien, Donald A Shroyer, Richard H Baker, and Liang-Chen Lu. Experimental vertical transmission of japanese encephalitis virus by culex tritaeniorhynchus and other mosquitoes. *The American journal of tropical medicine and hygiene*, 40(5):548–556, 1989.

[203] Leon Rosen, Robert B Tesh, Jih Ching Lien, and John H Cross. Transovarial transmission of japanese encephalitis virus by mosquitoes. *Science*, 199(4331):909–911, 1978.

[204] S Ruan. The effect of global travel on the spread of sars. *Mathematical biosciences and engineering: MBE*, 6(1):207, 2009.

[205] Gideon Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.

[206] Leah B Shaw and Ira B Schwartz. Fluctuating epidemics on adaptive networks. *Physical Review E*, 77(6):066101, 2008.

[207] Eunha Shim, Gretchen B Chapman, Jeffrey P Townsend, and Alison P Galvani. The influence of altruism on influenza vaccination decisions. *Journal of The Royal Society Interface*, page rsif20120115, 2012.

[208] T Solomon and PM Winter. Neurovirulence and host factors in flavivirus encephalitisxevidence from clinical epidemiology. In *Emergence and control of zoonotic viral encephalitides*, pages 161–170. Springer, 2004.

[209] Kathleen Stratton, Alicia Gable, Padma Shetty, Marie McCormick, Institute of Medicine (US) Immunization Safety Review Committee, et al. Immunization safety review: measles-mumps-rubella vaccine and autism. 2001.

[210] Gui-Quan Sun, Jun-Hui Xie, Sheng-He Huang, Zhen Jin, Ming-Tao Li, and Liqun Liu. Transmission dynamics of cholera: Mathematical modeling and control strategies. *Communications in Nonlinear Science and Numerical Simulation*, 45:235–244, 2017.

[211] Gui-Quan Sun and Zi-Ke Zhang. Global stability for a sheep brucellosis model with immigration. *Applied Mathematics and Computation*, 246:336–345, 2014.

[212] Jun-Sang Sunwoo, Keun-Hwa Jung, Soon-Tae Lee, Sang Kun Lee, and Kon Chu. Reemergence of japanese encephalitis in south korea, 2010–2015. *Emerging infectious diseases*, 22(10):1841, 2016.

[213] Åke Svensson. A note on generation times in epidemic models. *Mathematical biosciences*, 208(1):300–311, 2007.

[214] Ikuo Takashima and Leon Rosen. Horizontal and vertical transmission of japanese encephalitis virus by aedes japonicus (diptera: Culicidae). *Journal of medical entomology*, 26(5):454–458, 1989.

[215] Ratree Takhampunya, Heung-Chul Kim, Bousaraporn Tippayachai, Ampornpan Kengluecha, Terry A Klein, Won-Ja Lee, John Grieco, and Brian P Evans. Emergence of japanese encephalitis virus genotype v in the republic of korea. *Virology journal*, 8(1):449, 2011.

[216] Jean M Tchuenche, Nothabo Dube, Claver P Bhunu, Robert J Smith, and Chris T Bauch. The impact of media coverage on the transmission dynamics of human influenza. *BMC Public Health*, 11(1):S5, 2011.

[217] Huai-Yu Tian, Peng Bi, Bernard Cazelles, Sen Zhou, Shan-Qian Huang, Jing Yang, Yao Pei, Xiao-Xu Wu, Shi-Hong Fu, Shi-Lu Tong, et al. How environmental conditions impact mosquito ecology and japanese encephalitis: an eco-epidemiological approach. *Environment international*, 79:17–24, 2015.

[218] Joseph H Tien, Hendrik N Poinar, David N Fisman, and David JD Earn. Herald waves of cholera in nineteenth century london. *Journal of The Royal Society Interface*, page rsif20100494, 2010.

[219] Oyewale Tomori. Yellow fever: the recurring plague. *Critical reviews in clinical laboratory sciences*, 41(4):391–427, 2004.

[220] Sherry Towers, Oscar Patterson-Lomba, and Carlos Castillo-Chavez. Temporal variations in the effective reproduction number of the 2014 west africa ebola outbreak. *PLoS currents*, 6, 2014.

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

[222] Andrew F Van den Hurk, Scott A Ritchie, and John S Mackenzie. Ecology and geographical expansion of japanese encephalitis virus. *Annual review of entomology*, 54:17–35, 2009.

[223] Albert Jan van Hoek, Alessia Melegaro, Emelio Zagheni, W John Edmunds, and Nigel Gay. Modelling the impact of a combined varicella and zoster vaccination programme on the epidemiology of varicella zoster virus in england. *Vaccine*, 29(13):2411–2420, 2011.

[224] Pedro FC Vasconcelos, Sueli G Rodrigues, Nicolas Degallier, Mario AP Moraes, Jorge FS Travassos Da Rosa, Elizabeth S Travassos Da Rosa, Bernard Mondet, Vera LRS Barros, and Amelia PA Travassos Da Rosa. An epidemic of sylvatic yellow fever in the southeast region of maranhao state, brazil, 1993–1994: epidemiologic and entomologic findings. *The American journal of tropical medicine and hygiene*, 57(2):132–137, 1997.

[225] Cesar Gomes Victora, Lavinia Schuler-Faccini, Alicia Matijasevich, Erlane Ribeiro, André Pessoa, and Fernando Celso Barros. Microcephaly in brazil: how to interpret reported numbers? *The Lancet*, 387(10019):621–624, 2016.

[226] Wilmer Villamil-Gómez, Luz Alba Silvera, Jorge Páez-Castellanos, and Alfonso J Rodriguez-Morales. Guillain-barré syndrome after chikungunya infection: A case in colombia. *Enfermedades infecciosas y microbiologia clinica*, 34(2):140–141, 2016.

[227] Jacco Wallinga, Peter Teunis, and Mirjam Kretzschmar. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *American journal of epidemiology*, 164(10):936–944, 2006.

[228] Lulan Wang, Stephanie G Valderramos, Aiping Wu, Songying Ouyang, Chunfeng Li, Patricia Brasil, Myrna Bonaldo, Thomas Coates, Karin Nielsen-Saines, Taijiao Jiang, et al. From mosquitos to humans: genetic evolution of zika virus. *Cell host & microbe*, 19(5):561–565, 2016.

[229] Wendi Wang and G Mulone. Threshold of disease transmission in a patch environment. *Journal of Mathematical Analysis and Applications*, 285(1):321–335, 2003.

[230] Wendi Wang and Shigui Ruan. Simulating the sars outbreak in beijing with limited data. *Journal of theoretical biology*, 227(3):369–379, 2004.

[231] Wendi Wang and Xiao-Qiang Zhao. An epidemic model in a patchy environment. *Mathematical biosciences*, 190(1):97–112, 2004.

[232] Zhen Wang, Michael A Andrews, Zhi-Xi Wu, Lin Wang, and Chris T Bauch. Coupled disease–behavior dynamics on complex networks: A review. *Physics of life reviews*, 15:1–29, 2015.

[233] Zhen Wang, Chris T Bauch, Samit Bhattacharyya, Alberto d'Onofrio, Piero Manfredi, Matjaž Perc, Nicola Perra, Marcel Salathé, and Dawei Zhao. Statistical physics of vaccination. *Physics Reports*, 664:1–113, 2016.

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

[235] AC Wielanek, J De Monredon, M El Amrani, JC Roger, and JP Serveaux. Guillain-barré syndrome complicating a chikungunya virus infection. *Neurology*, 69(22):2105–2107, 2007.

[236] VB Wigglesworth. The adaptation of mosquito larvae to salt water. *Journal of Experimental biology*, 10(1):27–36, 1933.

[237] David T Williams, Peter W Daniels, Ross A Lunt, Lin-Fa Wang, Kim M Newberry, and John S Mackenzie. Experimental infections of pigs with japanese encephalitis virus and closely related australian flaviviruses. *The American journal of tropical medicine and hygiene*, 65(4):379–387, 2001.

[238] JP Woodall and TM Yuill. Why is the yellow fever outbreak in angola a 'threat to the entire world? *International Journal of Infectious Diseases*, 48:96–97, 2016.

[239] Jianyong Wu, Radhika Dhingra, Manoj Gambhir, and Justin V Remais. Sensitivity analysis of infectious disease models: methods, advances and their application. *Journal of The Royal Society Interface*, 10(86):20121018, 2013.

[240] Joseph T Wu, Corey M Peak, Gabriel M Leung, and Marc Lipsitch. Fractional dosing of yellow fever vaccine to extend supply: a modelling study. *The Lancet*, 388(10062):2904–2911, 2016.

[241] Zhi-Qiang Xia, Shi-Fu Wang, Shen-Long Li, Liu-Yu Huang, Wen-Yi Zhang, Gui-Quan Sun, Zhong-Tao Gai, and Zhen Jin. Modeling the transmission dynamics of ebola virus disease in liberia. *Scientific reports*, 5, 2015.

[242] Duo Yu, Qianying Lin, Alice PY Chiu, and Daihai He. Effects of reactive social distancing on the 1918 influenza pandemic. *PLoS ONE*, 12(7):1–16, 07 2017.

[243] Emilio Zagheni, Francesco C Billari, Piero Manfredi, Alessia Melegaro, Joel Mossong, and W John Edmunds. Using time-use data to parameterize models for the spread of close-contact infectious diseases. *American journal of epidemiology*, 168(9):1082–1090, 2008.

[244] Yi Zeng and Therese Hesketh. The effects of china's universal two-child policy. *The Lancet*, 388(10054):1930–1938, 2016.

[245] Jiu-Song Zhang, Qiu-Min Zhao, Xiao-Fang Guo, Shu-Qing Zuo, Jing-Xia Cheng, Na Jia, Chao Wu, Pei-Fang Dai, and Jun-Ying Zhao. Isolation and genetic characteristics of human genotype 1 japanese encephalitis virus, china, 2009. *PLoS One*, 6(1):e16418, 2011.