Epidemic Modeling for Travel Restrictions on the Pandemic Influenza A (H1N1)

CHONG, Ka Chun

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy
in
Public Health

The Chinese University of Hong Kong
September 2011

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Thesis/Assessment Committee

Professor WONG Yeung Shan Samuel (Chair)
Professor ZEE Chung Ying Benny (Thesis Supervisor)
Professor GOGGINS III William Bernard (Committee Member)
Professor YAU Kai Wing Kelvin (External Examiner)

Abstract

In an epidemic, the international traffic accelerates the spread of infections across wide geographic areas. Policy makers are interested to know the impact on the disease transmission once the international traffic has been re-scaled. Since the pharmaceutical interventions are usually not ready in the early stage of a new epidemic, the travel restriction is a high potential intervention that should be included into the containment and mitigation strategies for officials. According to some researches, the value of the travel restriction was controversial; and most importantly, we discovered several practical and theoretical limitations in the epidemic models. These problems largely motivated us to study the effectiveness of the travel restriction on the epidemic control in both *at-risk* countries and the *source* country. In the body of thesis, new methodologies of epidemic modeling were developed by making use of the influenza A (H1N1) pandemic in 2009 as a case study. Our result showed that the travel restriction was valuable on slowing down the growth of epidemics for both *at-risk* countries and the *source* country. The time delay of the epidemic would offer public health experts, policy makers, and scientists more time for preparation and decision making on control measures. Although solely imposing the travel restriction showed little benefit on reducing the final attack rate and the probability of cases exportation, it offered additional contribution on even halting the epidemic growth once other interventions such as antiviral and hospitalization could also be implemented. Therefore, the implementation of the travel restriction must be a potential intervention to control the

epidemic spread, especially for the next epidemics which could be lethal and highly intrusive.

摘要

在一個流行病疫情中，國際性交通加速了感染個案在廣泛的地理區域中傳播。政策制定者都會有興趣知道一旦國際交通比例重新調整後，對疾病傳播有多影響。由於抗病毒藥物通常都不能夠在一個新疫情的早期階段中準備好，旅遊限制絕對是一個具相當潛力的措施，而且應納入官方的疫情遏制和緩解計劃。根據一些研究，旅遊限制的價值是相當爭議性的，更最重要的，我們發現了幾個在傳染病模型中的實際和理論局限。這些問題大大地促使我們研究旅遊限制在存在風險國家和病源國的疫情控制成效。在這篇論文中，我們建購了新的流行病模型理論，並透過利用二零零九年的甲型流感(H1N1)大流行作爲個案研究。我們的結果顯示，旅遊限制爲減慢存在風險國家和病源國的疫情增長帶來一定價值。時間的延遲將會爲公共衛生專家，決策者和科學家提供更多空間去準備和制定控制疫情的策略。雖然單獨實施旅遊限制只能減少一小部分最終發病率及病源輸出的可能性，但一旦其他疫情控制措施例如抗病毒藥物及住院也可以同時間提供，旅遊限制絕對能夠提供額外的貢獻, 甚至停止疫情的增長。因此，旅遊限制絕對是一個具相當潛在能力的疫情控制措施，特別是對於未來的疫病大可能高侵襲性及致命。

Acknowledgments

I am thankful to my PhD advisor, Professor Benny Zee for supporting me on the PhD works. Professor Zee has been very patient and encouraging throughout my study. It is my great pleasure to have his guidance which is not only academic teaching but also experience sharing and personality improvement.

I would like to thank Maria and Professor Cheung Siu Hung for encouraging me to engage in the study of the PhD degree.

I must also thank Dr. David Wilmshurst and Miss Yip Siu Ying for helpful comments on editing.

Many people made my life enjoyable. I am lucky that I met many CCT and BMT buddies: Dawn, Namita, Chris, Lunan, Lewis, Jack, Philip, Chloe, Saren, Raymond, Eva, Peggy, Dolores, Jenny, Winnie, Terry, Josephine, Professor Goggins, Fred, Miranda, Carmen, and particularly Maisie. I am also happy to meet the SPIIPC schoolmates: Liqing, Hou Jing, Li Xue, Qin Ying, Cecilia, Minghui, Leon, Kenneth, Cai Yue, Hildy, Liz, Grace, Fu Zhenming, Yu Wenzhou, Xiaobing, Hao Chun, Joey, Daisy, Wang, David, Harry, Nicole, Vivi, Dorothy, Vivien, and my adminstrative supervisor Joyce.

In particular, I greatly appreciate encouragement and support of my family and my wife for the completion of my thesis.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Statement of the Problem

The novel influenza A (H1N1), swine flu has spreaded across more than 70 countries and the speed of disease transmission raised the public awareness on control measure globally. When the epidemic outbreak was started in the *source* country, Mexico, containing the epidemic was one of the priority actions. Unfortunately, the containment strategy was fail and the mass international travel pattern lead infected individuals to carry their virus to other *at-risk* countries. The prior action of *at-risk* countries is to delay the epidemic outbreak locally in order to squeeze more time to preparation and to minimize the attack rate afterward. Therefore, the objectives of imposing the control measures especially the travel restriction, between the *source* country and the *at-risk* countries are different. Recently, several literatures have developed different epidemic models to evaluate the effectiveness of various interventions. Among those interventions, the travel restriction gained most interest. Since there are several limitations for these epidemic models such as inadequate distribution of travel pattern and lack of uncertainty estimates, the value of the travel restriction has not yet been proven.

In this paper, methodologies will be newly developed, by making use of the influenza A (H1N1) pandemic in 2009 as a case study to demonstrate the

impact of travel restriction in views of the *source* country and the *at-risk* countries.

1.2 Importance of Epidemic Models

In reality, it is hard to have high standard experimental or observational studies to analyze the strategies on controlling the transmissions of infectious diseases. Most clinical trial designs are not practical for assessing the effectiveness of some interventions, such as face masks and isolation, because of the ethical considerations which relate to epidemics in general. Therefore, policy makers are hard to understand the effectiveness of their strategies. Fortunately, by using the mathematical models, the dynamic of the epidemic and the impact of the interventions can be demonstrated. The models are able to use mathematical science to describe the disease system under the constraints of interventions from the biological, political, and epidemiological data. The modeling result could thus explain and quantify how infectious diseases spread in the real world. So mathematical models play essential roles in offering valuable advice among different political settings for epidemics. Because of the above advantages, there is an increasing trend in applying mathematical models to explore the spread patterns of infectious diseases and the impacts of interventions in recent decades.

1.3 Applications of Epidemic Modeling

Epidemic models have been used to design optimum strategies in containing the disease at a source country, assignment of treatments or vaccination, preparation of the interventions, and controlling the antiviral resistance in epidemics over the past decades. Longini, *et al.* [75] estimated the disease transmission pattern in southeast Asia and assessed the possibility for containing the H5N1

influenza epidemic. Reily, *et al.* [97] designed the optimum dose coverage in order to maintain a lower illness attack rate in the United States. In order to prepare effective interventions, Ferguson, *et al.* [39] employed a stochastic simulation model to study the effectiveness of various control measures such as geographical treatments allocation and workplace, household or hospital quarantine. Besides, Gani, *et al.* [49] estimated the ratio of the antiviral resistance in view of different antiviral coverage situations during an influenza pandemic. Apart from that, epidemic models have also been used to predict the threats of bioterrorism from emerging virus.

1.4 Pandemic Influenza A (H1N1)

Novel influenza A (H1N1), also called swine flu, is a new influenza virus that caused its first illness in Mexico in 2009. In mid-march 2009, Mexico government identified an unreasonable increase in the number of influenza-like illness cases, even though it was not in peak seasons of the influenza outbreak [77]. After half a month, an acute respiratory illness was discovered on two children and was further confirmed as a new influenza A (H1N1) virus in mid-April 2009 [78]. Then the first notification of novel influenza A (H1N1) was announced by the World Health Organisation (WHO) on April 26, 2009. Because additional cases were successively discovered in the United States [79], WHO raised the pandemic alert level to phase five in the end of April. Because of insufficient information on this particular infectious disease, the World Health Organization (WHO) declared the first global influenza pandemic on June 11, 2009. In a recent clinical update, more than 214 countries and territories worldwide have reported laboratory-confirmed influenza A (H1N1) cases, and the disease has caused more than 18,000 deaths [116]. Its high transmissibility has raised the public awareness of disease control measures.

With Hong Kong's large-scale international travel pattern and a high population density, the Centre for Health Protection (CHP), Hong Kong has reported about 300 severe cases and 80 fatal cases from influenza A H1N1 in Hong Kong since May 1, 2009 [22]. The virus has been widely circulating locally, and it is therefore necessary to implement effective control measures in order to relieve the disease burden. According to CHP, the control measures for mitigation was adjusted and taken effect in Hong Kong after the disease outbreak [90]. Up to eight Designated Flu Clinics (DFCs) were implemented for managing patients with fever and influenza-like illnesses. Confirmed cases with mild symptoms were not required for admissions and were provided with symptomatic treatments and reassessments. Antiviral treatments were only given to influenza like illness patients with chronic diseases or in immuno-compromised states. Hospitalization was target for clinically more serious cases, confirmed cases from pregnancy, and cases presenting medical risk factors, which include those suffering from chronic diseases or having immuno-compromised states.

In the pandemic, the international traffic accelerated the spread of infections across wide geographic scales. The researchers, even publics, would like to know the impact on the influenza A (H1N1) disease transmission once the traffic has been either partially or completely blocked. Moreover, the value of the travel restriction is still not clear now, especially when the pharmaceutical interventions are not ready in the early stage of the pandemic. Therefore, the pandemic influenza A (H1N1) would be a good case study to explore the impact of interventions in both *source* country and *at-risk* countries, in order to have well planning of containment and mitigation strategies in the future.

1.5 Dissertation Outline

Chapter 2 of the dissertation is the literature review of the current mathematical models, the statistical methods, and the applications in epidemics. The research questions will also be identified in this chapter. Chapter 3 demonstrates the impact of travel restriction for *at-risk* countries with the corresponding methodology, the result, and the discussion which employing the pandemic influenza A (H1N1) as a case study. The impact of travel restriction for a *source* country is analyzed in Chapter 4; the corresponding methodology, the result, and the discussion will also be noted which used the pandemic influenza A (H1N1) as a case study. Chapter 5 is the summary and the conclusion from the findings.

Chapter 2

Literature Review

2.1 Epidemic Models

Mathematical modelling has been used for transmission mechanism of infectious disease for a long time. Hamer [57] has developed one of the earliest epidemic models in 1906. The model considered the probability of infection in one time step proportional to the product of the number of susceptible individuals and the number of infected individuals. Ross [99] adopted the method in a time series model and called it mass action principle in 1916. Until 1927, Kermack and McKendrick [66] developed a famous SIR model and it still works as a principal for various extensions of epidemic models for nowadays. The details of Kermack and McKendrick SIR model will be discussed in Section 2.1.2.

In an epidemic model design, it is necessary to know the requirement of identification of the questions and decide how much detail should be incorporated. For example, to evaluate the travel restriction, it is better to put the travel pattern into the model. However, for model with greater detail, the parameter values should be set carefully, otherwise bias will be introduced. On the other hand, some model structures are sensitive to parameters so the evaluation of the sensitivity is important.

Generally, the epidemic models are in compartmental form, that is, the

population is divided into different compartments or categories according to individual disease status, demographic details, and risk factors, so that, the compartmental epidemic models can have better resolution in disease transmission dynamic. However, the model complexity increase with the number of compartments as well as the available data; the statistical inference is hard to be drawn for complicated structural model especially for those with many latent variables.

A large amount of frameworks of infectious disease transmission model are from deterministic and stochastic structures. Due to the complexity of epidemic models, the deterministic structures are relatively easy to build and thus common in practice in order to demonstrate the average behavior of infectious disease transmission. But the main problem is that deterministic models do not capture any uncertainty along with the time series of disease propagation. On the other hand, the stochastic structures are often individual-based and incorporate the stochastic variation into the epidemic system, especially when the number of infected individuals is small or the chance event is important in the transmission dynamics [13]. For example, the epidemic can go to extinct provided that the number of initial infected subjects is small. In addition, the statistical inference for epidemiological parameters is more appropriate to be drawn in stochastic models. Typically, binomial chain method is adopted in discrete time stochastic models such as the Reed-Frost model [2] and the Greenwood model [56]. The stochastic epidemic models have been used extensively on infectious disease like foot and mouth disease [108], meningococcal disease [93], and Human immunodeficiency virus (HIV) [64]. As for a whole picture, the deterministic models are usually used as explanatory tools in describing the general picture of the epidemic and to estimate the transmission parameters followed-by the complex models.

2.1.1 Definitions of Epidemiological Quantities

Understanding the quantity of the disease mechanism is crucial to effective pandemic preparation. Basic reproductive numbers (R_0) is defined as the average number of secondary infections produced by a typical infected individual in a wholly susceptible population. It is usually difficult to measure as not all people in the population are susceptible due to the pre-existing immunity especially for influenza. So the reproduction number (or effective reproduction number), that is, the average number of secondary infections produced by a typical infected individual and denoted as R, is estimated. In order to prevent the pandemic, the quantity R which identifying the intensity of interventions used, should be maintained smaller than 1. The more control measures and interventions should be introduced if the quantity R is large. Some estimated basic reproduction numbers are listed in Table 2.1.

Table 2.1: Examples of the basic reproduction numbers (R_0) according to different infectious diseases

Infectious Disease	**Estimated R_0**	**Reference**
Measles	16-18	[6]
HIV	2-5	[5]
Foot-and-Mouth disease	3.5-4.5	[41]
Smallpox	3.5-6	[50]
SARS	2-5	[111]
Influenza	1-3	[82]

According to World Health Organization (WHO) [114], a simple figure for strategy from range of reproduction number and case fatality rate is presented (Figure 2.1). The higher the value of the reproduction number, the more complex to contain the disease. For the lower value of R, government bodies should consider the cost for containment and benefits of mitigations.

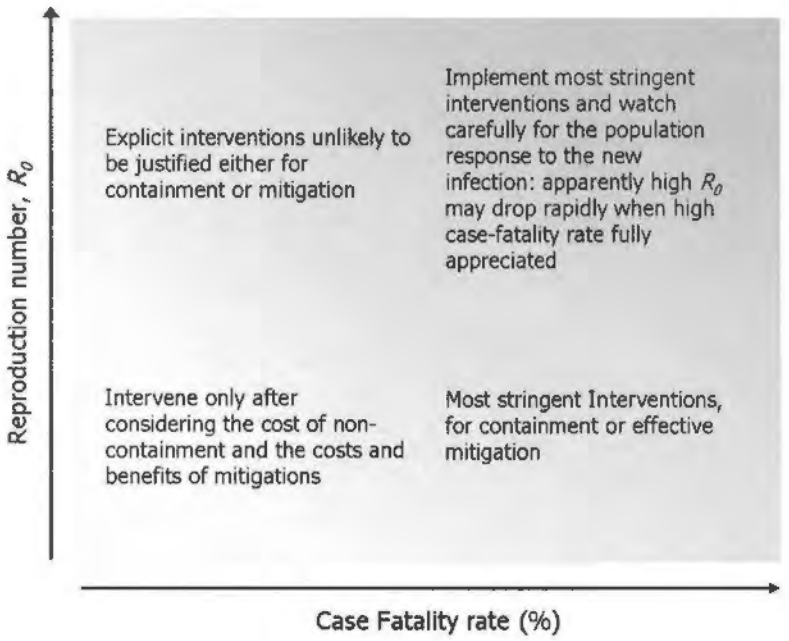


Figure 2.1: Strategies from range of reproduction number and case fatality rate, WHO

Generation time, or serial interval (T), is the average time between the onset of symptoms in a given infected individual and the onset of symptoms in individuals that person has infected. When the serial interval is large, health politician will have much more time to apply control measures.

Force of infection (λ), the rate of susceptible individuals become infected by an infectious disease, that is, the number of new infections divided by the product of number of exposed and average duration of exposure.

Transmission rate (β), is calculated by the product of the transmission probability and contacts rate.

2.1.2 Kermack and McKendrick SIR Model

Kermack and McKendrick (1927) [66] SIR model is one of the earliest mathematical models in the history of epidemic model. The model categorizes population into **S**usceptible, **I**nfected, and **R**ecovered. Susceptible individuals

in S-stage have chance to be infected and progress to Infection I-stage until recovery to R-stage. The flow is shown in the following figure 2.2:

Figure 2.2: Flow of SIR model

We denote β as the transmission rate so the force of infection λ (rate of susceptible individuals become infected) would be βI, where I is the number of infectious individuals. By mathematical convention, we denote S, I and R as the subpopulations in each compartment for time t. The total population size N, is equal to $S + I + R$ for any time and $N = S$ for time zero. The SIR model can be written as the following system of nonlinear differential equations:

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \tag{2.1}$$

where β and γ are the model parameters and γ is the recovery rate. As the infectious period is assumed exponential distributed, we denote $1/\gamma$ as the average infectious period. By linearising the system [32], the basic reproductive numbers R_0 is equal to $\beta N/\gamma$.

Although the Kermack and McKendrick model is extremely simple and the assumption of exponential infectious period may not be hold in some situations [112], it works as a fundamental model in epidemic disease transmission over the past 70 years.

Apart from the deterministic version, the Kermack and McKendrick model has been extended stochastically. Using the same notation with a time indicator, t, Greenwood [56] adopted the binomial chain idea and the probability of susceptible individuals becoming infected would be

$$P(I(t+1) = i + k | S(t) = s, I(t) = i) = \begin{pmatrix} s \\ k \end{pmatrix} p^k (1-p)^{s-k} \tag{2.2}$$

for $k = 0, 1, 2, ..., s$. The p is the probability of a susceptible individual got the virus in the community and s, i, and k are constants. A limitation of the model is that the value of p is fixed over time which causes the epidemic ended in an arbitrary fast rate. The problem is improved in Reed-Frost model [2]; the probability of disease transmission depends on the number of infectious individuals at time t, i.e., $p(i) = 1 - (1-p)^i$. Bartlett (1949) [12] incoprorated the stochastic effects in deterministic SIR model theoretically and applied equally well in modelling the disease transmission for measles.

2.1.3 Susceptible-Exposed-Infectious-Removed $SEIR$ Model

$SEIR$ model is another common epidemic model with adding an Exposed (latent) compartment on SIR model. Latent period is defined as the period of time that individuals get infected but not yet infectious. Once suspectible individual get infected, they will refer to the Exposed E-stage and followed by Infectious I-stage. The flow is shown in the following figure 2.3:

The latent period is also assumed exponential distributed, so the average latent period is equal to $1/\alpha$. Following similar configuration of SIR model,

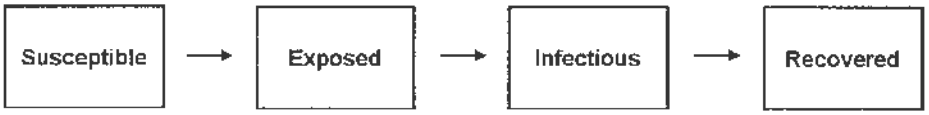


Figure 2.3: Flow of $SEIR$ model

the system of nonlinear differential equations of $SEIR$ model can be written as:

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dE}{dt} &= \beta SI - \alpha E \\ \frac{dI}{dt} &= \alpha E - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \tag{2.3}$$

The formula of the reproduction number in $SEIR$ model is the same as that in SIR model. However, $SEIR$ has a slower growth rate as the susceptible individuals require to pass through the latent class before contributing to the disease transmisson process.

2.1.4 SIR and $SEIR$ Model Extension

Since many essential factors contribute to the disease transmission in reality, simple epidemic SIR and $SEIR$ models cannot account for the effects of all factors; therefore, the classical models have been extended into different ways to incorporate various effects. For example, the time scale of the disease spread is slow and the demographical data such as population births and deaths would affect the disease dynamic, the system should incorporate the birth rate and mortality rate together. Another important factor that will affect the disease dynamic for a long time scale is the seasonal variation. Several literatures have

addressed this problem by replacing the infection parameter into time-varying periodic function or developing more compartments for measles, chickenpox, and influenza [73, 20]. Population heterogeneity, that is, is another important factor that different individuals may have different pattern from contracting to transmission, also affect to the disease dynamics. For example, children have higher risk to infect chickenpox due to the nature of contacts. With enough information, the model can offer essential information for those sub-classes such as influenza [85, 110].

Apart from individual characteristics, most of the disease models are developed to quantify the effectiveness on the pharmaceutical and the non-pharmaceutical interventions based on SIR model and $SEIR$ model during epidemic. Many articles have evaluated the effectiveness of pharmaceutical interventions such as antiviral treatments [3, 74] and vaccination [85] on influenza epidemics. In order to optimize the required resource, the final illness attack rate was assessed under different antiviral treatment and vaccine supply by multi-compartments models. Non-pharmaceutical interventions namely school closure [110], isolation, quarantine [21, 70, 119], and travel restriction [37] have also extended the $SEIR$ models to address the control measure properties. Given various pandemic situations, i.e., reproduction numbers, the models quantified the efficacy for these public health interventions. Because some single strategies may not be feasible, the combination strategies were also found to be effective for reducing the global spread of pandemic across a range of reproduction numbers [41, 75]. It enables policy makers to leverage on the effectiveness of some control measures and to reduce potential impact of others.

2.1.5 Travel Restriction in Epidemic Models

For the infectious disease with high transmissibility such as influenza, the global travel network plays a central role in geographical spread of disease. In current world, the speed and volume of international travel are unprecedented and the massive movement converge to favor the emergence of infectious disease. Travelers may carry pathogens in their bodies and thereby facilitate the introduction of a communicable disease into a new geographical area. Therefore, travel restrictions can reduce the rate of new infected people imported from or exported to different areas. So in model construction, it is necessary to consider the geographical feature, population characteristics, and travel patterns. Up to this moment, the epidemic models of global scale have been applied to specific outbreaks such as seasonal influenza [55], human immunodeficiency virus (HIV) [43], severe acute respiratory syndrome (SARS) [61], and recently, influenza A (H1N1) [44].

A large proportion of literature has studied the global disease spread problem from the meta-population, or 'patch' SIR and $SEIR$ model structures [100, 55, 44, 45, 27, 37] to assess the effectiveness of travel restrictions. The transmission between cities was connected by a symmetric air travel matrix (nxn) for which matrix elements (σ_{ij}) for row i and column j represented average daily passenger flow from city i to j. And a transportation operator (Ω_i) was implemented on the susceptible (S) and latent compartment (E) individuals at time t in the $SEIR$ model as follows:

$$\begin{aligned}\Omega(S_i(t)) &= S_i(t) + \sum_{j=1}^{N}[S_j(t)\frac{\sigma_{ji}}{n_j} - S_i(t)\frac{\sigma_{ij}}{n_i}] \\ \Omega(E_i(t)) &= E_i(t) + \sum_{j=1}^{N}[E_j(t)\frac{\sigma_{ji}}{n_j} - E_i(t)\frac{\sigma_{ij}}{n_i}]\end{aligned} \tag{2.4}$$

where N is the total number of cities and n_i is the populaton size of city i. Hence, the probability of travel for each time step is (σ_{ij}/n_i) from city i to city j. The meta-population approach would have thousands differential equations in the system as well as more than 3,000 airports in 220 countries [24]. However, some of the literatures did not consider the stochastic variation [100, 55] i.e. constant volume of passenger flow. Colizza (2006) [24], re-phased the model into stochastic version and the stochastic variable for the number of individuals travelled from city i to city $j, j+1, ...$ would follow a multinomial distribution. However, the model used the same infectious parameters for all cities which is inappropriate due to different contact patterns.

Besides the patch structure, there are some other model structures to study the traffic rescaling globally; for instance, multi-group based model and network model. Multi-group based model allows individuals to capture high levels of heterogeneity such as household structure, workplace structure, and school structure [39, 75]. These kinds of models can be used to investigate containment measures due to actual setting of locations [95]. Network model is another type of model structure that can also be applied in actual location settings represented by clusters and vertex. Riley, *et al.* has applied the spatial network transmission model in Great Britain locally to investigate the disease dynamic for smallpox [96]. However, the big challenge is that both model structures required large amount of information for groups. Moreover, the large number of parameters would make the parameter estimation much more complex. Additionally, those models usually study local spatial transmission, rather than global airline transmission as it is hard to incorporate the travel distance model in the social contact network.

2.2 Statistical Preliminaries: Relating Data to Models

In most of the epidemic models, parameter estimation is an essential part for model constructions. When relating data to model, we should maintain a balance between the completeness of data captured and the adequacy of the inference drawn on the data structure.

Least-square estimation, maximum likelihood (ML), and expectation maximization (EM) algorithm are common statistical estimation method employed in epidemic modeling studies. Least-square estimation method, minimising the sum of square differences between observed data and model prediction, is a typical approach dealing with estimation problems. Chowell, *et al.* [21] used least-square method to estimate the basic reproductive number (R_0) of the Ebola hemorrhagic fever outbreaks in Congo and Uganda. The method fitted the epidemiological data into a deterministic SEIR epidemic model. Maximum likelihood (ML) method is similar to least-square method which adapts the independent Gaussian errors in the epidemic models. Lekone, *et al.* [68] employed maximum likelihood estimation method to estimate the SEIR model parameters for an outbreak of Ebola in Congo in 1995. These methods are easy to be implemented; however, they are hard to solve the intractable likelihoods given high dimensional integral regions. Expectation maximization (EM) algorithm is another method to solve complex likelihood [14]. It works well for unimodal likelihoods, but it does not converge properly for multimodal likelihoods as the algorithm highly depends on the initial conditions.

Markov Chain Monte Carlo (MCMC) is a computational intensive approach to optimize the estimates via the Monte Carlo samples. High dimension of parameters space does not offer any obstacle to the MCMC method. Not only estimates parameters with uncertainty, the great flexibility of the MCMC method offers data augmentation for the unobserved process, especially the

non-reported data and unobserved compartments in the epidemic models during Markov chains generation. Therefore, the MCMC method has been employed in many epidemic modeling studies in the past [53, 35, 88].

The following sections will briefly introduce the background of MCMC and its relationship to Bayesian inference.

2.2.1 Markov Chain Monte Carlo

The MCMC optimize estimates by drawing samples from the **Monte Carlo** method and adapting the convergence through **Markov Chains**. The asymptotic property ensures the parameter converged in the realisations. The idea of Markov Chain Monte Carlo (MCMC) was first used in physics context in 1953 [76] and was generalised in statistical field by Hastings, *et al.* in 1970 [58]. With sufficient computing resource, MCMC became a famous computational algorithm for statistical community; it broadens horizons in Bayesian inference, stochastic processes, and statistical computing.

Monte Carlo sampling makes use of randomness to come up with the random variable estimates. It usually deals with the inferential problems which involve intractable integrations as well as multi-dimensionality. Given function f of interest and $p(x)$ a probability density corresponding to a random variable X, the expectation $\mathbb{E}_p(f)$ can be approximated by

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

if n is taken to be large enough. By the law of large numbers, the confidence of the random variable estimate increases with the number of sampling. Existing theories have proven that the degree of convergence from Monte Carlo sampling towards the true values of the variables. The greatest advantage of the estimation technique is easy in implementation.

A **Markov Chain** is a sequence of random variables $\{X_n\}$ which the probabilities of the values depend on its previous time. That is,

$$P(X_t = x_t | X_{t-1} = x_{t-1}, ..., X_0 = x_0) = P(X_t = x_t | X_{t-1} = x_{t-1})$$

There are two kinds of properties for a Markov Chain: Homogeneous and ergodicity. A homogeneous Markov Chain indicates that the transition probabilities will not change in the progression of states transitions. Provided that the number of iterations n approachs to infinity, the distribution, which is independent to the initial condition X_0, is equilibrium. If there is only one equilibrium distribution, the Markov Chain $\{X_n\}$ is called ergodic. That means, if $X_t \sim f$, then $X_{t-1} \sim f$. Most of the MCMC algorithms satisfy the above conditions and ensure the convergence to the target distributions; the Gibbs sampling and the Metropolis-Hastings algorithm are the common MCMC method to obtain the posterior estimates [48, 58, 76].

Gibbs sampling

Gibbs sampling is a MCMC scheme provided that the transition probabilities are formed by the full conditional distributions of parameters. It was firstly introduced by Geman, *et al.* [52] in an image processing publication. Suppose $\Theta = (\theta_1, ..., \theta_p)'$ and f are the parameter and distribution of interest respectively, the Gibbs sampling draws the samples from the successive generations from the full univariate conditional distributions alternatively i.e. $f_i(\theta_i | \theta_1, ..., \theta_{i-1}, \theta_{i+1}, .., \theta_p)$. Here is the algorithm [48]:

1. Start the iteration counter at $j = 1$ and set the initial values for $\Theta^{(0)} = (\theta_1^{(0)}, ..., \theta_p^{(0)})'$.

2. Sample candidate new value $\Theta^{(j)}$ from full conditional distributions

$$\begin{aligned}
\theta_1^{(j)} &\sim f_1(\theta_1|\theta_2^{(j-1)}, ..., \theta_p^{(j\ \ 1)}) \\
\theta_2^{(j)} &\sim f_2(\theta_2|\theta_1^{(j-1)}, \theta_3^{(j-1)}, ..., \theta_p^{(j-1)}) \\
&\vdots \\
\theta_p^{(j)} &\sim f_p(\theta_p|\theta_1^{(j-1)}, , ..., \theta_{p-1}^{(j-1)})
\end{aligned} \tag{2.5}$$

3. The **Gibbs sequence** is obtained. Change the iteration counter from j to $j+1$. Return to step 2 until convergence is reached.

By the law of conditional expectation, the distribution of interest can be estimated by the Monte Carlo average,

$$f(\theta_i) = E(f_i(\theta_i|.)) \approx \frac{1}{J}\sum_{j=1}^{J} f_i^j(\theta_i|.) \tag{2.6}$$

given J is sufficiently large enough. The Markov chains approach to stationary distribution given sufficient large number of iterations after the burn-in period. In extension of Gibbs sampler, random selection can be adopted for the number of parameter updates, and this sampling method is called random scan Gibbs sampler. The MCMC method has been demonstrated its wide variety of application in statistical aspects, and its great flexibility of usage in the Bayesian statistics [51, 103].

Metropolis-Hastings algorithm

The Metropolis-Hastings (MH) algorithm is one of the parameter updating schemes to generate a convergent distribution. Once the non-iterative generation of probability distribution f is complex especially in high dimension

integration, the Metropolis algorithm is able to draw f without its specific forms of parametric conditional distributions [76]. Given suitable conditions [48], the Markov chain correctly converges to target distribution.

Suppose paramters Θ, the Metropolis-Hastings algorithm produces sequence of draws as follows [48]:

1. Start the iteration counter at $j = 1$ and set the initial values for $\Theta^{(0)}$.

2. Sample candidate new values Θ' from the **proposal density** $\alpha(\Theta'|\Theta^{(j-1)})$, which is a probability of generating Θ' given previous values $\Theta^{(j-1)}$. The proposal density must be symmetric in Metropolis sampling, i.e. $\alpha(\Theta'|\Theta^{(j-1)}) = \alpha(\Theta^{(j-1)}|\Theta')$.

3. Accept the new values Θ' with probability $min(1, A)$, where

$$A = \frac{f(\Theta')\alpha(\Theta'|\Theta^{(j-1)})}{f(\Theta^{(j-1)})\alpha(\Theta^{(j-1)}|\Theta')} \tag{2.7}$$

   If accepted, $\Theta^{(j)} = \Theta'$; otherwise, $\Theta^{(j)} = \Theta^{(j-1)}$ and the chain does not move.

4. Change the iteration counter from j to $j + 1$ and return to step 2 until convergence is reached.

It can be proven that (not showed here) the Metropolis-Hastings algorithm ensures the reversibility property of the Markov chain for the pair (x, y) with respect to f, i.e.

$$f(x)\alpha(x|y) = f(y)\alpha(y|x) \tag{2.8}$$

Therefore, the transition probabilities for the chains are the same for both direction $x \rightarrow y$ and $y \rightarrow x$. Consequently, the algorithm generates a Markov chain $\{\Theta^0, \Theta^1, ..., \Theta^k, ...\}$ and the transition probability from Θ^j to Θ^{j+1} only depends on Θ^j but not $\{\Theta^0, \Theta^1, ..., \Theta^{j-1}\}$.

Random walks are common proposals for the MH algorithm and there are many random walk proposal distributions such as gamma, uniform, and Gaussian. Suppose a Gaussian random walk step, the new values Θ' is generated from,

$$\theta' = \theta^{(j-1)} + \epsilon \tag{2.9}$$

where ϵ follows a symmetric normal density i.e. $N(0, \sigma^2)$. The σ is a step size of the chain and $\theta' \sim N(\theta^{(j-1)}, \sigma^2)$. Tuning σ affects the acceptance rate for the parameter updates and the acceptance rate is suggested to be around 20% to 40% for a good convergence mixing. Roberts, *et al.* recommended around 23% acceptance rate is optimal for a Gaussian random walk MH algorithm [98]. Typically, the selection of the proposal distribution is somehow arbitrary; it is similar to that of priors, which poorly mismatch of high density region would likely to converge slowly.

Implementation and Diagnosis

In general, the objective of the MCMC algorithms is to obtain the stationary density for the chain in a number of runs. Since the rate of convergence usually depends on initial starting points, the sampler, and the posterior density space [98], a number of initial steps, which regarded as the *burn-in* period, are discarded in order to minimize the effect from the initial non-convergence. Given enough number of iterations, the Markov chain is most likely to converge

to its stationary distribution. Inefficient start points would greatly extend the required burn-in period; it is suggested to start the simulation with a point closed to the mode of our target distribution. Sometimes the chains are *thinned* in order to eliminate the correlation among samples. Suppose n is any fixed value, we usually take the every $1n$-th, $2n$-th, ... iterations as a kind of *thinning* methods with interval n.

MCMC diagnosis is essential to identify problems with convergence. We can monitor the convergence by the time series trace plot, that is, the plot of generated values versus the number of iterations. As for Metropolis-Hastings algorithm, the convergence can be diagnosed by the time series trace plot. Good mixing of chains would show no trend, presumably toward a stationary state. The lagged autocorrelations plot is another graph to monitor the underlying correlation structure for the time series. If the samples are highly correlated, slow convergence of the ergodic average posterior estimates would likely occur which means the chains receive small amount information from the iterations. The samplers behave good at autocorrelation if the chains have the geometric decay trend in the lagged autocorrelations plot.

2.2.2 Bayesian Inference

Different from classical likelihood inference, Bayesian does not treat parameters as fixed but draws the estimate by repeated sampling principle, which adapts prior distributions on the model parameters. Using Bayes theorem, the posterior distribution of interest is calculated by the combination of the prior and the likelihood. Suppose parameters are denoted by Θ, the posterior distribution is equal to

$$
\begin{aligned}
P(\Theta|data) &= \frac{L(\Theta)\pi(\Theta)}{P(data)} \\
&= \frac{L(\Theta)\pi(\Theta)}{\int_{\Theta} L(\Theta)\pi(\Theta)d\Theta} \\
&\propto L(\Theta)\pi(\Theta)
\end{aligned}
\tag{2.10}
$$

which $L(\Theta)$ is the likelihood function from observed data and $\pi(\Theta)$ is the prior distribution assigned to the parameters. The formula is called Bayes' theorem. The denominator in the formula is a normalising constant. It is hard to be calculated in Bayesian settings, however, it can be resolved by computational sampling method, like MCMC.

Regarding prior distributions, the selection of the priors is usually based on epidemiological beliefs. *Conjugate* priors, which lead the posterior belonging to the same family distribution, are common in practice and they are usually computational convenient in MCMC method. Besides, it is also common to use non-informative priors to provide a baseline assumption for analysis. An non-informative prior means that the probability of every candidate value of parameter θ is equal, i.e.,

$$
p(\theta) = \frac{1}{b-a}, a < \theta < b
$$

given a bounded continuous parameter space $[a, b]$. In MCMC methods, the data dominate the posteriors in stationary stages whatever the prior information is, so the exact forms of the priors are not important in most scenarios.

In Bayesian analysis, we are interested in drawing the posterior distribution; among the inference problems, most of them come down to expectation

calculation and MCMC is powerful enough to achieve the estimates. Once the converged posteriors are obtained, the point estimation and interval estimation in addition to the probability density plots would normally be our practical interest.

Mean, median, and mode are the common measures of central tendency for the posterior distributions. The decision of selecting the measures of central location depends on the shape of the posterior; for example, if the posterior is unimodal and symmetric, the three central tendency measures coincide. In asymmetric posteriors, median is usually prefered since mean is affected by the outliers heavily and mode maybe close to the non-representative peak. Apart from location measures, measures of dispersion, for example, credible intervals, are other statistics of interest. Typically, the lower bound and the upper bound of the credible interval are simply taken as the $\alpha/2$-th quantile and the $1-\alpha/2$-th quantile of $P(\Theta|data)$ respectively.

2.2.3 Application in Epidemic Models

Due to the great flexibility, the Bayesian inference and the MCMC method have been employed widely in the epidemic modeling studies [94, 118, 53, 54, 68, 35, 88]. In practice, the dynamic model structures would make the inference much more complex. Besides, epidemic data, such as the times of infection are usually unobserved [88, 54]. With extensive available computing power, computational simulation methods like MCMC perform efficiently on solving complicated likelihoods. Moreover, the MCMC methods are well-suited to data augmentation even for large dimensionality. Hence, the estimated epidemic models would be more adequate than that of using reference values to describe the disease transmission mechanism as well as the intervention effectiveness.

In order to draw the inference for epidemic models, the observations are

usually adopted from either the times of status or the incidence count. O'Neill, *et al.* used the times between each detection of measles cases in household outbreaks to analyze the distributions of infections by the MCMC method [87]. The literature also demonstrated the augmentation skill of MCMC to impute the unobserved infection times [88, 86, 87]. The MCMC methods like Gibbs sampling and Metropolis-Hastings algorithm were found performing well in some situations, but converged badly for some parameters; the convergence testing is thus important. Lekone, *et al.* [68] employed a Bayesian method on $SEIR$ model to study the Ebola outbreak from the daily counts of reported cases and reported deaths. The MCMC did well on imputing the unobserved series but had to use of final outbreak size in data augmentation. Besides, the article did not describe any method in solving the problems of non-reported data. Apart from them, the Bayesian inference has also been drawn for several epidemic models in order to solve the complex likelihoods and to provide uncertainty estimates [35, 118].

2.3 Identification of Research Question

In this thesis, we will fully evaluate the value of travel restriction. The travel restriction will be assessed for two wide areas: 1, the impact for *at-risk* countries and 2, the impact for the *source* country. In addition, some of the current limitations from the literature review will be overcome.

2.3.1 Impact of Travel Restriction: *At-risk* Countries

The impact of travel restriction for *at-risk* countries has been studied in many epidemics. However, the research topic is still of interest for policy makers and epidemiologists, as the implementation is controversial over past decades and there are several limitations in these studies.

Most of the literatures support that the travel restriction is a valuable intervention. Epstein , *et al.* [37] simulated the scenarios with air travel restrictions by a global $SEIR$ model and demonstrated reductions on the cumulative incidence in the early period of an epidemic. Besides, a 95% travel restriction was able to delay the first passage times (FPT) for more than two weeks. Compared to other control measures, a moderate proportion of travel rescaling could not reduce a certain amount of final illness attack rate. However, it was important to delay the epidemics especially when the initial growth rate was relatively low [117]. Colizza *et al.* [25] employed a meta-population stochastic model to indicate that air travel restriction was able to decrease the probability of global outbreaks. Wood, *et al.* [117] have investigated the effectiveness of the internal border control on limiting influenza spread in the context of Australia, and demonstrated that it could delay the pandemic for several weeks between two cities. The authors also noted that the travel restriction worked better in more isolated communities that lacked international ports. Brownstein, *et al.* [18] used weekly influenza and pneumonia mortality data to illustrate the association between the decrease of volume in air travel and the time delay to a influenza season.

In contrast to above findings, Cooper, *et al.* [27] showed that even if more than 90% of air travels had been blocked, the rate of the global spread only would have achieved a little reduction once major outbreaks were underway. Similar to the study of Cooper, *et al.*, Hollingsworth, *et al.* [60] employed a $SEIR$ model to conclude that the travel restriction only slowed down the export process of infected cases, instead of halting the spread even if 99% of air travel was banned.

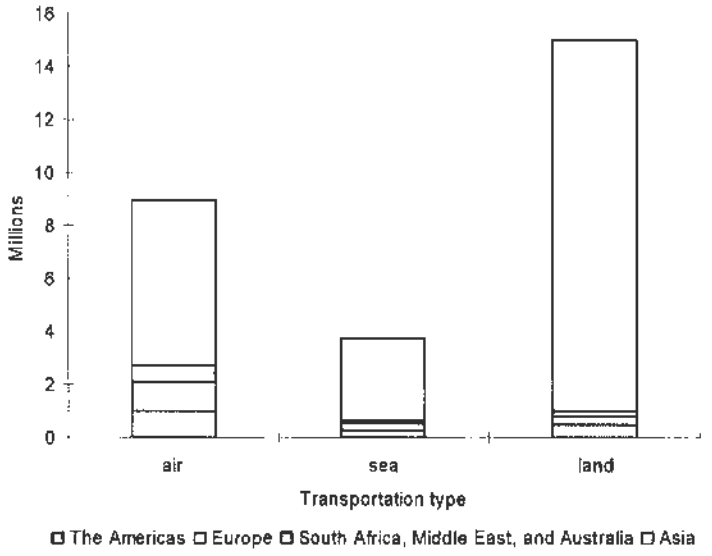
Apart from the above research motivation, there are several major limitations that can be improved in current literatures. Firstly, previous studies focused only on air travel restrictions, but in many cities, including Hong

Kong, it is not the main means of transport for arriving and departing travelers. Statistics show that annually more than half of the passengers who arrived in Hong Kong came either by sea or land [62]. Figure 2.4 illustrates that more than ten million visitors came to Hong Kong from Asia by land transport annually. Visitors from North America and Europe contributed a higher proportion of air transport arrivals. Therefore, the incorporation of air, sea, and land transport was necessary to demonstrate the effectiveness of travel restrictions. Secondly, most of the mathematical models only took into account the latent individuals who traveled between countries. But with only a limited screening sensitivity at the border points of entry [28], a large number of infected cases could enter, thereby resulting in a large increase in the rate of disease transmission locally [15]. Such studies may therefore provide misleading information on the effect of travel restrictions [23]. Thirdly, most of the geographical epidemic models were deterministic and they did not give consideration to the stochastic variation [10, 9]. Fourthly, they ignored the city heterogeneities in force of infections. For example, they assumed all countries had the same number of imported cases [4]. Given the information on disease transmission among various countries, meta-population models were preferred to distance transmission models and to network transmission models [95] which arbitrarily adopted meanings in point-to-point transmission.

In chapter 3, a mathematical model will be developed to describe the disease spread and to explore the impact of travel restriction on the influenza A (H1N1) pandemic by using Hong Kong as a case study. The model will adapt the following properties to improve the current limitations:

- Air, sea, and land transportation;
- Limited screening sensitivity of the border points of entry;
- Stochastic uncertainty;

Figure 2.4: Total arrivals (in millions) by air, sea, and land transport in 2007.



Countries are allocated to different categories: the Americas, Europe, Asia, South Africa, Middle East, and Australia. Forty-four countries were selected in total which contributed more than 95% of arrivals to Hong Kong.

- Spatial heterogeneities in force of infections i.e. different numbers of import cases according to the initial growths of the countries.

In addition, the effectiveness of antiviral drugs and hospitalization is also investigated according to the strategies from the Department of Health, Hong Kong [90].

2.3.2 Impact of Travel Restriction: *Source* Country

In the early phase of an epidemic, the top priority of mitigating strategies is to contain the pandemic outbreak at the *source* country. The public health

measures, like the travel restrictions, play an important role to reduce the possibility of infected cases exported to other areas as well as to delay the spread from the source area. In view of *source* country, epidemic studies usually focus on the time delays to epidemic seeded to other areas from the infected source rather than the illness attack rates, which is a different perspective to *at-risk* countries.

With the understanding to the distribution of the exported infections from the source country, experts are able to assess the possibility of disease containment and to have better preparation for the control measures, like the border control. However, researchers have to face the problems of the time delay until the first official disease confirmation and the non-reporting rate, while formulating the distribution. As for the influenza A (H1N1) pandemic, it was believed that the virus has been circulated within communities several months before the recognition of the disease outbreak [63, 71]. Before the active surveillance of influenza A (H1N1) and the confirmative diagnosis from clinicians and microbiologists, the virus was undetected over a period of time. Several studies estimated the initial point of the disease outbreak around the mid-January to late-Feburary through the analysis of the viral genetic sequence and the epidemic models [47, 106], and the delay would have significant impact on simulation results [19, 40]. Apart from the initial time delay, the reporting rate was low for the influenza A (H1N1) pandemic. Most of the ascertainment was particularly focused on cases with severe condition. Also, either asymptomatic or mild cases were not presented in medical consultation. A good example of official surveillance being under-estimated the disease transmission intensity in the community would be the telephone interviews from the Beijing Center for Disease Prevention and Control (CDC) [120], which showed that the consultation rate among influenza-like illness (ILI) patients was no more than 50% in Beijing, China.

Although the issues of initial reporting delay and non-reporting are important, most of the epidemic modeling studies have neglected these factors in model development. Caley, *et al.* [19] quantified the distribution of the initial time delay with justified factors, like in-flight transmissions and the probability of screening at exit and entry borders. Given p_d the probability of the epidemic initiated on day d followed by identification at source region, they drew the probability distribution of the time delay (D) until the epidemic was first initiated in the *at-risk* country as

$$Pr(D = d) = (1 - p_1)(1 - p_2)(1 - p_3)...(1 - p_{d-1})p_d$$

They concluded number of travelers who attempted to enter the at-risk countries largely determined by the rate of country-to-country spread. But the study did not deal with the non-reporting issue. Hollingsworth, *et al.* [60] constructed an epidemic model to investigate the impact of travel restriction. However, the probability of exported cases to countries was arbitrarily assumed. Most importantly, no estimation was done on the estimation of model parameters for the studies.

In addition, most of the epidemic models are deterministic without the consideration of the stochastic variation. The lack of the model uncertainty would be an obstacle to justify the significance of the modeling outcomes. In general, the use of bayesian approach is recommended, as it is able to incorporate the uncertainty of parameters along with the stochastic variation. The bayesian approach is also prefered to univariate-vary the parameters for a sensitivity analysis.

Consequently, a mathematical model which incorporated the effect of initial reporting delay and the reporting rate behind the surveillance data will be

developed in chapter 4 to demonstrate the impact of travel restriction by assessing the probability distribution of exported cases from the source country. The model will adapt the Bayesian approach and will employ the influenza A (H1N1) pandemic in Mexico as a case study. The model incorporates the following properties:

- Initial reporting delay;
- Under-reporting;
- Statistical Inference on model parameters.

The model is able to offer insights of the initial epidemic dynamic to epidemiologists, and to advise policy makers to have a better management on containing an epidemic at the source country.

Chapter 3

Impact of Travel Restriction: *At-risk* Countries

In this chapter, we studied the impact of travel restriction on the influenza A (H1N1) pandemic in views of *at-risk* countries through a stochastic compartmental model. As for most of the previous epidemic modeling studies, they only focused on air travel restriction and assumed 100% screening sensitivity at the border points of entry which were not realistic. In section 3.2, we developed an epidemic model which incorporated all means of air, land, and sea transport with stochastic uncertainty. In addition, the use of antiviral and hospitalization were also adopted in order to provide a more realistic compartment on the recovery, and also to compare the effectiveness of these control measures. The model was then applied to the influenza A (H1N1) pandemic in Hong Kong and the modeling results were demonstrated in section 3.3.

According to our result, we concluded that the greatest value of travel restrictions was in their ability to slow down the spread of the epidemic. With the imposition of other interventions that can suppress the disease transmission intensity, whether locally or not, the restrictions on all external travel reduced the local attack rates, and they even halted the disease spread. Similar to the findings of other previous research, solely implementing travel restrictions was not completely effective in reducing the attack rates, especially during the

severe scenarios. In practice, the pharmaceutical interventions, like vaccine and antiviral, are usually not available early enough once a new emerging virus has arrived in the community. So the travel restriction is a simple and direct non-pharmaceutical intervention to slow down the epidemic during the early stage, in order to allow a longer period for the preparation of the mitigation response, especially for the next emerging virus with unknown characteristics. The details of discussion were highlighted in section 3.4.

3.1 Introduction

For infectious diseases with high transmissibility, such as influenza, the traveling patterns of individuals play an essential role in the geographical spread of disease. Travelers may carry pathogens in their bodies and thereby facilitate the introduction of a communicable disease into a new geographical area. Travel restrictions are a kind of social control measure that have been evaluated in several epidemics such as influenza [55], human immunodeficiency virus (HIV) [43], SARS [61], and, recently, influenza A (H1N1) [44]. Nevertheless, not all the relevant literature supports the value of air travel restrictions for containing the epidemic [16-19]. The studies have shown that air travel restrictions have only a limited benefit in slowing the global spread of a pandemic influenza. Besides, travel restrictions have low social acceptability, and they may also have a huge impact on the economy.

Despite these factors, the investigation of the value of travel restrictions remains essential. Firstly, previous studies focused only on air travel restrictions, but in many cities, including Hong Kong, it is not the main means of transport for arriving and departing travelers. Statistics show that annually more than half of the passengers who arrived in Hong Kong came either by sea or land [62]. Figure 2.4 illustrates that more than ten million visitors came to Hong Kong from Asia by land transport annually. Visitors from North America and

Europe contributed a higher proportion of air transport arrivals. Therefore, the incorporation of air, sea, and land transport was necessary to demonstrate the effectiveness of travel restrictions. Secondly, most of the mathematical models only took into account the latent individuals who traveled between countries. But with only a limited screening sensitivity at the border points of entry [28], a large number of infected cases could enter, thereby resulting in a large increase in the rate of disease transmission locally [15]. Such studies may therefore provide misleading information on the effect of travel restrictions [23]. Thirdly, most of the geographical epidemic models were deterministic and they did not give consideration to the stochastic variation [10, 9]. Fourthly, they ignored the city heterogeneities in force of infections. For example, they assumed all countries had the same number of imported cases [4]. Given the information on disease transmission among various countries, meta-population models were preferred to distance transmission models and to network transmission models [95] which arbitrarily adopted meanings in point-to-point transmission. In addition, other control measures, such as antiviral drugs and hospitalization, should be included in the model in order to better manage the spread of the disease and the way it is controled.

In our study, an epidemic mathematical model was developed to describe the disease spread and to explore the impact of travel restrictions via air, sea, and land travel on the influenza A (H1N1) pandemic in Hong Kong. We also studied the effectiveness of antiviral drugs and hospitalization for the comparison. Furthermore, we investigated some important effects of changes, including reproduction numbers from non-local countries to Hong Kong, the screening sensitivity at entry border points, the implementation date on travel restrictions, and the length of latent period. The results will provide valuable information to government policy-makers and to public health experts.

3.2 Model Formulation

We extended the discrete stochastic $SEIR$ model [56, 2, 68] to study the influenza A (H1N1) dynamic and the impacts of the interventions locally. The model was developed to adapt the arrivals both of latent and infectious individuals by means of air, land, and sea transport and the use of antiviral and hospitalization with stochastic uncertainty.

3.2.1 Basic Stochastic $SEIR$ Model

Let Δt be a time step and $(t, t-\Delta t]$ be a time interval, we denote $S(t)$, $E(t)$, $I(t)$, and $R(t)$ as the number of individuals in **S**usceptible, **E**xposed, **I**nfected, and **R**ecovered compartments at time t, respectively. Suppose $B(t)$ is the incidence, the number of susceptible become infected and $C(t)$ is the number of infected individuals who start to be infectious at time t. And $D(t)$ is the number of individuals who recover or die from infectious state at time t. Assume the population is homogeneously mixed, the system of general SEIR stochastic model with no intervention is

$$\begin{aligned} S(t+\Delta t) &= S(t) - B(t) \\ E(t+\Delta t) &= E(t) + B(t) - C(t) \\ I(t+\Delta t) &= I(t) + C(t) - D(t) \\ R(t+\Delta t) &= R(t) + D(t) \end{aligned} \tag{3.1}$$

An individual would have a probability p to get into next stage which follows a bernoulli distribution. So by given n individuals, the number of individuals who get into next stage would follow a binomial distribution with probability m. We take $\text{bin}(m, n)$ as a binomial distribution with parameters

probability m and number of total individuals n. The corresponding distributions for the classes

$$\begin{aligned} B(t) &\sim bin(1 - exp[-\frac{\beta}{N} I(t)\Delta t], S(t)) \\ C(t) &\sim bin(1 - exp(-\alpha\Delta t), E(t)) \\ D(t) &\sim bin(1 - exp(-\gamma\Delta t), I(t)) \end{aligned} \tag{3.2}$$

where the rate of infection is equal to $\beta I(t)/N$ for a time step where β is the transmission rate and N is the population size. The α and γ are the constant transition rates from latent state to infectious state and from infectious state to removed state respectively. And the rates are transformed into probabilities assuming in poisson process.

3.2.2 Arrived and Departed Cases

In the disease transmission model, latent $(IM^E(t))$ and infectious $(IM^I(t))$ travelers arrive from other countries by transport k-th and come to the compartments $E(t)$ and $I(t)$. A single population model adapts the travel effect from 3 modes of transport: sea, land, and air. Suppose the probabilities of travel are the same for all individuals and the probability of travelers import from country i-th $(i = 1, 2, ..., p)$ are represented by $m^I_{k,i}$ by mode of transport k $(k = 1, 2, 3)$ for air, sea, and land respectively.

Here are the model compartments of imported cases in latent status,

$$IM^E(t) = \sum_{k=1}^{3} \sum_{i=1}^{p} bin(m^I_{k,i}, E_i(t)) \tag{3.3}$$

and infectious status,

$$IM^I(t) = \sum_{k=1}^{3} \sum_{i=1}^{p} bin(m_{k,i}^I, I_i(t)) \tag{3.4}$$

The number of latent subjects, $E_i(t)$, and the number of infectious subjects, $I_i(t)$, at time t of country i-th are generated from discrete-time $SEIR$ model based on the reproduction numbers of the countries,

$$\begin{aligned} E_i(t + \Delta t) &= E_i(t) + S_i(t)[1 - exp(-\beta_i \Delta t I_i(t)/N_i)] - E_i(t)[1 - exp(-\alpha \Delta t)] \\ I_i(t + \Delta t) &= I_i(t) + E_i(t)[1 - exp(-\alpha \Delta t)] - I_i(t)[1 - exp(-\gamma \Delta t)] \end{aligned} \tag{3.5}$$

where $1 - exp(-\beta_i \Delta t I_i(t)/N_i)$, $1 - exp(-\alpha \Delta t)$, and $1 - exp(-\gamma \Delta t)$ are the per capita probabilities of infection, becoming infectious, and becoming recovered respectively given transmission parameter β_i in population N_i. Individual transmission parameter β_i is calculated from the basic reproduction number (R_0) of country i-th. It is defined as the average number of secondary infections produced by a typical infected individual in a wholly susceptible population. In order to allow the transmission heterogeneities between non-local countries, we will estimate the reproduction numbers by the initial exponential growth rate method [21] employing two months after dates of their first onset cases daily surveillance data [115] [34] [47],

$$R_0 = 1 + \frac{r^2 + (\alpha + \gamma)r}{\alpha\gamma} \tag{3.6}$$

where r is the initial exponential growth rate estimated by the least square fitting to the model, i.e. $logarithm$(cumulative number of cases at time t) $\propto rt$.

At the same time, a number of infected individuals will leave and carry the pathogens away from the local city. Let m_k^E be the probability of departure from local area by the mode of transport k, the compartments of exported cases in latent status, $EX^E(t)$, and in infectious status, $EX^I(t)$, will be $\sum_{k=1}^{3} bin(m_k^E, E(t))$ and $\sum_{k=1}^{3} bin(m_k^E, I(t))$ respectively. The compartments of exported cases in latent status will be,

$$EX^E(t) = \sum_{k=1}^{3} bin(m_k^E, E(t)) \tag{3.7}$$

and in infectious status will be,

$$EX^I(t) = \sum_{k=1}^{3} bin(m_k^E, I(t)) \tag{3.8}$$

Given ν is the sensitivity of the entry screening board, so only $1-\nu$ proportion of imported infectious individuals are able to be scanned [28], the system of the stochastic equations:

$$\begin{aligned}
S(t+\Delta t) &= S(t) - B(t) \\
E(t+\Delta t) &= E(t) + B(t) + IM^E(t) - EX^E(t) - C(t) \\
I(t+\Delta t) &= I(t) + C(t) + (1-\nu)IM^I(t) - EX^I(t) - D(t) \\
R(t+\Delta t) &= R(t) + D(t)
\end{aligned} \tag{3.9}$$

A simple schematic flow is shown in Figure 3.1.

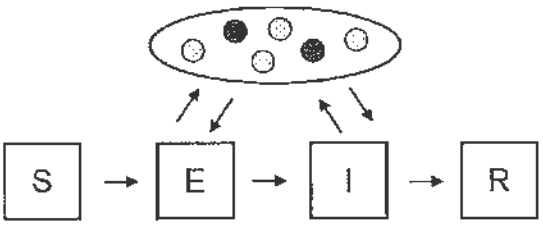


Figure 3.1: Schematic flow of $SEIR$ model which incoporates import-export latent and infectious individuals

3.2.3 Antiviral and Hospitalization

Two new compartments are added into the model, antiviral Treatment $T(t)$ and Hospitalization $H(t)$. Once individuals become infectious, they seek for antiviral treatment and hospitalization with proportions p_T and p_H respectively. With regard to limited resources, part of them may be untreated as proportions p_U. We adapt a ψ fraction reduction of infectiousness for individuals who receive antiviral. Suppose classes $M(t)$ and $N(t)$ are the number of infectious individuals who take antiviral treatment and hospitalization at time t respectively. The $P(t)$ and $Q(t)$ are the number of removed individuals from antiviral treatment and hospitalization with transition rates γ_T and γ_H to the removed status. Let f_k be the restriction fraction for import transportation k-th, the stochastic system is as follow,

$$
\begin{aligned}
S(t+\Delta t) &= S(t) - B(t) \\
E(t+\Delta t) &= E(t) + B(t) + (1-f_k)IM^E(t) - EX^E(t) - C(t) \\
I(t+\Delta t) &= I(t) + C(t) + (1-f_k)(1-\nu)IM^I(t) - EX^I(t) - D(t) - M(t) - N(t) \\
T(t+\Delta t) &= T(t) + M(t) - P(t) \\
H(t+\Delta t) &= H(t) + N(t) - Q(t) \\
R(t+\Delta t) &= R(t) + D(t) + P(t) + Q(t)
\end{aligned}
\tag{3.10}
$$

Because infectious individuals include those being treated and hospitalized, the probability of a susceptible person becoming infected is equal to $1 - exp[\beta[I(t) + (1-\psi)T(t) + H(t)]/N]$ for a time step Δt where β is the transmission rate. The corresponding distributions for the classes,

$$
\begin{aligned}
B(t) &\sim bin(1 - exp[-\frac{\beta}{N}[I(t) + (1-\psi)T(t) + H(t)]\Delta t]), S(t)) \\
C(t) &\sim bin(1 - exp(-\alpha\Delta t), E(t)) \\
M(t) &\sim bin(p_T\Delta t, I(t)) \\
N(t) &\sim bin(p_H\Delta t, I(t)) \\
D(t) &\sim bin(p_U[1 - exp(-\gamma_R\Delta t)], I(t)) \\
P(t) &\sim bin(1 - exp(-\gamma_T\Delta t), T(t)) \\
Q(t) &\sim bin(1 - exp(-\gamma_H\Delta t), H(t))
\end{aligned}
\tag{3.11}
$$

In this model, γ_R is the transition rates from infectious state to removed state. The γ_T and γ_H are the transition rates from the treatment state to the removed state and from the hospitalization state to the removed state

respectively. The detailed descrptions of the parameters are highlighted in Table 3.2. A simple schematic flow is showed in Figure 3.2.

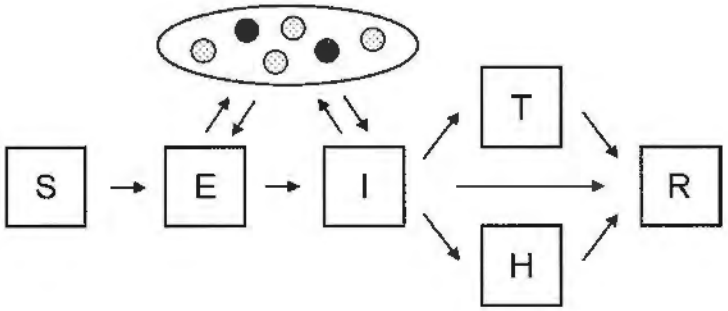


Figure 3.2: Schematic flow of $SEIR$ model which incorporates the compartments treatment and hospitalization with import-export latent and infectious individuals

3.3 Case Study: Effectiveness of Travel Restriction for 2009 Influenza A (H1N1) Pandemic in Hong Kong

Using the 2009 influenza A (H1N1) pandemic as a case study, the methodology was employed to assess the effectiveness of the travel restriction via different transportation in Hong Kong. The effectiveness of the antiviral and hospitalization was also studied for the comparison. Furthermore, the effects of changes in the reproduction numbers from the non-local countries, the screening sensitivity on entry border points, the implementation date of travel restrictions, and the length of latent period were investigated. The results would offer valuable advice to the government policy makers and the public health experts.

3.3.1 Materials

Population and Transportation

Population data were taken from the International Database (IDB), U.S. Census Bureau [89]. The individual probability of travel for each country was calculated by the daily rate of travel divided by the population size. The arrival data were taken from visitor arrival statistics provided by the Hong Kong Tourism Board [62]. The statistics included the total number of arrivals by countries with their modes of transport. Forty-four countries were selected which annually contributed more than 95% of the arrivals to Hong Kong (Figure 2.4). The yearly statistics for frequency of departures from Hong Kong residents by different modes of transport were collected from the Census and Statistics Department, Hong Kong [38]. The data were assumed to be uniformly distributed daily.

Epidemiological Details

Since we did not have any available information of the cross-immunity from past influenza infections, the initial population was set to be 100% susceptible. The local daily surveillance of confirmed infected cases was from the press releases on human swine flu, Department of Health, Hong Kong [92]. The average latent and infectious periods of influenza A (H1N1) were set to 1.45 and 2.9 days respectively [44, 16]. The length of the latent period would be tested in a sensitivity analysis for the values of a half-day and two days.

The latent ($IM^E(t)$) and infectious ($IM^I(t)$) travelers were based on the discrete-time $SEIR$ model which depended on the basic reproduction numbers. The reproduction numbers were estimated by the initial exponential growth rate method [21] where daily surveillance data ,from the World Health Organization (WHO) [115] and the European Centre for Disease Prevention and Control (ECDC) [34] pandemic H1N1 situation updates, of two months

Table 3.1: Frequency of departures and arrivals by countries with the modes of transports in 2007

Country	Total	Mode of transport		
		Air	Sea	Land
Departure				
Hong Kong	80,682,000	6,141,000	8,871,000	65,670,000
Arrival				
United States	1,230,927	724,023	191,178	315,726
Canada	395,167	219,469	59,004	116,694
Honduras	1,662	675	225	762
Mexico	35,706	21,260	5,821	8,625
Argentina	10,515	5,690	1,805	3,020
Brazil	40,339	19,861	8,061	12,417
Venezuela	10,896	4,356	1,612	4,928
United Kingdom	601,168	448,647	68,007	84,514
Netherlands	110,816	70,592	15,712	24,512
Denmark	30,013	18,734	4,193	7,086
Finland	21,830	13,365	3,448	5,017
Norway	18,624	12,381	2,327	3,916
Sweden	49,810	30,909	7,449	11,452
Austria	24,046	14,514	4,529	5,003
Germany	234,763	149,370	38,523	46,870
Switzerland	46,870	32,529	6,561	7,780
France	231,091	135,291	41,515	54,285
Belgium	32,413	20,190	5,114	7,109
Italy	118,841	73,043	17,564	28,234
Portugal	18,639	9,419	8,199	1,021
Spain	65,131	38,460	10,757	15,914
Russia	32,858	21,256	4,314	7,288
South Africa	72,897	47,001	4,357	21,539
Bahrain	2,500	1,833	106	561
Egypt	16,361	7,764	579	8,018
Israel	63,435	38,692	9,537	15,206
Jordan	11,084	4,809	333	5,942
Kuwait	4,366	2,838	283	1,245
Saudi Arabia	19,435	13,616	787	5,032
Turkey	41,011	20,619	2,764	17,628
United Arab Emirates	11,881	9,358	615	1,908
Australia	633,599	418,760	83,173	131,666
New Zealand	117,215	82,461	10,762	23,992
Japan	1,324,336	748,478	273,334	302,524
South Korea	876,231	507,872	136,095	232,264
Indonesia	366,217	185,197	63,102	117,918
Malaysia	504,487	237,542	105,036	161,909
Philippines	552,942	365,490	70,956	116,496
Singapore	631,963	393,423	93,794	144,746
Thailand	387,219	246,732	47,800	92,687
India	317,510	178,018	33,588	105,904
Taiwan	2,238,731	1,248,228	123,793	866,710
Macau	626,103	30,547	553,682	41,874
China	15,485,789	2,069,683	1,618,643	11,797,463

The statistics were from the Hong Kong Tourism Board [62] and the Census and Statistics Department, Hong Kong [38].

after the first onset case was used for model fitting. The surveillance data from different countries were listed on internet [113]. A sensitivity analysis was performed on the R_0s of the non-local countries for values with 20% increases and 20% decreases respectively.

3.3.2 Scenario Design

The mathematical model was developed to assess the effectiveness of: (i) the travel restrictions relating to different means of transport; and (ii) the use of the antiviral and the hospitalization for the influenza A (H1N1) pandemic in Hong Kong locally. The travel restrictions were supposed to take effect on the day after the first global onset case. Different start dates were tested in the sensitivity analysis. The antiviral and the hospitalization strategies were implemented locally 3.5 months after the first global onset case, which was similar to the strategies from the Department of Health, Hong Kong [90].

Travel restrictions relating to sea, land, and air travel. We applied 90% and 99% import restrictions (f_k) on different combinations of k-th transport. We also considered that only one-third of infectious cases was identifiable at the entry borders at the baseline scenario[28]. A 95% value and a 5% value of screening sensitivities were tested in the sensitivity analysis section. In addition, the start date on the travel restrictions was tested for three months and five months after the first global onset case, respectively.

antiviral and hospitalization. We assumed that 12% (p_T) of the infectious subjects were offered antiviral and 6% (p_H) of the infectious subjects were selected for hospitalization, according to the experience in influenza pandemic [49, 119]. The remaining 82% (p_U) of infectious individuals were untreated. The antiviral would reduce 60% infectiousness for the individuals [40]. Both interventions would reduce the average infectious period by 1.5 days [7]. Compartments for antiviral $T(t)$ and hospitalization $H(t)$ were developed in the

model for assessment.

The detailed descriptions of the parameters were highlighted in Table 3.2.

Epidemic evolution

The influenza A (H1N1) epidemic was seeded according to the start dates (Table 3.3) of the countries [115] [34]. The earliest epidemic was seeded at Mexico on March 11, 2009 [47]. Each country developed its own infected cases which generated from the discretized-time $SEIR$ model based on the estimated reproduction number. At the same time, the countries sent their infected cases to Hong Kong and the local epidemic evolution was initiated by the successive imported cases via air, sea, and land traffic. The first passage times (FPT) and first one hundred passages times (FHPT) were calculated for different restriction strategies.

Baseline scenario

Since the Hong Kong government confirmed the first imported case of influenza A (H1N1) on May 1, 2009 [91], the initial numbers of latent cases $E_i(0)$ and infectious cases $I_i(0)$ were iteratively estimated, and this minimized the difference between the reported date and the simulated first passage time (FPT). Adapting the stochastic nature, the baseline transmission rate, β was fitted for the first two months after the day of the first local onset case without any travel restrictions and intervention. The local daily surveillance of confirmed infected cases, shown in Figure 3.3, was from the press releases on human swine flu, published by the Department of Health, Hong Kong [92]. Optimum parameter was chosen which had average minimum relative mean square error between empirical and estimated cumulative incidence by Monte Carlo simulation. The reproduction number was the product of the transmission rate and the average infectious period. We adopted the range of parameter space for the reproduction numbers according to previous influenza A (H1N1) studies

Table 3.2: Parameters, definitions, and values for the model

Parameter	Definition	Value	Ref/remarks
R_0	Basic reproductive number	Estimated	Local baseline estimated about 1.3; test in range from 1.1 to 1.7
$T_E = 1/\alpha$	Average latent period (days)	1.45	[44][16]; test with 0.5 and 2 days
T_I	Average infectious period (days)	2.9	[44][16]
T_T	Average infectious period (days) for individuals treated with antiviral treatment	1.4	[7]
T_H	Average infectious period (days) for hospitalized individuals	1.4	[7]
p_T	Proportions of infectious subjects selected for treatment	0.12	[49]
p_H	Proportions of infectious subjects selected for hospitalization	0.06	[119]
p_U	Proportions of untreated infectious subjects	$1 - p_T - p_H$=0.82	
γ_R	Transition rates from infectious state to removed state	$1/T_I$=0.34	
γ_T	Transition rates from treatment state to removed state	$1/(T_T - 1)$=2	
γ_H	Transition rates from hospitalization state to removed state	$1/(T_H - 1)$=2	
f_k	Restriction fraction for k-th transportation	90%, 99%	Assumption
ψ	Fraction of infectiousness reduction for antiviral treatment	60%	[40]
ν	Sensitivity of the screening board for infectious subjects	0.3	[28]; test with 0.95 and 0.05

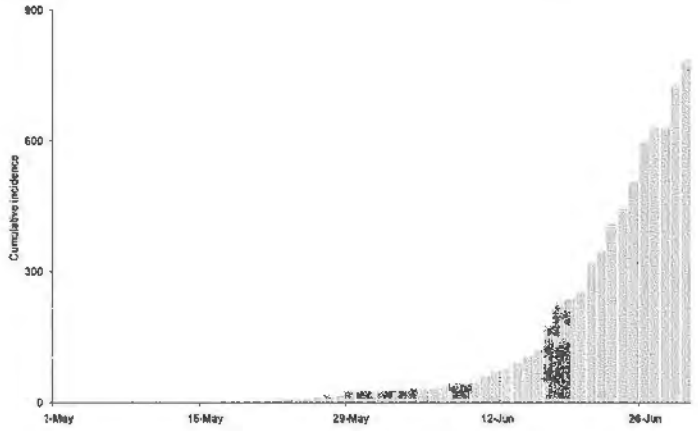
Table 3.3: Start date of epidemic (2009) and estimated R_0 (CI)

Country	Start date MM/DD	R_0 (CI)
United States	04/21	1.62 (1.52, 1.72)
Canada	04/28	1.42 (1.38, 1.47)
Honduras	05/23	1.39 (1.31, 1.48)
Mexico	03/11	1.56 (1.52, 1.59)
Argentina	05/09	1.81 (1.73, 1.90)
Brazil	05/09	1.45 (1.42, 1.49)
Venezuela	05/29	1.43 (1.37, 1.49)
United Kingdom	04/28	1.51 (1.47, 1.54)
Netherlands	05/01	1.42 (1.38, 1.47)
Denmark	05/02	1.37 (1.32, 1.42)
Finland	05/13	1.32 (1.30, 1.35)
Norway	05/11	1.27 (1.26, 1.28)
Sweden	05/07	1.39 (1.37, 1.41)
Austria	04/30	1.24 (1.20, 1.27)
Germany	04/30	1.37 (1.34, 1.39)
Switzerland	05/01	1.38 (1.35, 1.41)
France	05/02	1.33 (1.31, 1.35)
Belgium	05/14	1.24 (1.23, 1.26)
Italy	05/03	1.29 (1.27, 1.31)
Portugal	05/06	1.24 (1.21, 1.27)
Spain	04/28	1.30 (1.25, 1.35)
Russia	05/23	1.06 (1.04, 1.08)
South Africa	06/18	1.69 (1.62, 1.76)
Bahrain	05/27	1.35 (1.31, 1.40)
Egypt	06/03	1.35 (1.30, 1.40)
Israel	04/29	1.42 (1.39, 1.45)
Jordan	06/17	1.26 (1.23, 1.30)
Kuwait	05/25	1.10 (1.09, 1.11)
Saudi Arabia	06/03	1.48 (1.43, 1.54)
Turkey	05/17	1.30 (1.27, 1.32)
United Arab Emirates	05/25	1.30 (1.25, 1.34)
Australia	05/09	1.87 (1.77, 1.98)
New Zealand	04/29	1.35 (1.30, 1.41)
Japan	05/09	1.44 (1.35, 1.53)
South Korea	05/03	1.43 (1.39, 1.46)
Indonesia	06/24	1.69 (1.62, 1.75)
Malaysia	05/16	1.59 (1.54, 1.64)
Philippines	05/22	1.66 (1.60, 1.71)
Singapore	05/27	1.58 (1.53, 1.64)
Thailand	05/14	1.80 (1.71, 1.88)
India	05/17	1.56 (1.51, 1.60)
Taiwan	05/20	1.28 (1.23, 1.32)
Macau	06/19	1.39 (1.34, 1.45)
China	05/12	1.52 (1.50, 1.55)

The epidemic start dates of the countries were from WHO [115] and ECDC [34].

[47] [44] [17] [31].

Figure 3.3: Local cumulative incidence from May 1, 2009 to June 30, 2009.



3.3.3 Computer simulation

The model was implemented in software SAS 9.1.3. Simulation was started from the day of first global onset case with one day time step. The program generated one hundred realizations for each scenerio. The medians, means, and the 95% non-parametric confidence intervals of the incidence and the time of imported case arrivals were calculated over the realizations among different scenarios. The syntax and the functions of the SAS programs were highlighted in Appendix A.

3.3.4 Result

Scenarios with no interventions

Given no intervention in the early epidemic period, the local baseline R_0 for influenza A (H1N1) was estimated to be around 1.4 during the first two months after the reported FPT. The estimated parameters $E_i(0)$ and $I_i(0)$ were equal

to 90 individuals which obtained May 4 as a mean FPT with a 95% confidence interval [Apr 14, May 16]. Values of R_0 were adopted for the mild ($R_0 = 1.1$) and the severe ($R_0 = 1.7$) scenarios in Hong Kong, and these were in line with other studies [47, 121].

Showed in Table 3.3, the R_0s were ranged from 1.1 to 1.9 for other countries. All of the estimated initial growth rates were fitted significantly ($p - value < 0.05$). In the baseline scenario ($R_0 = 1.4$), the medians of FPT and first one hundred passage time (FHPT) of infected cases to Hong Kong were the 55th and the 90th day, respectively (Table 3.4). Because the influenza A (H1N1) was initiated in the Americas, the infected cases arrived in Hong Kong at the fourth month by air travel, which was the main means of transport from the Americas (Figure 3.4). The number of imported cases by air transport was more than that by land transport during the first six months. Afterwards, because the emerging virus had circulated to most the Asian countries, including China, the number of imported cases thus increased exponentially by land transport during the seventh month after the first global case onset. Because ships were not the main external means of transport to Hong Kong, they did not deliver a large number of cases during the epidemic period (Figure 3.4).

Given no control measures, the cumulative AR was 4.7% in the first five months, and it exceeded more than 50% after seven months, when the $R_0 = 1.4$ in Hong Kong. The seven months' cumulative AR was close to that of the final AR (Table 3.5). For a mild scenario ($R_0 = 1.1$), there was no more than 2% cumulative AR in the first five months locally, and the seven months' cumulative AR only reached two-thirds of the final cumulative AR (Table 3.6). If the local scenario was severe ($R_0 = 1.7$), the seventh month was near the end of the influenza A (H1N1) epidemic, and the cumulative AR exceeded 70% (Table 3.7).

Table 3.4: Median FPTs and FHPTs (in days) with confidence intervals (CI) at the baseline scenario.

Control measure	Transportation	FPT (95% CI)	FHPT (95% CI)
No travel restriction		55 (35, 67)	90 (89, 92)
90% travel restriction	Air	62 (42, 72)	99 (97, 100)
	Sea	56 (34, 67)	92 (90, 93)
	Land	58 (44, 69)	93 (91, 95)
	Air, Sea	66 (51, 77)	102 (101, 104)
	Air, Land	69 (45, 81)	106 (104, 107)
	Sea, Land	58 (30, 69)	95 (93, 96)
	All transports	94 (88, 98)	114 (114, 115)
99% travel restriction	Air	61 (37, 72)	99 (97,101)
	Sea	57 (28, 68)	92 (90, 94)
	Land	59 (38, 69)	93 (92, 95)
	Air, Sea	65 (39, 78)	104 (101, 105)
	Air, Land	68 (49, 82)	107 (108, 110)
	Sea, Land	59 (34, 70)	95 (93, 96)
	All transports	117 (116, 118)	138 (138, 139)

Travel restrictions took effect on the day after the first global case onset. The medians and the non-parametric 95% confidence intervals were obtained from 100 simulation runs.

Table 3.5: Median cumulative ARs (in %) with confidence intervals (CI) for different control measures without AH at the baseline scenario.

Control measure	Transportation	No antiviral and hospitalization		
		5 months	7 months	End of epidemic
No travel restriction		4.7 (4.3, 5.2)	54.1 (53.9, 54.3)	57.8 (57.6, 57.9)
90% travel restriction	Air	2.0 (2.0, 2.5)	49.1 (48.7, 49.8)	56.2 (56.1, 56.3)
	Sea	4.1 (3.7, 4.6)	52.9 (52.7, 53.2)	57.1 (57.0, 57.2)
	Land	3.7 (3.3, 4.1)	50.5 (50.2, 50.9)	55.0 (54.9, 55.1)
	Air, Sea	1.4 (1.2, 1.8)	45.9 (45.3, 47.0)	55.5 (55.4, 55.6)
	Air, Land	1.0 (0.9, 1.2)	39.5 (38.5, 40.6)	53.2 (53.0, 53.3)
	Sea, Land	3.0 (2.8, 3.5)	48.8 (48.5, 49.2)	54.3 (54.2, 54.4)
	All transports	0.3 (0.3, 0.4)	25.9 (24.6, 27.2)	52.3 (52.2, 52.5)
99% travel restriction	Air	1.8 (1.6, 2.2)	48.4 (47.8, 49.0)	56.0 (55.9, 56.2)
	Sea	4.0 (3.7, 4.9)	52.8 (52.6, 53.2)	57.0 (57.0, 57.2)
	Land	3.6 (3.3, 4.8)	50.1 (49.8, 50.8)	54.7 (54.6, 54.9)
	Air, Sea	1.13 (1.0, 1.4)	44.0 (43.1, 45.1)	55.2 (55.1, 55.4)
	Air, Land	0.7 (0.6, 1.1)	34.7 (33.2, 37.6)	52.6 (52.5, 52.8)
	Sea, Land	2.9 (2.7, 3.4)	48.1 (47.7, 48.7)	53.9 (53.7, 54.0)
	All transports	0.0 (0.0, 0.0)	2.4 (1.9, 2.9)	51.7 (51.5, 51.8)

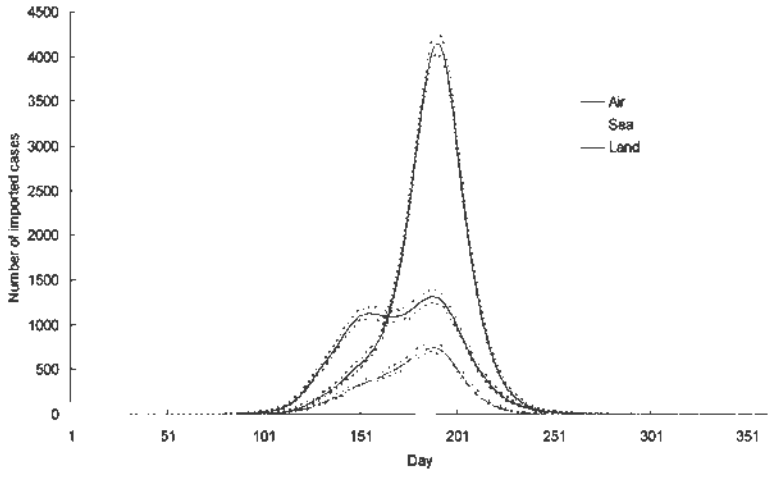
Travel restrictions took effect on the day after the first global case onset. The medians and the 95% non-parametric confidence intervals of each scenario were obtained from 100 simulation runs.

Table 3.6: Median cumulative ARs (in %) with confidence intervals (CI) for different control measures without AH at the mild scenario.

Control measure	Transportation	No antiviral and hospitalization		
		5 months	7 months	End of epidemic
No travel restriction		1.6 (1.5, 1.7)	23.5 (23.3, 23.7)	33.8 (33.6, 33.9)
90% travel restriction	Air	0.7 (0.7, 0.8)	17.1 (16.9, 17.3)	30.7 (30.5, 30.9)
	Sea	1.4 (1.3, 1.4)	21.5 (21.2, 21.7)	32.5 (32.4, 32.7)
	Land	1.3 (1.2, 1.3)	17.0 (16.8, 17.4)	28.2 (28.0, 28.5)
	Air, Sea	0.5 (0.5, 0.5)	14.1 (13.9, 14.3)	29.3 (29.1, 29.5)
	Air, Land	0.4 (0.3, 0.4)	7.9 (7.7, 8.1)	23.6 (23.5, 23.9)
	Sea, Land	1.0 (1.0, 1.1)	14.3 (14.0, 14.6)	26.5 (26.2, 26.8)
	All transports	0.1 (0.1, 0.2)	3.6 (3.4, 3.8)	21.1 (20.7, 21.4)
99% travel restriction	Air	0.6 (0.6, 0.7)	16.3 (16.1, 16.5)	30.4 (30.2, 30.6)
	Sea	1.4 (1.3, 1.4)	21.2 (21.0, 21.6)	32.4 (32.2, 32.6)
	Land	1.2 (1.2, 1.3)	16.3 (16.1, 16.6)	27.6 (27.4, 27.8)
	Air, Sea	0.4 (0.4, 0.4)	12.9 (12.6, 13.1)	28.8 (28.6, 29.0)
	Air, Land	0.2 (0.2, 0.3)	5.6 (5.5, 5.9)	22.1 (21.8, 22.4)
	Sea, Land	1.0 (0.9, 1.0)	13.2 (12.9, 13.5)	25.5 (25.3, 25.8)
	All transports	0.0 (0.0, 0.0)	0.3 (0.3, 0.3)	19.5 (19.0, 19.9)

Travel restrictions took effect on the day after the first global case onset. The medians and the 95% non-parametric confidence intervals of each scenario were obtained from 100 simulation runs.

Figure 3.4: Number of imported cases to Hong Kong by different transports vs. days with no travel restriction.

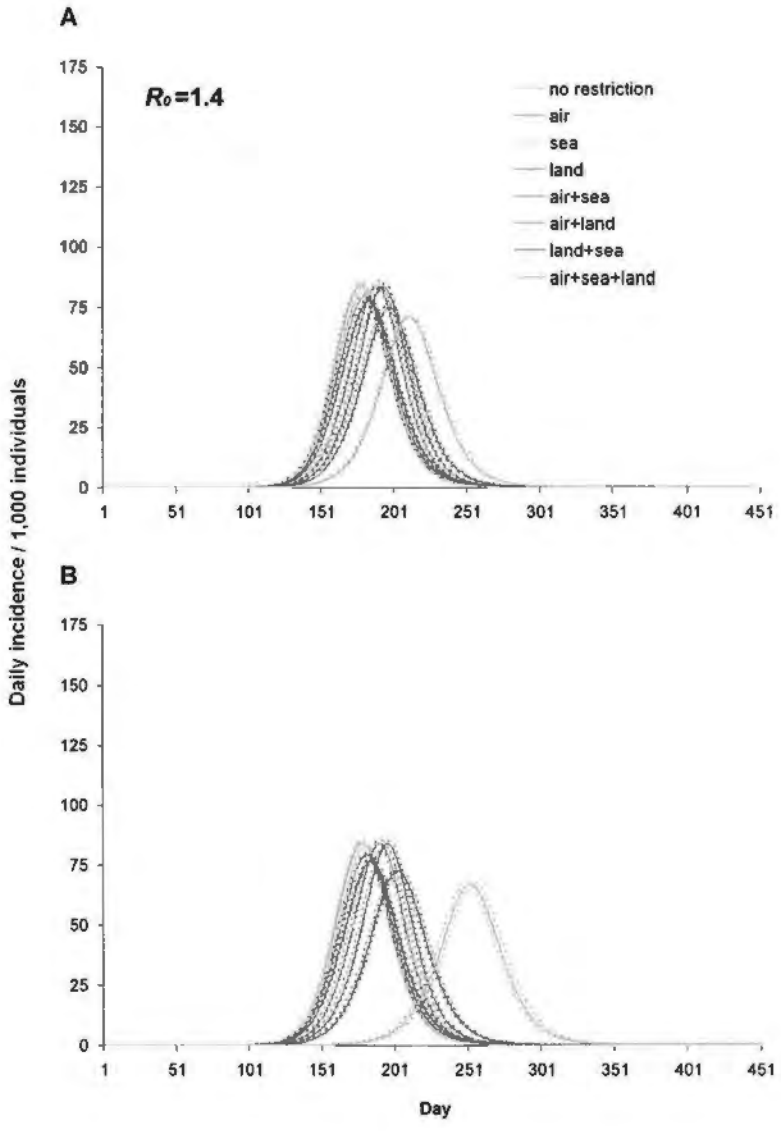


Day one was taken to be March 11, 2009 (the time of the first global case onset). The solid lines represent the average cases; the dotted lines represent the corresponding lower and upper bounds of the 95% non-parametric confidence intervals.

Impact of the interventions

Table 3.4 shows that travel restrictions worked well for slowing down the local spread of the influenza A (H1N1) epidemic in Hong Kong at the baseline fitted scenario ($R_0 = 1.4$). Excepting all air, sea, and land travel restrictions, there were no big differences for FPTs and FHPTs between 90% and 99% restrictions of one or two kinds of transport (Table 3.4). Among the three kinds of single transport restrictions, air travel restrictions worked best in slowing down the FPT and FPHT; they delayed the passage times for one more week than when no travel restrictions were used. The FPT and FHPT could have an additional one to two weeks' delay when both air and land transports were restricted. Once the volume of all transports was reduced 90%, more than one month delay to FPT and FHPT was observed compared to that with no travel

Figure 3.5: Daily incidences vs. days at the baseline scenario ($R_0 = 1.4$) without the uses of the antiviral and hospitalization.



The upper panel (A) and the lower panel (B) illustrate the 90% and the 99% restriction rescaling respectively. Day one was taken to be March 11, 2009 (the time of the first global case onset). The solid lines represent the average cases; the dotted lines represent the corresponding lower and upper bounds of the 95% non-parametric confidence intervals.

Table 3.7: Median cumulative ARs (in %) with confidence intervals (CI) for different control measures without AH at the severe scenario.

Control measure	Transportation	No antiviral and hospitalization		
		5 months	7 months	End of epidemic
No travel restriction		21.6 (16.5, 48.7)	72.2 (72.1, 72,3)	72.9 (72.8, 72.9)
90% travel restriction	Air	9.9 (6.5, 28.2)	71.2 (71.1, 71.4)	72.0 (71.9, 72.0)
	Sea	18.2 (13.9, 33.8)	71.9 (71.8, 72.0)	72.5 (72.4, 72.6)
	Land	16.7 (12.6, 28.5)	70.9 (70.8, 71.0)	71.3 (71.2, 71.4)
	Air, Sea	5.8 (4.2, 11.4)	70.7 (70.5, 70.9)	71.6 (71.5, 71.7)
	Air, Land	4.5 (2.7, 15.5)	69.4 (69.0, 69.9)	70.3 (70.2, 70.4)
	Sea, Land	14.0 (10.2, 28.7)	70.5 (70.4, 70.7)	70.9 (70.8, 71.0)
	All transports	0.9 (0.8, 1.1)	66.3 (65.8, 66.9)	69.8 (69.8, 70.0)
99% travel restriction	Air	8.9 (6.1, 18.8)	71.1 (71.0, 71.3)	71.9 (71.8, 72.0)
	Sea	17.4 (13.6, 25.5)	71.8 (71.7, 71.9)	72.4 (72.4, 72.5)
	Land	16.5 (12.5, 29.2)	70.7 (70.7, 70.9)	71.1 (71.1, 71.2)
	Air, Sea	5.4 (3.4, 11.5)	70.5 (70.2, 70.8)	71.4 (71.3, 71.5)
	Air, Land	3.3 (1.9, 9.3)	68.8 (68.3, 69.6)	70.0 (69.9, 70.1)
	Sea, Land	13.2 (9.8, 22.3)	70.3 (70.2, 70.4)	70.7 (70.6, 70.8)
	All transports	0.0 (0.0, 0.0)	23.9 (18.7, 28.9)	69.5 (69.4, 69.6)

Travel restrictions took effect on the day after the first global case onset. The medians and the 95% non-parametric confidence intervals of each scenario were obtained from 100 simulation runs.

fraction reduction. Moreover, a 99% travel restriction could have additional two months' delay to FPT and FHPT (Table 3.4).

Among the three kinds of transport, the restriction on air travel was effective in controling the five months' cumulative ARs; the ARs kept no more than a half of the one from baseline (Table 3.5). The peak time could have two more weeks' delay if a single 99% air travel restriction had been imposed (Figure 3.5B). Once the land travel was also blocked in either 90% or 99%, seven months' cumulative ARs would have a 15% to 20% decrease and five months' cumulative ARs could be maintained at around 1% on average. They also deferred the peak time for about 3.5 weeks. Most importantly, both the 90% and the 99% travel restrictions for all means of transport were able to keep no more than 1% of the five months' cumulative ARs. The 90% rescaling of all means of transport could maintain the seven months' cumulative AR as a half of when there were no travel restrictions, and it delayed the peak

time from the 25th week to the 30th week (Figure 3.5A). Compared to when no travel restrictions were implemented, the 99% travel restriction kept the seven months' cumulative AR to below 3% on average, and it also deferred the peak time for 11 weeks (Figure 3.5B). Nevertheless, the blocking of sea or land transport did not confer any large reduction in the five months' and seven months' cumulative ARs. The travel restrictions also showed no large reduction in cumulative ARs at the end of the influenza A (H1N1) epidemic. Even when all transports were 99% rescaled, there was only a 5% cumulative AR drop compared to that when no travel restrictions were implemented (Table 3.5).

With the combined use of antiviral and hospitalization (AH), the travel restrictions made greater impacts on slowing down the ARs increase and deferring the incidence peak time. Even if no external travel restrictions had been implemented, the seven months' and final cumulative ARs had still decreased from 54% to 29% and from 58% to 37%, respectively (Table 3.8). When travel restrictions were implemented, the five months' and the seven months' cumulative ARs were reduced by more than half of those when no intervention was implemented. Although the blockings on a single route did not greatly slow down the cumulative ARs' growths, the restrictions on air travel could delay the peak time for more than three weeks (Figure 3.6A and Figure 3.6B). As shown in Table 3.8, the blocking of air and land travel was one of the more effective ways of slowing down the growth of the cumulative ARs. In addition to the use of AH, a 99% restriction of air and land travel could maintain the seven months' cumulative AR to below 10%. They could also delay the peak time for more than six weeks (Figure 3.6B). Once all the external means of transport were 99% rescaled with the use of AH, the increase in the cumulative ARs was greatly deferred, and no more than 1% of the ARs were shown. Most importantly, it greatly extended the peak time, which occurred about five months after when no interventions were implemented (Figure 3.6B). Apart from the

epidemic delay, the restrictions on all means of transports could reduce the peak incidence for more than half of that when only using AH (Figure 3.6A and Figure 3.6B). The 99% travel restriction on all means of transport was also able to reduce the final cumulative AR to about 14% in addition to the use of AH (Table 3.8).

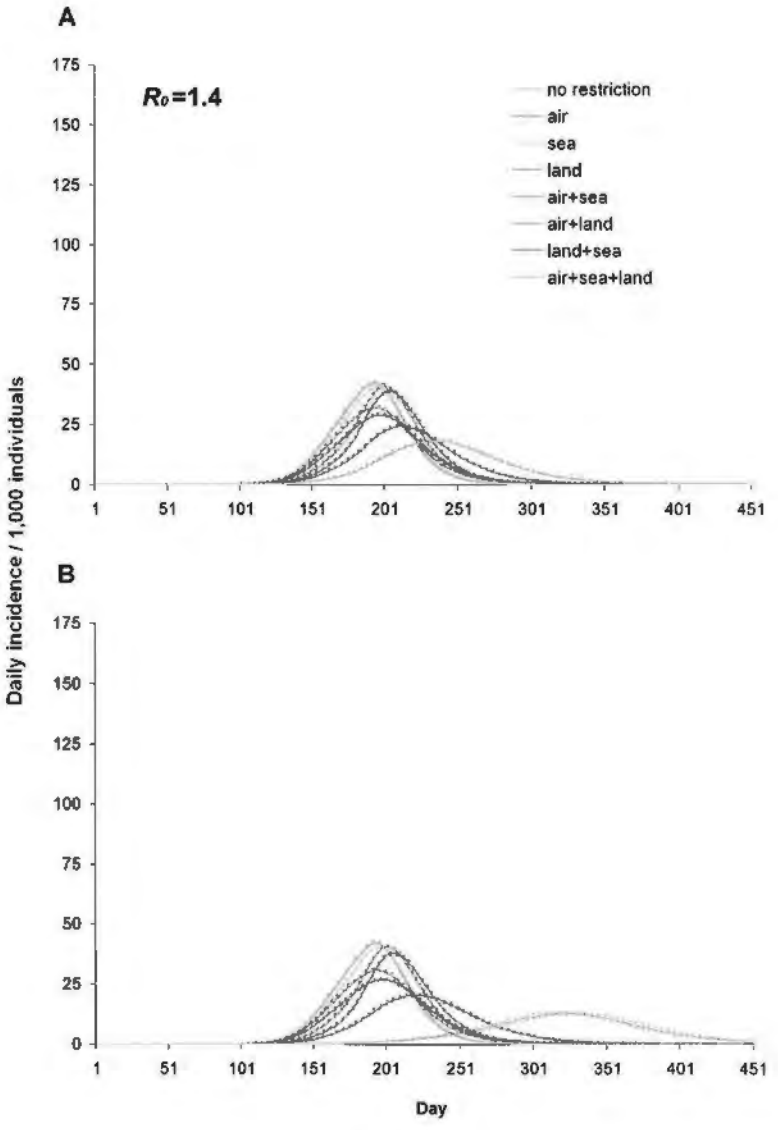
Table 3.8: Median cumulative ARs (in %) with confidence intervals (CI) for different control measures with AH at the baseline scenario.

Control measure	Transportation	antiviral and hospitalization		
		5 months	7 months	End of epidemic
No travel restriction		2.4 (2.3, 2.9)	29.2 (28.9, 29.6)	36.5 (36.3, 36.6)
90% travel restriction	Air	1.0 (1.0, 1.1)	22.5 (22.2, 22.8)	33.6 (33.4, 33.8)
	Sea	2.1 (2.0, 2.6)	27.2 (26.9, 27.8)	35.3 (35.2, 35.5)
	Land	1.9 (1.8, 2.5)	22.8 (22.5, 23.5)	31.3 (31.1, 31.5)
	Air, Sea	0.7 (0.7, 0.8)	19.2 (18.8, 19.5)	32.3 (32.1, 32.5)
	Air, Land	0.5 (0.5, 0.6)	11.9 (11.4, 12.4)	27.1 (26.9, 27,4)
	Sea, Land	1.5 (1.4, 1.7)	20.0 (19.5, 20.3)	29.7 (29.5, 29.9)
	All transports	0.2 (0.2, 0.2)	5.6 (5.4, 6.0)	24.9 (24.6, 25.2)
99% travel restriction	Air	0.9 (0.9, 1.0)	21.7 (21.3, 22.0)	33.3 (33.1, 33.4)
	Sea	2.1 (1.9, 2.3)	27.0 (26.7, 27.4)	35.2 (35.0, 35.4)
	Land	1.8 (1.7, 2.1)	22.0 (21.7, 22.5)	30.7 (30.4, 30.9)
	Air, Sea	0.6 (0.5, 0.7)	17.7 (17.3, 18.1)	31.8 (31.6, 32.0)
	Air, Land	0.4 (0.3, 0.5)	8.8 (8.4, 9.3)	25.7 (25.5, 26.1)
	Sea, Land	1.5 (1.4, 1.7)	18.7 (18.3, 19.2)	28.8 (28.6, 29.0)
	All transports	0.0 (0.0, 0.0)	0.5 (0.4, 0.6)	22.9 (22.6, 23.2)

Travel restrictions took effect on the day after the first global case onset, whereas the antiviral and hospitalization were implemented 3.5 months after the first global case onset. The medians and the 95% non-parametric confidence intervals of each scenario were obtained from 100 simulation runs.

For a milder local scenario ($R_0 = 1.1$), the travel restrictions were more effective on the delay of the influenza A (H1N1) epidemic. The blocking of air travel was still the best choice among the three means of transport for controling the increase of cumulative AR in the first five months of the epidemic (Table 3.6). Because the disease transmissions were comparatively slow and mild, both the 90% and the 99% land import restrictions decreased the peak ARs by one-third. If the air travel was also restricted, the peak time would be deferred by three to four weeks (Figure 3.7A and Figure 3.7B). In addition

Figure 3.6: Daily incidences vs. days at the baseline scenario ($R_0 = 1.4$) with the uses of the antiviral and hospitalization.



The upper panel (A) and the lower panel (B) illustrate the 90% and the 99% restriction re-scaling respectively. Day one was taken to be March 11, 2009 (the time of the first global case onset). The solid lines represent the average cases; the dotted lines represent the corresponding lower and upper bounds of the 95% non-parametric confidence intervals.

to keeping the seven months' cumulative below 1%, the 99% rescaling of all means of transport was able to delay the peak time for a year after the first global case arose (Table 3.6; Figure 3.7B). The 90% all travel restrictions could also delay the peak time for about seven weeks. In the presence of AH, the air, land and all transports restrictions showed an obvious impact on ARs reduction (Figure 3.8A and Figure 3.8B). The 90% and 99% blockings of air and land travel controled the final ARs by no more than 5%. Once all routes of travel were restricted with the use of AH, the spread of the local epidemic was halted; the 99% travel restriction was able to keep the value of final cumulative AR 0.2% on average (Table 3.9).

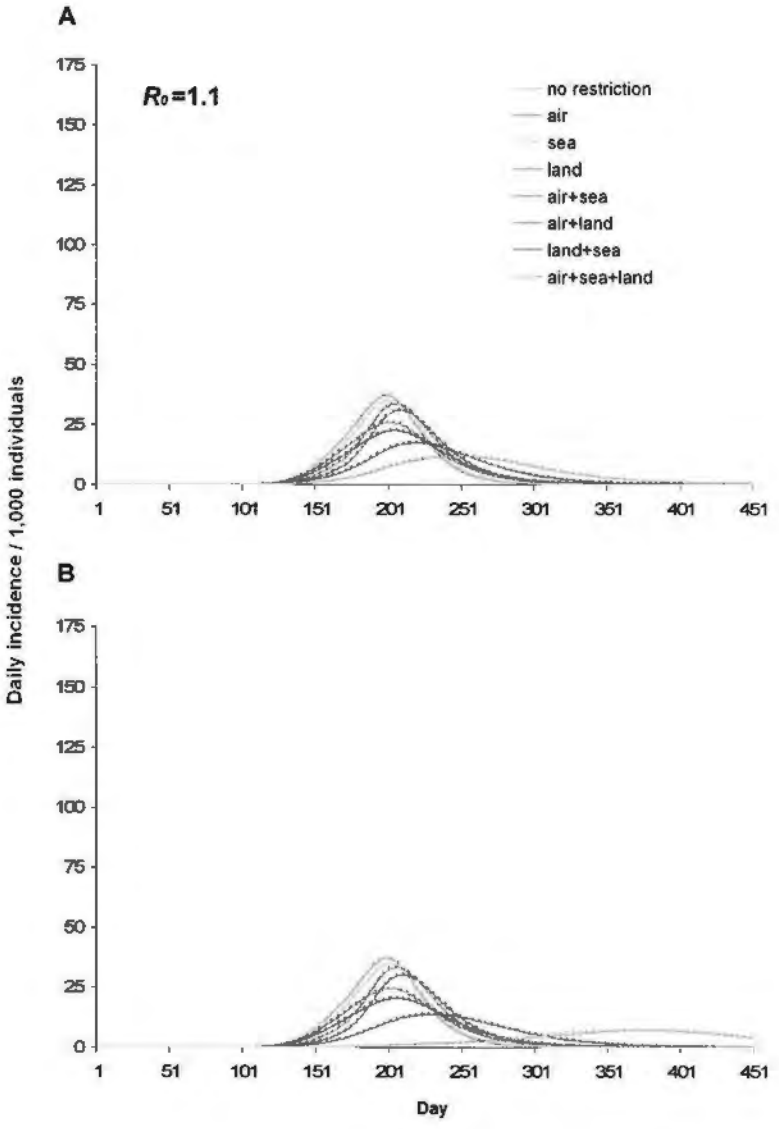
Table 3.9: Median cumulative ARs (in %) with confidence intervals (CI) for different control measures with AH at the mild scenario.

Control measure	Transportation	antiviral and hospitalization		
		5 months	7 months	End of epidemic
No travel restriction		1.0 (1.0, 1.0)	11.5 (11.4, 11.6)	15.9 (15.8, 16.0)
90% travel restriction	Air	0.5 (0.4, 0.5)	8.1 (8.0, 8.2)	12.6 (12.4, 12.7)
	Sea	0.9 (0.8, 0.9)	10.2 (10.1, 10.3)	14.6 (14.4, 14.7)
	Land	0.8 (0.8, 0.8)	6.9 (6.8, 7.1)	9.9 (9.7, 10.0)
	Air, Sea	0.3 (0.3, 0.3)	6.7 (6.6, 6.8)	11.0 (10.9, 11.2)
	Air, Land	0.2 (0.2, 0.2)	3.0 (2.9, 3.1)	5.1 (5.0, 5.2)
	Sea, Land	0.6 (0.6, 0.7)	5.4 (5.4, 5.6)	8.0 (7.9, 8.1)
	All transports	0.1 (0.1, 0.1)	1.3 (1.3, 1.4)	2.5 (2.5, 2.6)
99% travel restriction	Air	0.4 (0.4, 0.4)	7.8 (7.7, 7.8)	12.2 (12.1, 12.3)
	Sea	0.9 (0.8, 0.9)	10.0 (10.0, 10.1)	14.4 (14.3, 14.6)
	Land	0.8 (0.7, 0.8)	6.4 (6.4, 6.6)	9.2 (9.0, 9.3)
	Air, Sea	0.2 (0.2, 0.3)	6.1 (6.1, 6.2)	10.4 (10.3, 10.6)
	Air, Land	0.2 (0.1, 0.2)	2.1 (2.0, 2.1)	3.5 (3.4, 3.6)
	Sea, Land	0.6 (0.6, 0.7)	4.8 (4.7, 4.9)	7.0 (6.9, 7.2)
	All transports	0.0 (0.0, 0.0)	0.1 (0.1, 0.1)	0.2 (0.2, 0.3)

Travel restrictions took effect on the day after the first global case onset, whereas the antiviral and hospitalization were implemented 3.5 months after the first global case onset. The medians and the 95% non-parametric confidence intervals of each scenario were obtained from 100 simulation runs.

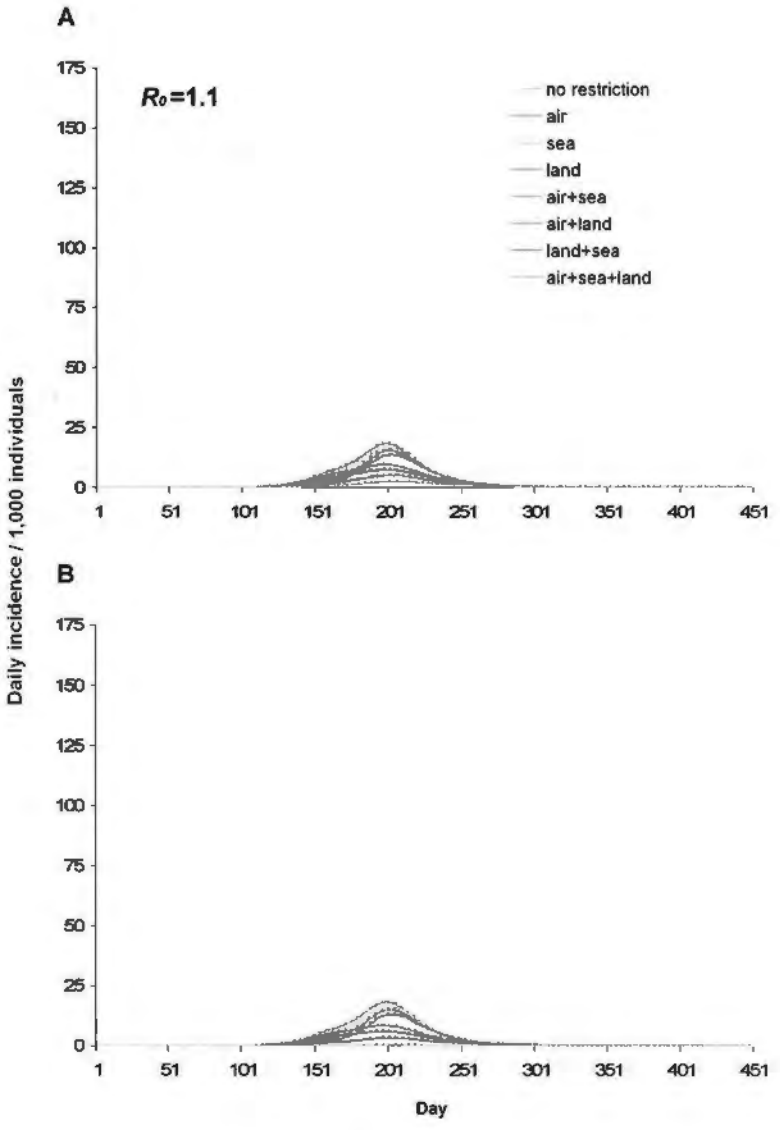
However, the import travel restrictions became less effective in the case of the severe scenario ($R_0 = 1.7$), especially for the ARs' reduction. Compared to 1.6% of the five months' cumulative AR in the mild scenario, the five months'

Figure 3.7: Daily incidences vs. days at the mild scenario ($R_0 = 1.1$) without the uses of the antiviral and hospitalization.



The upper panel (A) and the lower panel (B) illustrate the 90% and the 99% restriction re-scaling respectively. Day one was taken to be March 11, 2009 (the time of the first global case onset). The solid lines represent the average cases; the dotted lines represent the corresponding lower and upper bounds of the 95% non-parametric confidence intervals.

Figure 3.8: Daily incidences vs. days at the mild scenario ($R_0 = 1.1$) with the uses of the antiviral and hospitalization.



The upper panel (A) and the lower panel (B) illustrate the 90% and the 99% restriction re-scaling respectively. Day one was taken to be March 11, 2009 (the time of the first global case onset). The solid lines represent the average cases; the dotted lines represent the corresponding lower and upper bounds of the 95% non-parametric confidence intervals.

cumulative AR attained 22% on average with a large variation which could be up to 49% in the severe scenario (Table 3.7). Imposing restrictions on air travel reduced the five months' cumulative AR by more than 50%, and it delayed the peak time for more than one week (Figure 3.9A and Figure 3.9B). In the severe scenario, the 99% restrictions on all means of transport was still able to halt the local spread during the first five months, and it kept the cumulative AR to about one-third of that without intervention use; it also deferred the epidemic peak time for about eight weeks. Nevertheless, the travel restrictions did not greatly contribute to the decrease in the final cumulative ARs and peak ARs; this was because of rapid disease transmission (Table 3.7; Figure 3.9A and Figure 3.9B). The use of the antiviral and hospitalization became more important in this scenario; the five months' cumulative AR dropped to 9% even when there were no travel restrictions. Because the incidence growth was suppressed by the use of AH, the travel restrictions worked better in slowing down the epidemic. The 90% rescaling of all means of transport reduced the seven months' cumulative AR from 72% to 32% (Table 3.10), and it delayed the epidemic peak for a further seven weeks (Figure 3.10A). When a 99% restriction was imposed on all means of transport, the seven months' cumulative AR would be kept to no more than 4% on average, and the peak time would be delayed for about 12 weeks (Figure 3.10B).

Effect of R_0 from non-local countries

We adopted 20% increases and 20% decreases to the R_0s from a total of 44 non-local countries, in order to test these effects on our results. The R_0s ranged from 1.3 to 2.2 with a median value of 1.7 and ranged from 0.8 to 1.5 with a median value of 1.1 for the 20% increases and the 20% decreases, respectively. Although five countries did not occur any outbreak i.e. $R_0 < 1$, it made small impact on the size of infected cases exportation among all countries.

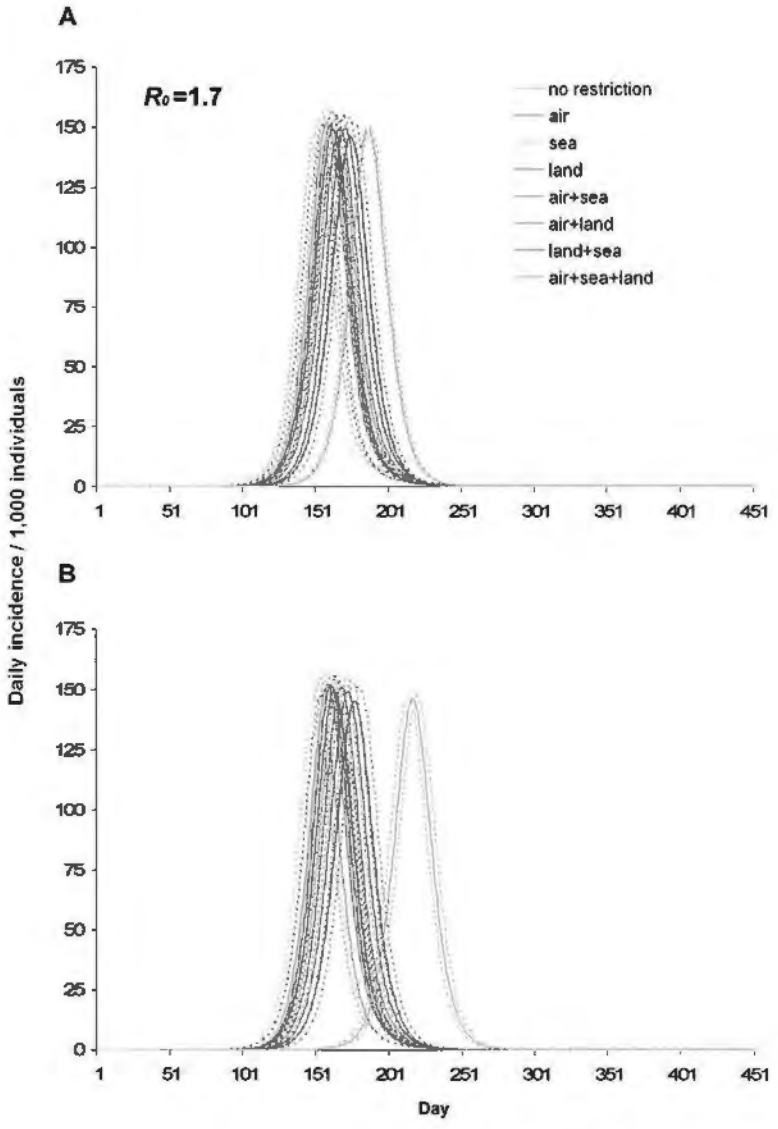
Showed in Figure 3.11, the external travel restrictions performed slightly

Table 3.10: Median cumulative ARs (in %) with confidence intervals (CI) for different control measures with AH at the severe scenario.

Control measure	Transportation	antiviral and hospitalization		
		5 months	7 months	End of epidemic
No travel restriction		9.0 (7.2, 12.9)	54.0 (53.7, 54.3)	55.9 (55.8, 56.0)
90% travel restriction	Air	4.2 (3.0, 13.1)	51.0 (50.5, 52.5)	54.2 (54.1, 54.4)
	Sea	7.9 (6.3, 17.6)	53.1 (52.9, 53.8)	55.2 (55.1, 55.5)
	Land	7.5 (5.6, 22.1)	51.1 (50.8, 52.8)	53.0 (52.8, 54.0)
	Air, Sea	2.4 (1.8, 4.5)	48.8 (48.1, 50.3)	53.5 (53.3, 53.6)
	Air, Land	2.0 (1.2, 9.2)	44.4 (42.9, 48.0)	51.0 (50.8, 51.1)
	Sea, Land	6.0 (4.5, 17.9)	49.9 (49.4, 51.3)	52.2 (52.0, 52.6)
	All transports	0.4 (0.4, 0.5)	32.2 (31.1, 33.6)	50.0 (49.9, 50.2)
99% travel restriction	Air	3.8 (2.6, 13.3)	50.5 (49.9, 51.8)	54.0 (53.9, 54.3)
	Sea	8.1 (6.0, 18.7)	53.1 (52.7, 53.7)	55.0 (55.0, 55.4)
	Land	6.8 (5.3, 11.3)	50.7 (50.4, 51.3)	52.6 (52.5, 52.8)
	Air, Sea	2.1 (1.4, 5.1)	47.8 (46.7, 50.1)	53.2 (53.1, 53.3)
	Air, Land	1.4 (0.9, 3.4)	41.6 (39.5, 45.9)	50.4 (50.2, 50.5)
	Sea, Land	5.9 (4.2, 13.0)	49.5 (48.9, 50.6)	51.7 (51.6, 51.9)
	All transports	0.0 (0.0, 0.0)	3.7 (3.0, 4.9)	49.3 (49.2, 49.3)

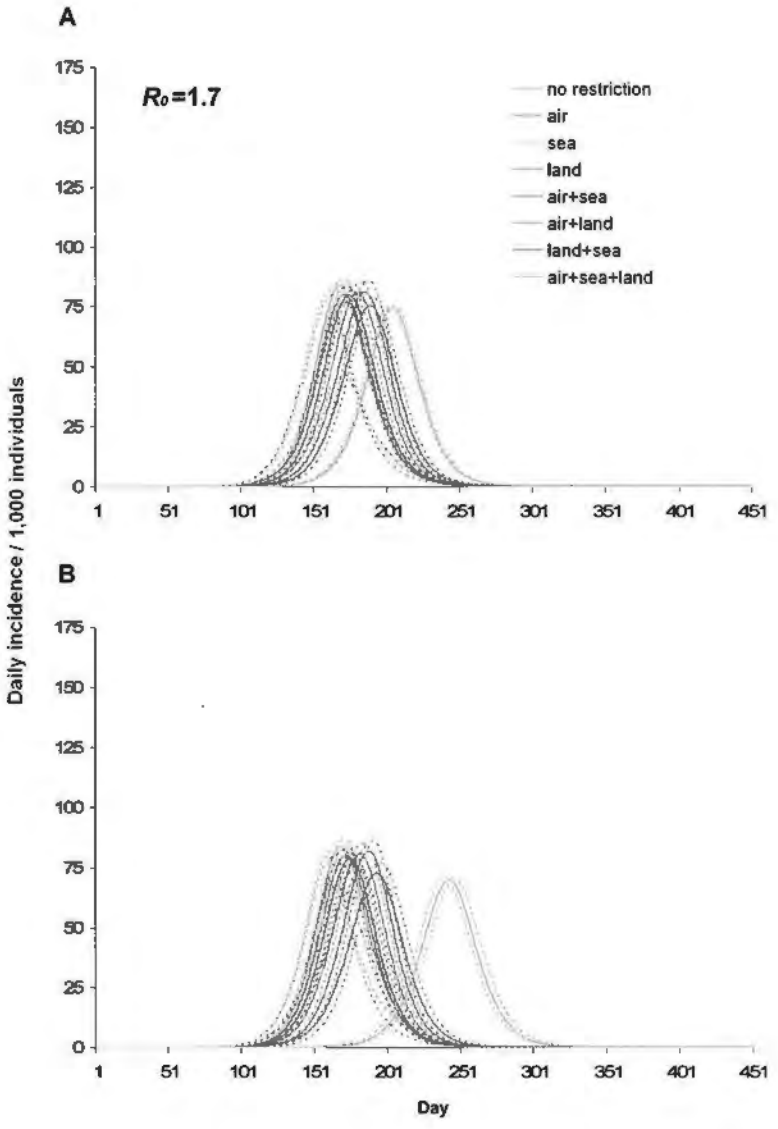
Travel restrictions took effect on the day after the first global case onset, whereas the antiviral and hospitalization were implemented 3.5 months after the first global case onset. The medians and the 95% non-parametric confidence intervals of each scenario were obtained from 100 simulation runs.

Figure 3.9: Daily incidences vs. days at the severe scenario ($R_0 = 1.7$) without the uses of the antiviral and hospitalization.



The upper panel (A) and the lower panel (B) illustrate the 90% and the 99% restriction re-scaling respectively. Day one was taken to be March 11, 2009 (the time of the first global case onset). The solid lines represent the average cases; the dotted lines represent the corresponding lower and upper bounds of the 95% non-parametric confidence intervals.

Figure 3.10: Daily incidences vs. days at the severe scenario ($R_0 = 1.7$) with the uses of the antiviral and hospitalization.

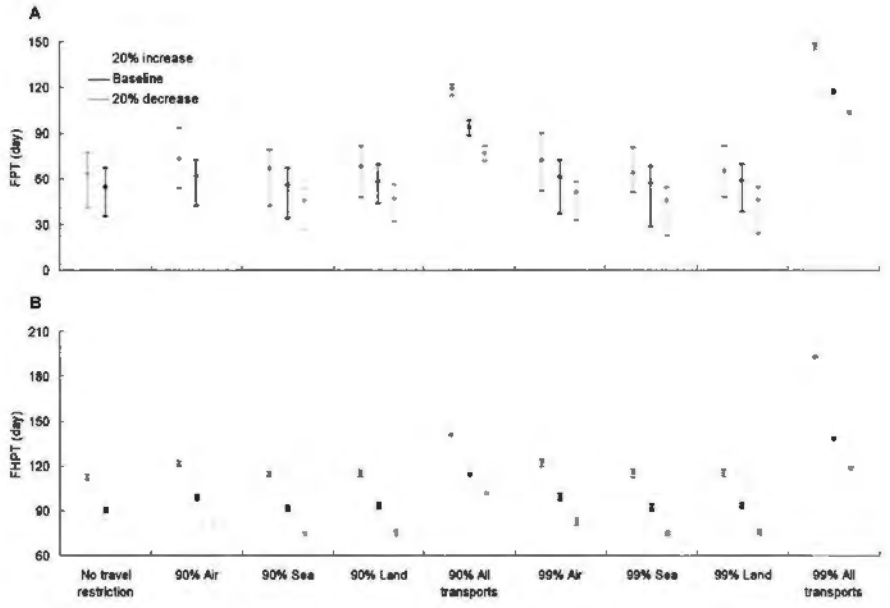


The upper panel (A) and the lower panel (B) illustrate the 90% and the 99% restriction re-scaling respectively. Day one was taken to be March 11, 2009 (the time of the first global case onset). The solid lines represent the average cases; the dotted lines represent the corresponding lower and upper bounds of the 95% non-parametric confidence intervals.

better in deferring the FPTs and the FHPTs when the R_0s from non-local countries decreased. Given the R_0s increased by 20%, the medians of FPT and the FHPT were day 44-th and day 74-th respectively with no travel restriction; the medians of FPT and the FHPT were day 63-th and day 112-th respectively when the R_0s decreased 20%. Amongst all situations for the changes of the R_0s, either 90% or 99% of air travel rescaling could have about 1 week delay for the FPTs; but once all means of transport were 90% or 99% restricted, the FPT would have one month more delay when the R_0s decreased 20% compared to that of the R_0s with 20% increases. Moreover, the FHPT could be delayed for more than 2.5 months with 20% decreases of the R_0s, whereas the FHPT was delayed for 1.5 months with 20% increases of the R_0s for a 99% restriction of all means of transport.

Since the number of imported cases depended on the changes of the R_0 from the non-local countries, the growth of the local epidemic was affected by the cases passage times (Figure 3.12). When the R_0s increased by 20%, the five months' cumulative AR attained 19% and the epidemic ended at the seventh month since the first global case arose. During the first five months, the blockings of all external means of transport were still effective on controlling the cumulative ARs. A 99% travel restriction maintained about 12% of seven months' cumulative AR (Figure 3.12A). Similar to the baseline scenario, the travel restriction made greater impacts on slowing down the ARs increase with the use of antiviral and hospitalization (AH); a 99% rescaling of means of all transport controlled the final AR at about 20% in addition to the use of AH (Figure 3.12B). When the R_0s decreased by 20%, the travel restrictions performed better in slowing down the disease transmission. Even if only the air travel was either 90% or 99% restricted, the seven months' cumulative ARs would have reduced about 15% compared to that of no intervention (Figure 3.12C). A 99% restriction of all means of transport would have halted the local spread i.e. cumulative ARs $< 0.1\%$ in seven months' time whether or not the

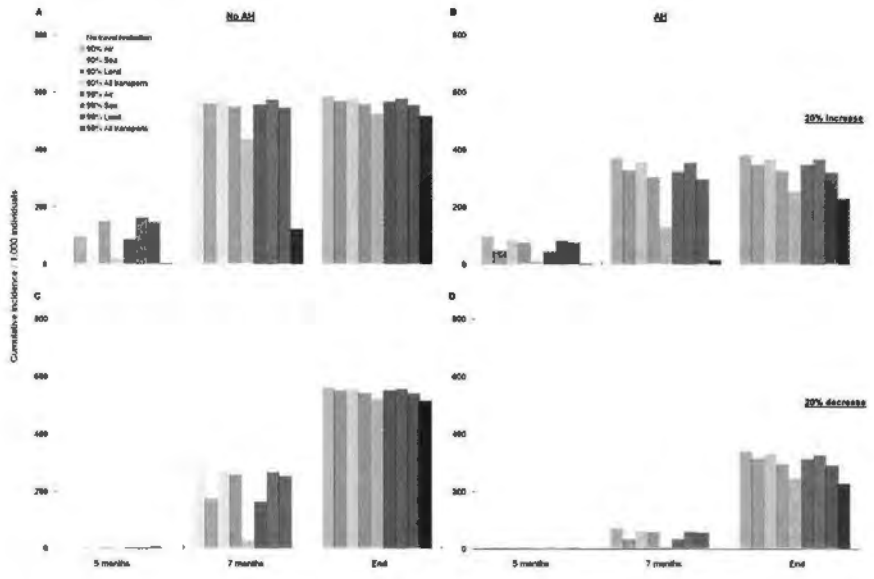
Figure 3.11: FPT and FHPT when non-local countries R_0 increased by 20% or decreased by 20%.



The upper panel (A) and the lower panel (B) illustrate the FPT and the FHPT respectively. Day one was taken to be March 11, 2009 (the time of the first global case onset). The medians are demonstrated as the dots in the interpolations; the corresponding lower and upper bounds of the 95% non-parametric confidence intervals are demonstrated as the lower cups and upper cups respectively.

AH had been used (Figure 3.12C and 3.12D). However, the final cumulative ARs would not be affected by the changes of the R_0s from non-local countries.

Figure 3.12: Median cumulative ARs on different time points when non-local countries R_0 increased by 20% or decreased by 20%.



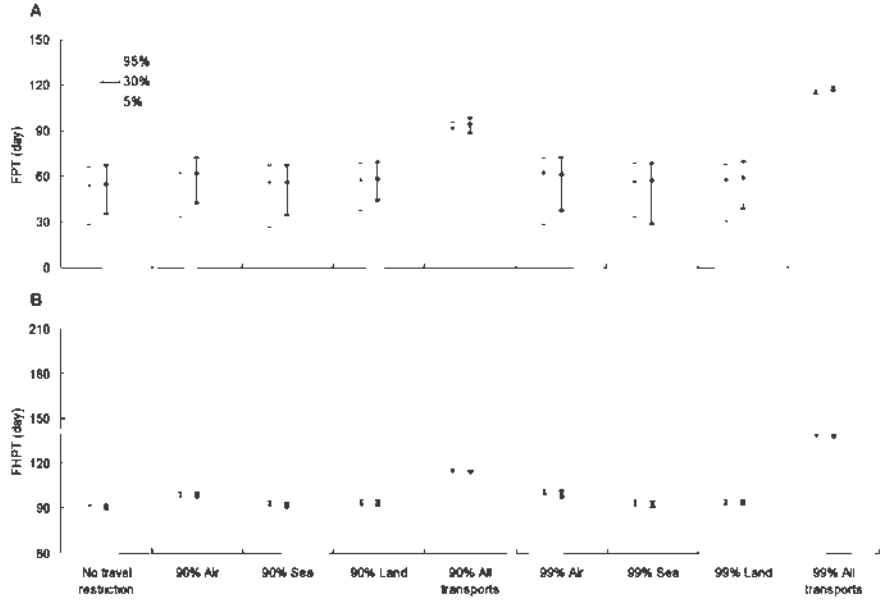
The upper panel (A and B) and the lower panel (C and D) show the cumulative ARs with the non-local countries' R_0s increased by 20% and decreased by 20% respectively. The absences and the presences of the uses of the antiviral and hospitalization are illustrated in left column (A and C) and right column (B and D) respectively. The baseline scenario ($R_0 = 1.4$) was adopted.

Effect of screening sensitivity at entry border points

In the baseline scenarios, we set 30% as the value of the screening sensitivity at entry border points; the value of the screening sensitivity was tested as high as 95% and as low as 5% in our results. According to Figure 3.13 and 3.14, the screening sensitivities at entry border points affected slightly on the times of cases arrival. Amongst most of the travel restriction strategies, a 95%

screening sensitivity showed at most only a one to two weeks additional delay to the FHPTs compared to that of a 5% screening sensitivity (Figure 3.13).

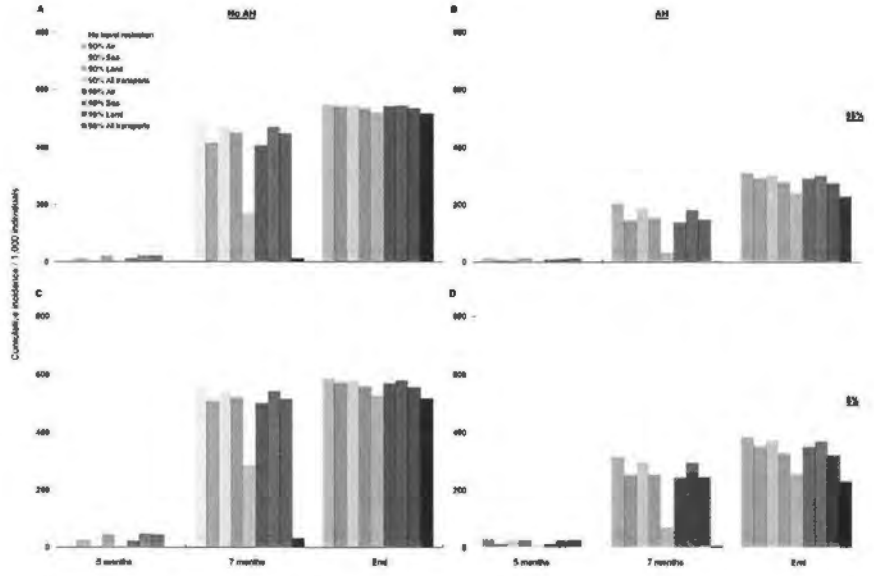
Figure 3.13: FPT and FHPT when screening sensitivity increased to 95% or decreased to 5%.



The upper panel (A) and the lower panel (B) illustrate the FPT and the FHPT respectively. Day one was taken to be March 11, 2009 (the time of the first global case onset). The medians are demonstrated as the dots in the interpolations; the corresponding lower and upper bounds of the 95% non-parametric confidence intervals are demonstrated as the lower cups and upper cups respectively.

The increase of the screening sensitivity at entry border points offered a moderate benefit on slowing down the growths of cumulative ARs. Showed in Figure 3.14A-D, a 95% screening sensitivity showed only half of five months' cumulative ARs compared to that of a 5% screening sensitivity. The 95% screening sensitivity also decreased the seven months' cumulative ARs by about 10% in most of the restriction strategies whether or not the AH had been imposed.

Figure 3.14: Median cumulative ARs on different time points when screening sensitivity increased to 95% or decreased to 5%.



The upper panel (A and B) and the lower panel (C and D) show the cumulative ARs with the screening sensitivities increased to 95% and decreased to 5% respectively. The absences and the presences of the uses of the antiviral and hospitalization are illustrated in left column (A and C) and right column (B and D) respectively. The baseline scenario ($R_0 = 1.4$) was adopted.

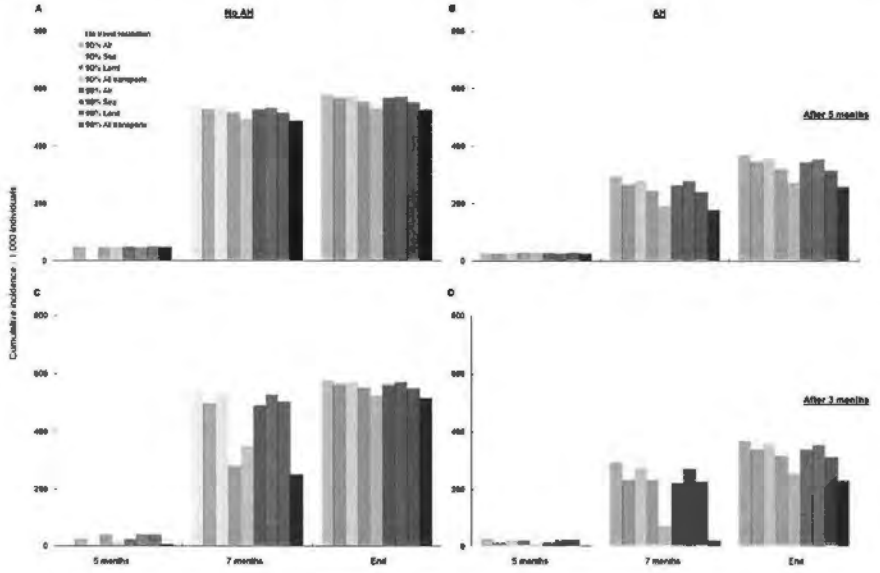
Effect of implementation date on travel restrictions

In the simulation, we started the implementations of travel restrictions at the day after the first global case arose. We also tested the impacts on our results by pushing the implementation date of the travel restrictions to either five months or three months after the first global case arose. Showed in Figure 3.15A and 3.15B, imposing travel restrictions five months after the first global case arose would be too late obviously. Even if all means of transport had been 99% rescaled, the reduction in the cumulative AR was too small. However, it could still decrease the seven months' cumulative AR by no more than 10% if the growth of the epidemic was slowed down by the use of AH. Showed in Figure 3.15C, imposing the travel restrictions three months after the first global case arose would be a little bit late; but fractional blockings on all means of transport worked well in deferring the growth of the ARs. The 99% restriction would reduce the five months' and seven months' cumulative ARs more than half of that without intervention. With the use of AH, imposing the 99% restriction of all mean of transport was able to control the cumulative AR by no more than 2% in the first seven months; a 90% restriction could still maintain the average seven months' cumulative AR about 6% to 7% (Figure 3.15D).

Effect of length of latent period

In the baseline scenario, the length of the latent period was set as 1.45 days, in accordance with the reference value. The impacts on the FPTs and the FHPTs were tested given that the latent period increased to two days or decreased to a half-day. Showed in Figure 3.16, our result was insensitive to the changes on the lengths of the latent period. The difference of the latent period's lengths did not show large variations on the FPT (Figure 3.16A) among all restrictions strategies; even if all means of transport had been 99% blocked, the FHPT

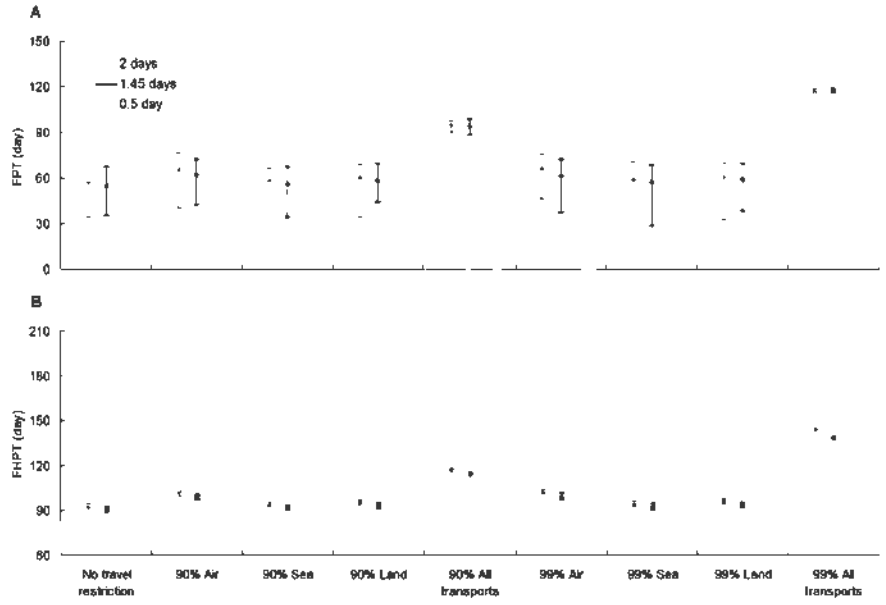
Figure 3.15: Median cumulative ARs on different time points when implementation date of travel restrictions delayed for five months or three months.



The upper panel (A and B) and the lower panel (C and D) show the cumulative ARs with the implementation dates on travel restrictions delayed for five months and three months respectively. The absences and the presences of the uses of the antiviral and hospitalization are illustrated in left column (A and C) and right column (B and D) respectively. The baseline scenario ($R_0 = 1.4$) was adopted.

would only have several days' difference between two days and a half-day in the latent periods (Figure 3.16B). Since the effect of the latent period's length on the local epidemic growth was beyond our scope, only the rate of passages i.e. FPT and FHPT had been investigated.

Figure 3.16: FPT and FHPT when latent period increased to 2 days or decreased to 0.5 day.



The upper panel (A) and the lower panel (B) illustrate the FPT and the FHPT respectively. Day one was taken to be March 11, 2009 (the time of the first global case onset). The medians are demonstrated as the dots in the interpolations; the corresponding lower and upper bounds of the 95% non-parametric confidence intervals are demonstrated as the lower cups and upper cups respectively.

3.4 Discussion

The choice of intervention use is usually an issue both for the public and for the policy-makers during the epidemic period. Previous mathematical modeling

studies demonstrated various impacts on disease control for different interventions [119, 39, 49, 74, 110, 97, 107]. In this paper, a mathematical model was developed to quantify the impact of the travel restrictions on air, sea, and land travel with stochastic uncertainty for human influenza A (H1N1). The infectious disease has spread to more than 214 countries and territories, and it caused almost 20,000 deaths; the large and dense international travel network should be one of the risk factors. Most of Hong Kong's visitors arrived by land transport; however, the efficacy of travel restrictions has been strongly argued [60, 37, 18]. Because of the limited data and the limitations of the methodology used in the epidemic models, it was hard to quantify the impact of travel restrictions other than that on air travel [55, 23]. In our study, the statistics of the number of arrivals in Hong Kong from 44 countries using air, sea, and land transport were collected [62], and they were adopted in a mathematical model to demonstrate the disease dynamic for influenza A (H1N1) and the impacts of the travel restrictions on all modes. The use of antiviral and hospitalization was also incorporated into the model in order to allow a proper comparison of the effectiveness of the transport restrictions.

Deferring the disease spread is important to the pandemic management of the early phase, whereas public health experts, policy-makers and scientists usually require a period of time for decision-making on epidemic control. Once the epidemic is not eliminable in the source country, another effective approach is to delay the disease spread in the at-risk countries. We adopted the influenza A (H1N1) pandemic in 2009 as a case study and the results showed that the greatest impact of travel restrictions was to slow down the spread of the disease. The local baseline reproduction number (R_0) was estimated to be about 1.4. Because the influenza A (H1N1) was initiated in the Americas, the restriction on air travel, which was the main means of transport from the Americas to Hong Kong, was most effective in delaying the time to the FPT and FPHT among the types of single transport restriction. Six months after the first global

case arose, the emerging virus had circulated to most Asian countries, including China, and the number of imported cases therefore increased exponentially by land transport; the five months' cumulative ARs could be maintained around 1% on average and the peak time could be additionally deferred for 3.5 weeks when both air and land transport were thus restricted. Expressed simply, the 99% restriction of all means of transport was the most efficient strategy for defering the local epidemic. It kept the cumulative AR below 3% during the first seven months and also it deferred the peak time for 11 weeks. With the use of antiviral and hospitalization (AH), the travel restrictions were more successful in deferring the growth of the ARs and the incidence peak time. The local epidemic was halted during the first seven months and the peak time was delayed for an additional five months once all external means of transport were 99% rescaled with the use of AH. Most importantly, the restrictions on all means of transport decreased the peak incidence by more than half of that when using only AH.

The travel restrictions worked better in the mild scenario ($R_0 = 1.1$), but they performed less well in the severe scenario ($R_0 = 1.7$). When $R_0 = 1.1$, the 99% rescaling of all means of transport also greatly delayed the peak time for a year. It finally halted the spread of the local epidemic with the use of AH. When $R_0 = 1.7$, the 99% restriction on all means of transport was still capable of halting the local spread during the first five months. However, the disease spread at a higher rate locally and the local infectious cases transmitted the H1N1 virus to others successively far more than the imported cases did. Therefore, the travel restrictions did not greatly contribute to the decrease in the peak ARs. In line with previous findings [60], the transport network only had a major role when the infected case numbers were low globally. The use of AH became more important in the severe scenario because it could suppress the incidence growth in the epidemic; the travel restrictions would still be effective in the scenario. A 99% restriction on all means of transport with the

use of AH could still keep the seven months' cumulative AR at no more than 4%, and the peak time would be delayed for about 12 weeks.

Apart from taking control measures locally, the effectiveness of the travel restrictions increased with the reduction of the R_0s from a total of 44 non-local countries. Because the number of imported cases depended on the changes of the R_0s, the growth of the local epidemic was greatly affected by the successive disease transmission from the cases. If control measures had taken effect in those non-local countries that decreased the R_0s by 20% on average, a 99% restriction on all external means of transport would possibly have halted the local spread, i.e., the cumulative ARs $< 0.1\%$ in seven months' time, whether or not the AH had been used. Moreover, increasing the screening sensitivity at the entry border points was beneficial for slowing down the growth of cumulative ARs; a 95% screening sensitivity showed half of the five months' cumulative ARs compared to that of a 5% screening sensitivity amongst most of the travel restriction strategies. Our results also suggested that it would be necessary to impose the restrictions no later than three months after the first global case arose. The implementation of the travel restrictions at the end of the fifth month would be almost useless; this is because the local epidemic would have by then evolved to a mature stage, in which the disease transmission would depend on the local exponentially increased cases rather than on the successive imported cases.

Due to economical, legal, and social consequences, it is hard to rigorously enforce the travel restriction even in a single country. However, the importance and the potential of imposing the travel restriction cannot be neglected. In recent decades, more serious diseases, such as SARS and influenza A (H1N1), successively emerged into our society and affected wider age groups compared to epidemics in the past. It is predictable that a more lethal virus might emerge in the near future. According to our result, the travel restriction is able to reduce the rate of the disease spread and even perform better with

the combination use of other control measures. Therefore, implementation of travel restrictions must be a potential public health control measure to reduce and to delay the community disease spread, especially for the next epidemics which could be highly intrusive and invasive. Moreover, the pharmaceutical interventions, like vaccine and antiviral, are usually not available early enough once a new emerging virus has arrived in the community. The travel restriction is a simple and direct non-pharmaceutical intervention to slow down the epidemic during the early stage, in order to allow a longer period for the preparation of the mitigation response, especially for the emerging virus with unknown characteristics. In an additional cost-effectiveness study (Appendix C), we examined the costs and benefits of imposing travel restrictions before the availability of effective interventions such as antiviral and hospitalization on potential extensions of influenza A (H1N1) virus' transmissibilities and case-fatality rates. According to our results, the cost of a 99% travel restriction of all means of transports would be $11,636 million if it was imposed for 3.5 months before the availability of effective interventions; the travel restriction was cost-effective only if the R_0 increased to 8 and the case-fatality rate $> 15\%$. However, the effect of epidemic delay from the travel restriction reduced a large portion of health care costs for imposing 5 months and 6.5 months before the availability of effective interventions once the disease transmission intensity was comparatively mild with 6% to 15% case-fatality rate; the travel restriction was also cost-effective for a late delivery of treatments when case-fatality rate attained 25%.

Not exactly mandatory, the travel rescaling can be implosed by several ways, such as travel advisories and screenings at border points. Our analysis also demonstrated that increasing the screening sensitivity at the borders would beneficial to delay the passage times of the infected cases as well as to slow down the epidemic growth locally. As described by Hollingsworth *et al.* [60], travel restrictions were better applied to the source country during

the containment phase. They would be effective in minimizing the number of cases being exported, and hence in reducing the successive disease transmission country by country over the global transport network.

In our study, we used a major city, Hong Kong, as a place to demonstrate the effectiveness of travel restrictions. Travel restrictions were likely to show a better illness rate reduction when the local disease transmission intensity was mild. In some rural areas, the disease transmission intensities as well as the reproduction numbers were not too high because of limited human mobility and contacts. In addition, these areas may not receive a large number of travelers from different source regions. So these areas may obtain more benefit from imposing the travel restrictions. As mentioned by Caley *et al.* [19], an additional delay in importing an epidemic was obtainable.

Chapter 4

Impact of Travel Restriction: *Source* Country

In this chapter, a new method was developed to evaluate the possibility of the disease containment by travel restrictions in view of the *source* country. In most of the epidemic modeling studies, surveillance data are important to the parameter configurations in the mathematical models. However, the surveillance data are usually subject to the time delay until the first disease confirmation and also to the non-reporting rate. So in section 4.2, we developed an Markov Chain Monte Carlo (MCMC) method, which imputed the unobserved data, to estimate the model parameters subject to the abovementioned problems. The method was validated by a series of simulations in section 4.3 and was applied to the influenza A (H1N1) outbreak in Mexico which showed in section 4.4. the model was able to demonstrate the probability distribution of exported cases, and thus the possibility of the disease containment by the travel restrictions was assessed.

According to our result, all of the estimates were consistent with other studies. Most importantly, we concluded that only strict restrictions on travelling, i.e., allowing three to 15 travelers exported per day, could have a chance of preventing an at-risk country from importing cases from the source region. If the travel restrictions had been implemented in combination with other

interventions, such as vaccination and antiviral drugs, to reduce the disease transmission locally, the possibility of the containment was enhanced. Besides, early control measures in the source region were crucial to contain the epidemic. In practice, the travel restrictions are not suggested being the first priority of the interventions in view of the source country once a new epidemic is initiated; other intervention, such as isolation and antiviral should be adopted in the community in order to suppress the growth of disease. As long as the incidence is controlled at a low rate, the containment at the source area is possible along with other effective interventions.

4.1 Introduction

In mid-march 2009, the Mexican government identified an unexpected increase in the number of influenza-like illness cases, even though it was not the peak season for influenza outbreaks [77]. After about a fortnight, an acute respiratory illness was discovered in two children and it was confirmed as a new influenza A (H1N1) virus in mid-April 2009 [78]. The first notification of this novel influenza A (H1N1) was announced by the World Health Organization (WHO) on April 26, 2009. Because additional cases were successively discovered in the United States [79], the WHO raised the pandemic alert level to phase five at the end of April. After the first global influenza A (H1N1) case had been confirmed, the containment phase was initiated. Therefore, most countries had taken control measures on border points to screen out the suspected cases, especially for those with traveling from Mexico and the United States, in order to prevent the possibility of local disease transmission from the source country.

With the understanding to the distribution of the exported infections from the source country, researchers are able to assess the possibility of disease containment and to have better preparation for the control measures, like the

border control. However, researchers have to face the problems of the time delay until the first official disease confirmation and the non-reporting rate, while formulating the distribution. As for the influenza A (H1N1) pandemic, it was believed that the virus had been circulating within communities before the disease outbreak was recognized [63, 71]. Before the active surveillance of influenza A (H1N1) and the confirmative diagnosis from clinicians and microbiologists, the virus had been undetected for a period of time. Several studies estimated the initial point of the disease outbreak as being around mid-January to late-February through the analysis of the viral genetic sequence and the epidemic models [47, 106], and the delay would have had a significant impact on the simulation results [19, 40]. Apart from the initial time delay, the reporting rate was low for the influenza A (H1N1) pandemic. Most of the ascertainment was particularly focused on cases with severe condition. Also, asymptomatic or mild cases were not presented at medical consultation. A good example of official surveillance being underestimated in the disease transmission intensity would be the telephone interviews from the Beijing Center for Disease Prevention and Control (CDC) [120], which showed that the consultation rate among influenza-like illness (ILI) patients was no more than 50% in Beijing, China.

Although the issues of initial reporting delay and non-reporting are important, most of the epidemic modeling studies have neglected these factors in model development. Caley, *et al.* [19] developed a complex probability distribution model accounted for the time delay between the start of an epidemic and the subsequent cases exported. But the study did not deal with the non-reporting issue. Hollingsworth, *et al.* [60] constructed an epidemic model to investigate the impact of travel restrictions. However, the probability of exported cases to countries was arbitrarily assumed. Most importantly, no estimations have been carried out on the model parameters for the studies.

In our study, a stochastic mathematical model was constructed to improve

the above limitations. Typically, infectious disease transmission model are either deterministic or stochastic. The deterministic models are easier to build and used to demonstrate the average behavior, but they do not incorporate the stochastic uncertainty into the epidemic systems. Besides, the transmission intensity in the models is determined by the basic reproduction number (R_0), which is defined as the average number of secondary infections produced by a typical infectious individual in a wholly susceptible population. The R_0 was estimated in the model and adopted some characteristics of the following model. First, we incorporated the time of the first exported case (FET) from Mexico to the at-risk countries. With unknown severity and transmissibility for the emerging virus, the at-risk countries were alert to the cases arose; given the information of the travel intensity, the FET would be an indicator to the determination of the disease outbreak stage in the source country. Observed from the data source (Table 4.2), most of the reported cases have traveled to Mexico or have exposed to the groups with Mexico travelling history and the fact gave us insight on the estimation. Second, the model was based on Bayesian Theorem, an inference that capable to incorporate the expert suggestions, multiple data sets, and the reference information when establishing prior distributions given unknown parameters. Incorporation of the experience and the available knowledge regarding the disease characteristic in models is useful for the appropriate management of infectious disease [72, 84, 112]. Besides, the use of Bayesian approach is able to take the uncertainty of parameters into account along with the stochastic variation. The Bayesian approach is also preferred to vary the parameters for a sensitivity analysis.

In this chapter, we developed a mathematical model to estimate the reproduction number subjected to the initial reporting delay and the reporting rate behind the surveillance data. Most importantly, we applied the method to the influenza A (H1N1) outbreak in Mexico, in order to demonstrate the probability distribution of exported cases, and thus studied the possibility of the

disease containment by the travel restriction. The methodology and the case study would be able to offer the insights of disease transmission and a better management in intervention preparation to officials, public health experts, and epidemiologists.

4.2 Methodology

4.2.1 System of Stochastic Disease Transmission Model

We adopted a simple stochastic SIR model to describe the disease dynamic [66, 56, 2]. Let Δt be a time step and $(t, t + \Delta t]$ be a time interval, and the population size, N, is divided into three classes: susceptible $(S(t))$; infectious $(I(t))$; and recovered $(R(t))$, at each time point t. Because we could not confirm whether the imported cases were in latent status or infectious status when they arrived in the countries, we employed an SIR model in convention. In the stochastic model, as soon as the susceptible individuals in compartment $S(t)$ become infected, they will move to the compartment $I(t)$ and stay for the infectious period T_I. The incidence, $X(t)$, follows a binomial distribution with the probability of an individual becoming infected, $p_I(t)$, in the compartment $S(t)$ at time t. Let $bin(m, n)$ be a binomial distribution with parameters probability m and number of total individuals n, the incidence

$$X(t) \sim bin(p_I(t), S(t)) \tag{4.1}$$

and the probability distribution of the incidence,

$$P(X(t + \Delta t) = x | S(t) = s, I(t) = i) = \begin{pmatrix} s \\ x \end{pmatrix} p_I(t)^x \left[1 - p_I(t)\right]^{s-x} \tag{4.2}$$

where s, i, and x are constants. Assume the population is randomly mixed, the probability of an infection

$$p_I(t) = 1 - exp\left[-\left(\frac{R_0}{T_I}\right)\left(\frac{I(t)}{N}\right)\Delta t\right] \tag{4.3}$$

where $0 \leq p_I(t) \leq 1$; R_0 is the basic reproduction number, which is defined as the average number of secondary infections produced by a typical infectious individual in a wholly susceptible population. The R_0/T_I is the transmission coefficient (β) for the infectious disease.

When the infectious period is over, the individuals in compartment $I(t)$ will recover and move to the compartment $R(t)$ with rate $1/T_I$ for one time step. Hence, the number of the removed subjects becomes

$$Y(t) \sim bin(p_R, I(t)) \tag{4.4}$$

where the probability of recovering

$$p_R = 1 - exp\left[-\left(\frac{1}{T_I}\right)\Delta t\right] \tag{4.5}$$

and $0 \leq p_R \leq 1$. In summary, the system of the stochastic epidemic model could be presented by the following equations:

$$
\begin{aligned}
S(t+\Delta t) &= S(t) - X(t) \\
I(t+\Delta t) &= I(t) + X(t) - Y(t) \\
R(t+\Delta t) &= R(t) + Y(t)
\end{aligned}
\tag{4.6}
$$

4.2.2 Probability Distribution of Exported Cases from Source Region

During the period of the epidemic, travelers may carry the virus from the source region to other at-risk countries; and the global pandemic would be likely to occur from successive disease transmissions. The travelers are considered as having an equal probability of being exposed to the disease as are the individuals at the source country [71]. We assume the count of infectious cases includes the travelers, instead of separating the imported cases and local cases [65]; this is because it is hard to know exactly how many travelers there are and for how long they have stayed in the source country. Let M be the daily rate of travel to particular country, the probability distribution of exporting Z infected travelers from the source region on day t would be

$$
Z \sim bin(p_I(t), M) \tag{4.7}
$$

assumed the travelers expose the same daily risk $p_I(t)$ as the resident cases in the source country.

Let m_i be the daily rate of travel to country i-th, the probability of exporting z_i infected travelers from the source region to country i-th on day t would be

$$q_{t,z_i} = P_t(Z = z_i) = \binom{m_i}{z_i} p_I(t)^{z_i}(1 - p_I(t))^{m_i - z_i} \tag{4.8}$$

where $0 \leq q_{t,z_i} \leq 1$; and $q_{t,0}$ is the probability of no exported cases from the source region on day t.

Suppose τ is the number of days from the epidemic initiation to the first global surveillance count and k $(k = 0, 1, 2, 3, ...)$ is the number of days after the first global surveillance count, the probability distribution of the time with the first exported case (FET) from the source region on day $\tau + k$ thus become

$$P(T = \tau + k) = q_{1,0}q_{2,0}q_{3,0}\cdots q_{\tau+k-1,0}(1 - q_{\tau+k,0}) \tag{4.9}$$

which is a geometric distribution.

4.2.3 Reporting Rate

The reporting proportion is incorporated in the model development. Because there is a period of time of non-reporting delay before the confirmation of the emerging virus, the influenza A (H1N1) surveillance count is zero before day τ-th. Suppose an actual incidence $X(t)$ and a constant reporting rate r over time, the observed surveillance, $U(t - \tau)$, is proportional to $rX(t)$ after the day of first surveillance τ. The model of the surveillance time series could be formulated as

$$U(t - \tau) = \begin{cases} 0, & \text{if } t < \tau; \\ rX(t), & \text{if } t \geq \tau. \end{cases} \tag{4.10}$$

We assume that no false positive and no false negative is presented from the laboratory testing.

4.2.4 Statistical Inference

The estimation of parameters is based on the Bayesian approach. Suppose Θ was our parameters of interest, the posterior distribution for the parameters would be

$$\begin{aligned} f(\Theta|data) &= \frac{L(\Theta)\pi(\Theta)}{P(data)} \\ &\propto L(\Theta)\pi(\Theta) \end{aligned} \tag{4.11}$$

where $\pi(\Theta)$ is the prior distribution of parameters, and $L(\Theta)$ is the full likelihood function. The full likelihood function is then constructed as

$$L(\Theta) = \prod_{t=\tau}^{t_*} \binom{S(t)}{U(t-\tau)} p_I(t)^{U(t-\tau)}(1-p_I(t))^{S(t)-U(t-\tau)} \tag{4.12}$$

where t_* is the total number of days from the time series surveillance data. And the corresponding log-likelihood function would be,

$$\begin{aligned} LL(\Theta) = \sum_{t=\tau}^{t_*} \Bigg[& log\binom{S(t)}{U(t-\tau)} + U(t-\tau)log(p_I(t)) \\ & + [S(t)-U(t-\tau)]\, log(1-p_I(t)) \Bigg] \end{aligned} \tag{4.13}$$

The uncertainties of the model parameters are represented by the posterior distribution given the data depending on the prior distributions and the likelihood function.

Our estimates are based on the posterior distributions of the parameters which are obtained by the MCMC method. The MCMC method is able to update the posterior distibution by sampling from the prior distributions. It has been employed in many epidemic modeling studies [53, 35, 88]; this is because of its powerful ability to augment data. In our situation, as the time series $[S(1), S(2), S(3), ..., S(\tau-1)]$ and $[I(1), I(2), I(3), ..., I(\tau-1)]$ are unobserved, the dynamic equation (Eq.4.6) could not be constructed. We simply denote the latent variables $\mathbf{h} = [S(\tau-1), I(\tau-1)]$ rather than the whole series, as the likelihood function base on them only. The marginal likelihood of Θ would be

$$L^m(\Theta) = \int P(data|\Theta, \mathbf{h}) f(\mathbf{h}|\Theta) d\mathbf{h} \tag{4.14}$$

where $P(data|\Theta, \mathbf{h})$ is either a full information or augmented likelihood function and $f(\mathbf{h}|\Theta)$ is a latent states distribution. Direct maximum likelihood (ML) estimation is typically preferred to solve the likelihood because of its strong theoretical properties; however, maximizing the $L^m(\Theta)$ would be impossible because it is hard to directly draw samples from $f(\mathbf{h}|\Theta)$. The expectation maximization (EM) algorithm is able to iteratively maximize the $L^m(\Theta)$ between E-step (Expectation) and M-step (Maximization), but it could be trapped into local maxima. As for this situation, the MCMC algorithm is able to construct the Markov chains by drawing alternating back and forth between the conditional distributions $f(\mathbf{h}|\Theta, data)$ and $f(\Theta|\mathbf{h}, data)$. The recursive simulation eliminates the computation of the integral with respect to $\mathbf{h}$. Unlike the EM algorithm, the update of parameters does not solely depend

on the increase of the likelihood function at each step; it could justify the effect of being trapped into the local maxima. By the Bayes theorem, the posterior target density

$$f(\Theta, \mathbf{h}|data) \propto P(data|\Theta, \mathbf{h})\pi(\Theta) \tag{4.15}$$

where $\pi(\Theta)$ is the prior distribution of parameters.

In the MCMC algorithm, the infectious period T_I was fixed generally [94, 118], as it was highly correlated to R_0. We also fix the reporting rate r; however, it could be post-validated with the probability distribution of the first exported case from the source region (Section 4.2.2). Besides, the update of τ has to go through the state variable $\mathbf{h}$ as τ does not directly relate to the likelihood. So the most convenient way to draw the $\mathbf{h}$ is from the distribution of Θ. Because the random walk proposal is employed for the update of Θ, given univariate value θ, new values θ' are drawn from

$$\theta' = \theta^{(j-1)} + \epsilon \tag{4.16}$$

where ϵ has a symmetric normal density, *i.e.*, $N(0, \sigma^2)$. The σ is the step size of the chain and $\theta' \sim N(\theta^{(j-1)}, \sigma^2)$. So the proposal distribution is a symmetric function

$$\alpha(\theta'|\theta^{(j-1)}) = \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{1}{2\sigma^2}(\theta^{(j-1)} - \theta')^2\right] \tag{4.17}$$

In the construction of latent variables $\mathbf{h} = [S(\tau-1), I(\tau-1)]$, the variability could be large along with the increase of τ in Eq.4.6 , and the acceptance of τ will be problematic. Without loss of generality, we take the expectation of $\mathbf{h}$ from each update of τ in order to adjust the variance inflation, *i.e.*,

$$\begin{aligned} S(\tau-1) &= N - \left[\sum_{t=0}^{\tau-2} p_I(t)S(t)\right] \\ I(\tau-1) &= I(0) + \left[\sum_{t=0}^{\tau-2} p_I(t)S(t)\right] - \left[\sum_{t=0}^{\tau-2} p_R(t)I(t)\right] \end{aligned} \tag{4.18}$$

So that the jump probability of the latent state is equal to that of τ,

$$\alpha(\mathbf{h}'|\mathbf{h}^{(j-1)}) = \alpha(\tau'|\tau^{(j-1)}) \tag{4.19}$$

which does not deviate the principle of the MH algorithm. The approach is similar to the EM algorithm which latent samples are generated from the current estimates of interest, but the parameter updates depending on the Metropolis step instead of maximization of the expectation of the log-posterior function.

The steps of the corresponding Metropolis-Hastings algorithm of reproduction number estimation (R_0) would be:

1. Start the iteration counter at $j = 1$ and set the initial values for $R_0^{(0)}$ and $\tau^{(0)}$;

2. Sample $\mathbf{h}_{|R_0,data}$ from the new value τ', i.e. $\alpha(\mathbf{h}'|\mathbf{h}^{(j-1)})$. Accept the new values with probability $min(1, A)$, where

$$A = \frac{P(\mathbf{h}'|R_0^{(j-1)}, data)\alpha(\mathbf{h}^{(j-1)}|\mathbf{h}')}{P(\mathbf{h}^{(j-1)}|R_0^{(j-1)}, data)\alpha(\mathbf{h}'|\mathbf{h}^{(j-1)})} \tag{4.20}$$

If we accept, $\tau^{(j)} = \tau'$; otherwise, $\tau^{(j)} = \tau^{(j-1)}$ and the chain does not move;

3. Sample $R_{0|\mathbf{h},data}$ by generating new values Θ' from the proposal density $\alpha(R_0'|R_0^{(j-1)})$. Accept the new values with probability $min(1, B)$, where

$$B = \frac{P(R_0'|\mathbf{h}^{(j)}, data)\alpha(R_0^{(j-1)}|R_0')}{P(R_0^{(j-1)}|\mathbf{h}^{(j)}, data)\alpha(R_0'|R_0^{(j-1)})} \tag{4.21}$$

If we accept, $R_0^(j) = R_0'$; otherwise, $R_0^(j) = R_0^{j-1}$ and the chain does not move; and

4. Change the iteration counter from j to $j + 1$ and return to step 2 until convergence is reached.

The probability A and B are equivalent to the product of the likelihood ratio, prior densities ratio, and the proposal densities ratio given the current and the modified Markov chains. Random step sizes are adopted and they are tuned to allow the acceptance rates within 20% to 40%. The MCMC estimation is iterated 100,000 times, in addition to the 10,000 iterations for burn-in period. The burn-in iterations are discarded in order to eliminate the bias from the initially chosen points. The 100,000 iterations are obtained for the posterior distributions. The convergence of Markov chains mixing in the MCMC process will be diagnosed by the autocorrelation function and the

time series trace plot. Stationary chains represent a good mixing pattern in the MCMC.

4.3 Simulation Study

4.3.1 Simulation Scenarios

In this section, we tested the performance of the estimation method. The settings of the simulations were mostly according to the experience of influenza epidemics. We simulated three sets of time series data according to the following settings:

1. P1: $R_0 = 1.2$, $\tau = 28$, and $r = 0.3$;
2. P2: $R_0 = 1.5$, $\tau = 16$, and $r = 0.15$; and
3. P3: $R_0 = 1.8$, $\tau = 7$, and $r = 0.05$.

We also fixed the infectious period (T_I) to three days. About one month expected epidemic time series data were generated from our SIR stochastic model. We fixed the population to 1,000,000 residents. The prior information is showed in Table 4.3.

4.3.2 Results

In the simulation, we reparameterized the transmission coefficient as $\beta = R_0/T_I$ for simulation convenience. Each MCMC estimation took about one hour. The posterior mean, standard deviation (SD), and 95% credible interval from the MCMC estimation are showed in Table 4.1; for the three simulated datasets, all of the estimated parameters were close to the actual values. Convergence was easily obtained for the parameters.

Table 4.1: Simulated posterior mean, standard deviation (SD), and 95% credible interval from MCMC estimation

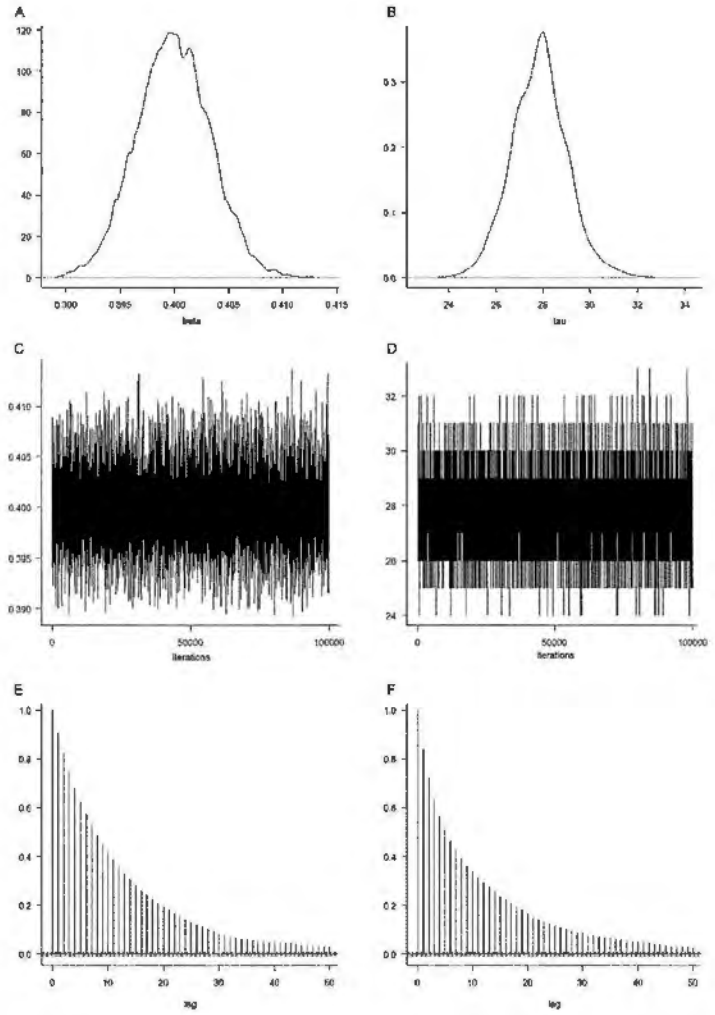
	Basic reproduction number (R_0)			Initial reporting delay (τ)		
Data	**Actual value**	**Mean(SD)**	**95% CI**	**Actual value**	**Mean(SD)**	**95% CI**
P1	1.2	1.2(0.009)	[1.179, 1.221]	28	28(1)	[26, 30]
P2	1.5	1.5(0.002)	[1.491, 1.509]	16	15(1)	[14, 17]
P3	1.8	1.8(0.001)	[1.794, 1.806]	7	7(1)	[6, 8]

As showed in Figures 4.1(A, B), 4.2(A, B), and 4.3(A, B), all of the posterior distributions were distinct from the non-informative priors. The acceptance rates were maintained within 20% to 40%. According to Figures 4.1(C, D), 4.2(C, D), and 4.3(C, D), the MCMC chains were well-mixing with random patterns after discarding the burn-in iterations. The convergence of parameters also showed low correlations to their lags (Figures 4.1(E, F), 4.2(E, F), and 4.3(E, F)). Starting values were randomly selected with finite log-likelihood. Different starting values did not deviate so much in convergence of the estimation.

4.4 Case Study: Contain the Influenza A (H1N1) outbreak at Mexico

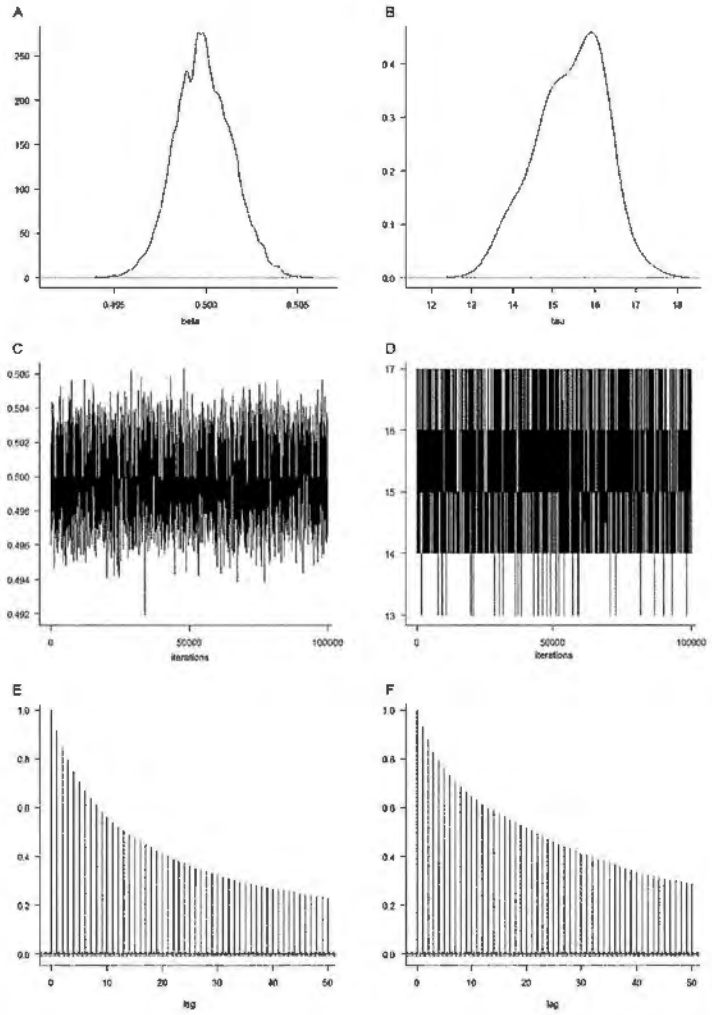
In this section, the methodology was applied to the influenza A (H1N1) pandemic in Mexico in 2009. The objective was to estimate the basic reproduction number subjected to the initial reporting delay and the under-reporting in the first wave of the influenza A (H1N1) pandemic at Mexico. Followed by the estimation, we studied the distribution of the exported cases, and thus assessed the possibility of containing the disease in the source country, Mexico.

Figure 4.1: Posterior distributions, time series trace and autocorrelation plots of β and τ for simulation set P1.



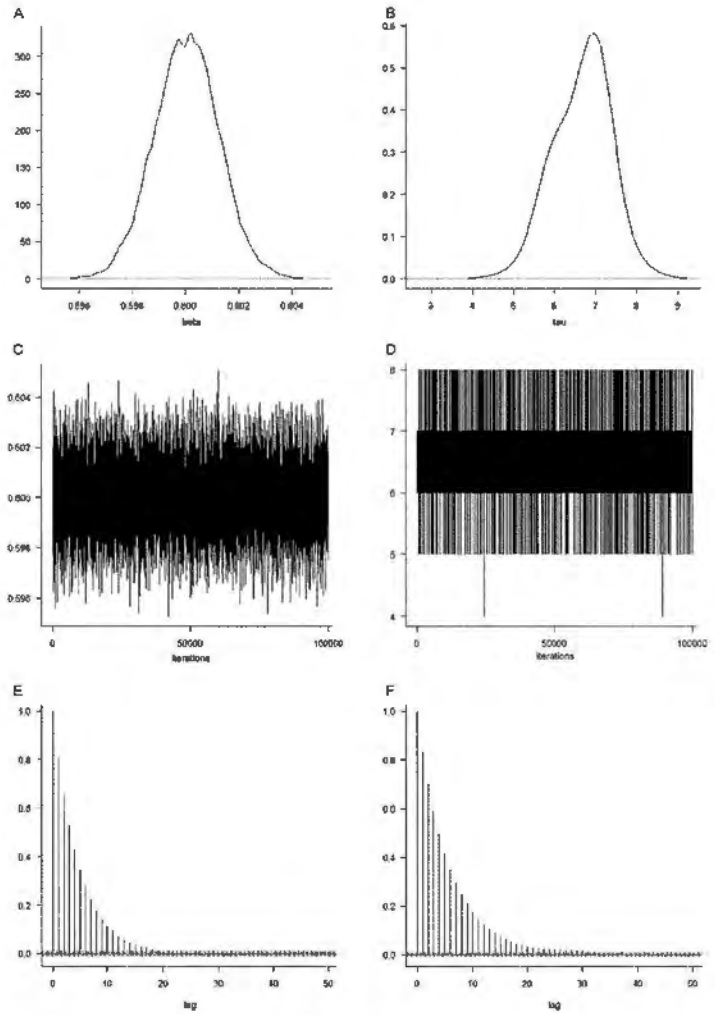
The kernel smoothed posterior distributions are showed on the upper panel (A and B). The time series trace plot illustrated the jumps of 100,000 iterations and are showed on the middle panel (C and D). The fifty-lag autocorrelation plots are showed on the bottom panel (E and F). The β and τ are aligned on the left column (A, C, and E) and on the right column (B, D, and F) respectively.

Figure 4.2: Posterior distributions, time series trace and autocorrelation plots of β and τ for simulation set P2.



The kernel smoothed posterior distributions are showed on the upper panel (A and B). The time series trace plot illustrated the jumps of 100,000 iterations and are showed on the middle panel (C and D). The fifty-lag autocorrelation plots are showed on the bottom panel (E and F). The β and τ are aligned on the left column (A, C, and E) and on the right column (B, D, and F) respectively.

Figure 4.3: Posterior distributions, time series trace and autocorrelation plots of β and τ for simulation set P3.

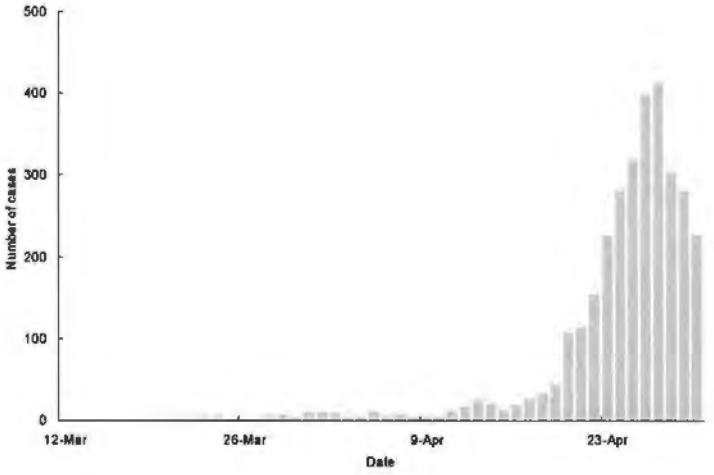


The kernel smoothed posterior distributions are showed on the upper panel (A and B). The time series trace plot illustrated the jumps of 100,000 iterations and are showed on the middle panel (C and D). The fifty-lag autocorrelation plots are showed on the bottom panel (E and F). The β and τ are aligned on the left column (A, C, and E) and on the right column (B, D, and F) respectively.

4.4.1 Materials

The population of Mexico (N) was set as 106,682,518 in 2009 which figure was provided by the National Council for Population of Mexico [81]. The influenza A (H1N1) surveillance data shown in Figure 4.4 were obtained from the Ministry of Health of Mexico for the months between March 14, 2009 and April 30, 2009 [80]. This period covered the first wave of the influenza A (H1N1) epidemic.

Figure 4.4: Confirmed cases in Mexico between March 14, 2009 and April, 30, 2009.



In order to validate the model, we adopted the number of passengers on flights from airports in Mexico, which data were provided by Fraser, *et al.* [47]. The estimation only involved these countries given that their first imported influenza A (H1N1) cases had traveled to Mexico with evidence. The passenger counts, the date of arrival, and the corresponding reference are showed in Table 4.2. The daily rates of travel from Mexico to particular country i-th, m_i, were roughly calculated by dividing the passenger count in March and April 2009

by 61 days. We did not take the United States into account because flying was not the only means of cross-border transport between the two countries. The rest of the daily travel rates from different countries were used in further simulation.

Table 4.2: Number of travelers exported from Mexico by air in March and April, 2009 and date of the first imported cases from Mexico to the corresponding countries

Destination country	**Travelers counts**	**Date of the first exported case from Mexico (2009)**	**Reference**
Canada	101,313	28, Apr	[47]
Spain	65,724	28, Apr	[47]
United Kingdom	20,513	28, Apr	[47]
Costa Rica	16,950	29, Apr	[47]
Germany	35,772	30, Apr	[47]
Netherlands	27,640	30, Apr	[83]
France	61,960	1, May	[46]
Columbia	24,535	3, May	[26]
El salvador	15,090	4, May	[36]
Argentina	24,609	7, May	[8]
Belgium	5,240	-	-
Brazil	38,749	-	-
Chile	18,535	-	-
Cuba	42,802	-	-
Guatemala	39,460	-	-
Honduras	2,340	-	-
Italy	12,060	-	-
Japan	4,675	-	-
Nicaragua	3,101	-	-
Panama	48,717	-	-
Peru	15,478	-	-
Venezuela	9,150	-	-

* USA was not taken into account in estimation as airline was not the only main cross-border transportation from Mexico.

4.4.2 Epidemiological Details

The epidemiological details were mostly from the previous findings of influenza A (H1N1). The reproduction number (R_0) was in the range of 1 to 3 and it followed a uniform prior distribution [47, 121, 118]. The length of the infectious period (T_I) was fixed as three days, and its sensitivity was tested. We also set the initial delay (τ) ranging from two days to 120 days, and the previous estimates of the start date of the outbreak were around mid-January 2009 to late-February 2009.

Because we did not have any information of the reporting rate (r) during the first wave of the epidemic, r was chosen within 0.1% to 40% by grid searching and was fixed before each MCMC estimation; the optimum r would be chosen which minimized the mean absolute difference (MAD) between the simulated and the actual day of the first exported case (FET) from Mexico (Table 4.2 and Section 4.2.2) along with other estimated parameters.

The details of the parameters are highlighted in Table 4.3.

Table 4.3: Parameter definitions, prior distributions, reference, and remarks

Parameter	**Definition**	**Prior distribution**	**Ref/remarks**
R_0	Reproduction number	Uniform(1, 3)	[47, 121, 118]
τ	Initial delay before the first global surveillance	Uniform(1, 120)	[47, 106]
T_I	Infectious period	Fixed at 3 days	[44, 16, 121]
m_i	Daily rates of travel from Mexico to country i-th	Fixed	See Table 4.2

4.4.3 Computer simulation

The methodology was implemented in software R 2.12.1. The syntax and the functions of the R programs were highlighted in Appendix B.

4.4.4 Result

Initial outbreak in Mexico

The developed MCMC algorithm was applied to the influenza A (H1N1) data. 100,000 iterations were drawn after discarding the 10,000 burn-in period. As showed in the autocorrelation plot (Figure 4.5(E, F)), there were no strong lags for the iterations. According to the time series trace plots (Figure 4.5(C, D)), the convergences were obtained and the posterior distributions were drawn; the posterior distributions were not highly skewed. The estimated results are summarized in Table 4.4.

Table 4.4: Acceptance rates, estimated posterior mean, standard deviation (SD), and 95% credible interval for surveillance data

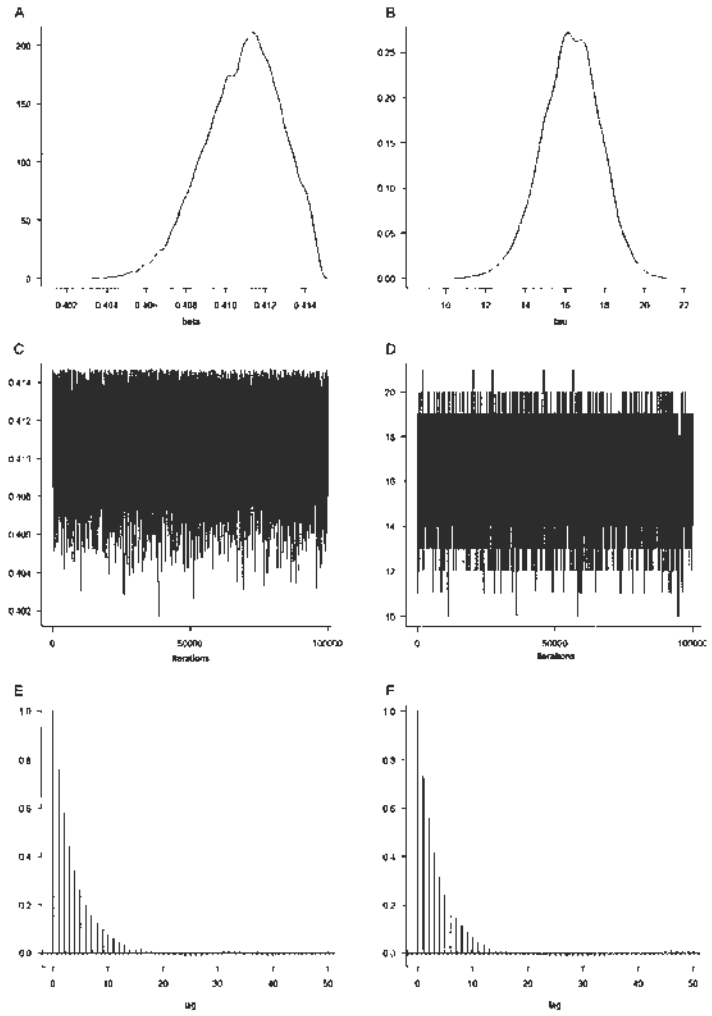
Parameter	**Acceptance rate**	**Mean(SD)**	**95% CI**
R_0	36%	1.233(0.006)	[1.221 to 1.242]
τ	37%	16(1)	[13 to 19]

By a grid search method, the r was iteratively estimated and it was about 8% with a MAD between the simulated and the actual days of the first exported case from different countries (Figure 4.6). The only disadvantage of this method is that it is time-consuming. The result was similar to the value of 5.2% as estimated by Wu, *et al.* [118] in the Hong Kong influenza A study at the end of June, 2009.

From Table 4.4, the estimated R_0 was about 1.233 with a credible interval [1.221 to 1.242]. Compared to the studies from Fraser, *et al.* [47] and Yang, *et al.* [121], our estimated R_0 was slightly lower than the range of 1.4 to 1.6 and the range of 1.3 to 1.7 from the epidemiological analyses, respectively; but close to the value of 1.22 from a genetic study described by Fraser, *et al.* Our precision of the estimated R_0 was higher than that of the 95% confidence interval [1.05 to 1.60] in the genetic analysis.

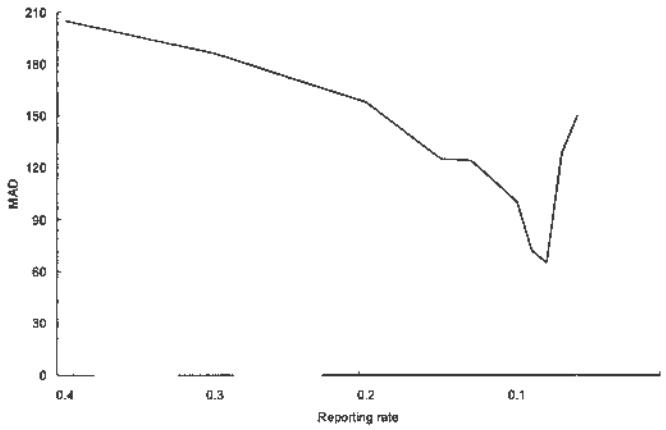
As showed in Table 4.4, the estimated start date of the outbreak was around

Figure 4.5: Posterior distributions, time series trace and autocorrelation plots of β and τ for influenza A (H1N1) surveillance data.



The kernel smoothed posterior distributions are showed on the upper panel (A and B). The time series trace plot illustrated the jumps of 100,000 iterations and are showed on the middle panel (C and D). The fifty-lag autocorrelation plots are showed on the bottom panel (E and F). The β and τ are aligned on the left column (A, C, and E) and on the right column (B, D, and F) respectively.

Figure 4.6: Mean absolute error (MAD) for different reporting rates in MCMC estimation

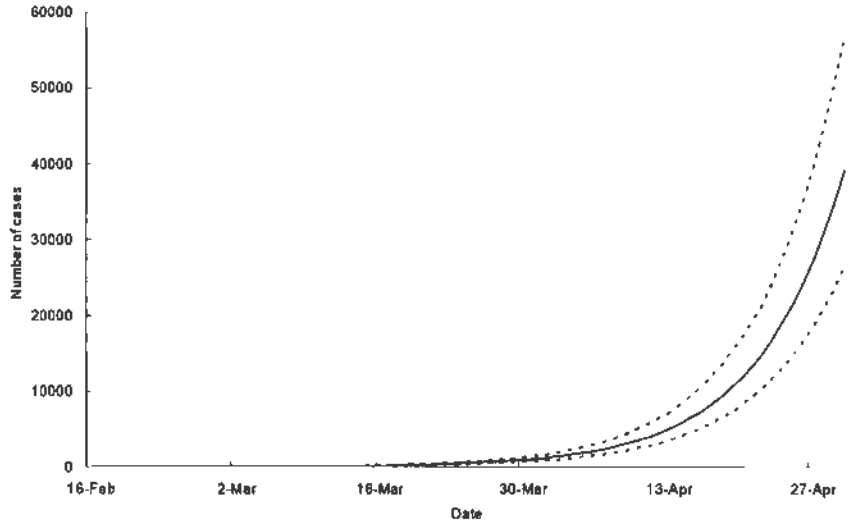


February 25, 2009 with a 95% credible interval [February 23, 2009 to March 1, 2009]. In line with the study from Towers, *et al.* [106], our estimated start date of the outbreak was similar to their estimate value of late-February by fitting an SIR model. Our estimate was also in good agreement with the finding from Fraser, *et al.* [47] which estimated that the initial time of the outbreak was within a 95% credible interval [November 3, 2008 to March 2, 2009].

By using the MCMC samples, we simulated the initial growth of the epidemic until the end of April, 2009. One thousand incidence curves were simulated by randomly choosing the estimated values from the previous MCMC samples. The result is demonstrated in Figure 4.7 and Table 4.5. According to the table, more than 1,000 individuals had been infected with the influenza A (H1N1) virus at the end of March, at which time the emerging virus had still not been confirmed. The infections increased exponentially and nearly 6,000 local cumulative incidence arose in Mexico in mid-April, 2009. At the same time, officials started to confirm that the emerging virus was a new influenza

A (H1N1) virus [78]. By the end of April, almost 40,000 individuals in Mexico had been infected with influenza A (H1N1), and the severity of disease transmission caused the WHO to raise the pandemic alert level to phase five. Our findings were in agreement with the study of Fraser, *et al.* [47], which suggested more than 20,000 individuals were infected with influenza A (H1N1) in Mexico by the end of April. Marc, *et al.* [71] estimated at least 113,000 to 375,000 influenza A H1N1 incidence during the month of April, 2009 by means of the person-at-risk approach. Our estimated cumulative incidence was lower than that of Marc, *et al.*, which maybe caused largely by the spatial effect.

Figure 4.7: Estimated cumulative incidence up to the end of April 2009.



The observed cases were showed in orange bar while the estimated cases were showed in green line; the solid line was the median incidence, and the dotted lines were the non-parametric 95% upper confidence interval and lower confidence interval respectively in 1,000 simulations.

An additional sensitivity analysis was performed in order to test how sensitive of the length of the infectious period (T_I) was to our results. A shorter T_I = (2.5 days) and a longer T_I (3.5 days) were set. The MCMC results were

Table 4.5: Estimated cumulative incidence on March 31, April 15, and April 30, 2009.

Date	Median (95%CI)
March 31	1,045 [733 to 1,456]
April 15	6,366 [4,439 to 9,135]
April 30	39,175 [26,518 to 56,915]

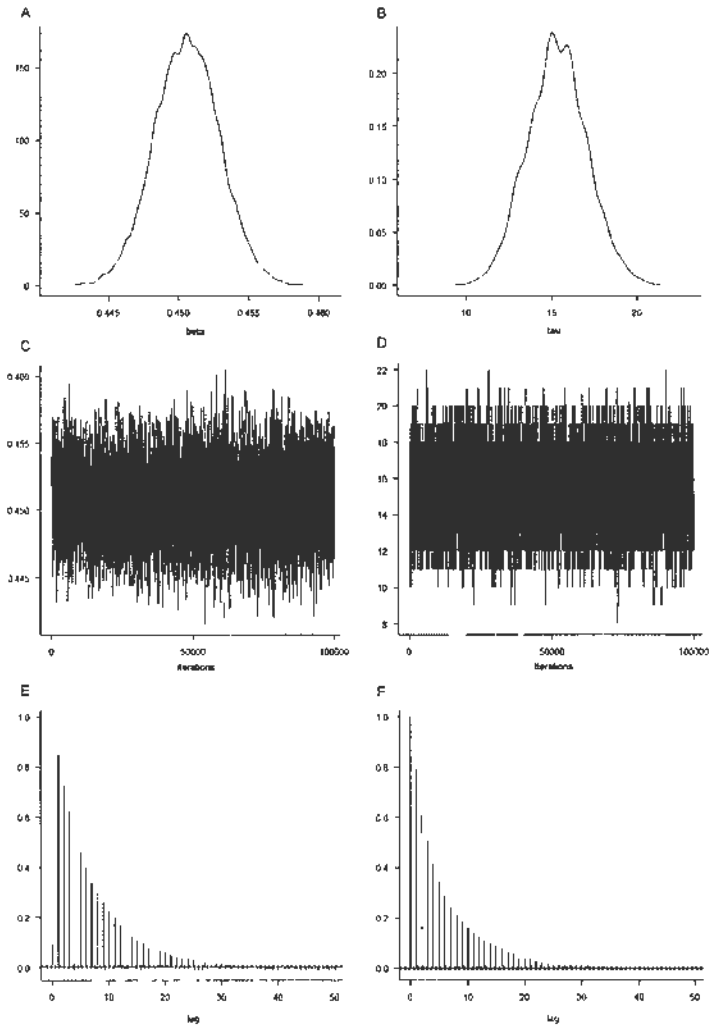
showed in Figure 4.8 and 4.9. Given a shorter infectious period ($T_I = 2.5$), the mean R_0 and τ were estimated as 1.128 with 95% CI [1.115 to 1.138] and 15 with 95% CI [12 to 19] respectively. The result did not deviate greatly from our original result. As showed in Figure 4.8(C-F), the iterations were well-mixed and did not violate the MCMC diagnosis. And for the longer infectious period ($T_I = 3.5$), the mean R_0 and τ was estimated as 1.313 with 95% CI [1.306 to 1.314] and 16 with 95% CI [13 to 18] respectively. The result was not too sensitive, although the distribution of R_0 was slightly skewed. According to Figure 4.9(C-F), the iterations were also well-mixed and did not violate the MCMC diagnosis.

Possibility of containing the disease at the source country by travel restrictions

Followed by the estimation, the distributions of the infected cases' exportation time from Mexico, which related to the rates of travels were studied. We further investigated the possibility of containing the disease in the source country by travel restrictions.

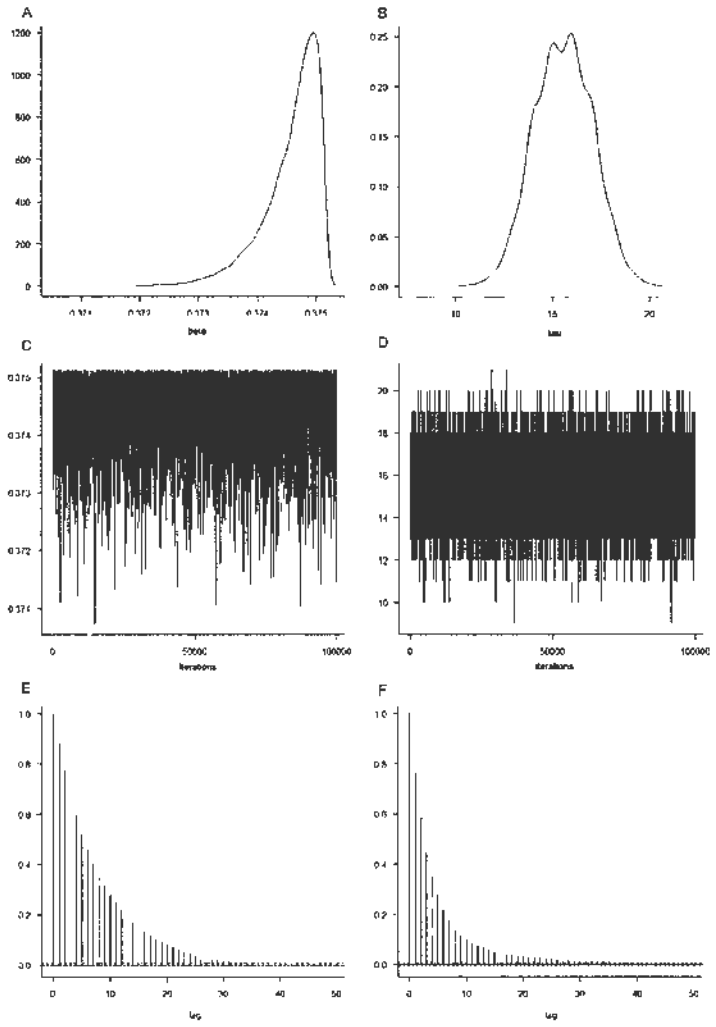
In this section, the distributions of the cases exportation time in the baseline scenario was mostly simulated from the MCMC samples. The days of exporting a certain numbers of infected cases such as the first infection exportation time (FET), were the main endpoints in the study. As showed in Table 4.2, the daily rates of travel (m) were in the range $\approx$ 40 to 1,600 travelers from Mexico to the other countries. We adopted $m = 300$ and $m = 1,500$

Figure 4.8: Posterior distributions, time series trace and autocorrelation plots of β and τ given $T_I = 2.5$.



The kernel smoothed posterior distributions are showed on the upper panel (A and B). The time series trace plot illustrated the jumps of 100,000 iterations and are showed on the middle panel (C and D). The fifty-lag autocorrelation plots are showed on the bottom panel (E and F). The β and τ are aligned on the left column (A, C, and E) and on the right column (B, D, and F) respectively.

Figure 4.9: Posterior distributions, time series trace and autocorrelation plots of β and τ given $T_I = 3.5$.



The kernel smoothed posterior distributions are showed on the upper panel (A and B). The time series trace plot illustrated the jumps of 100,000 iterations and are showed on the middle panel (C and D). The fifty-lag autocorrelation plots are showed on the bottom panel (E and F). The β and τ are aligned on the left column (A, C, and E) and on the right column (B, D, and F) respectively.

to study the distribution of the cases exportation time in Figure 4.10 and 4.11. The statistics are summarized in Table 4.6. In the baseline scenario ($R_0 = 1.23$), a country with a lower travel rate to Mexico, such as Chile and the UK ($m \approx 300$), would directly receive an influenza A (H1N1) infected case about 2.5 months after the first local case arose. Up to ten cases arrived in an at-risk country at the fourth month (Table 4.6). The daily probability of exporting infected cases from Mexico was no more than 0.1 for the first two months since the local case arose (Figure 4.10A). Once a country had a larger passenger flow from the source country, such as Canada ($m \approx 1,500$), the FET would mostly arise at the beginning of the third month, followed by additional three weeks time for the tenth imported case (Table 4.6); the daily probability of exporting at least one case would approach 0.05 (Figure 4.10B).

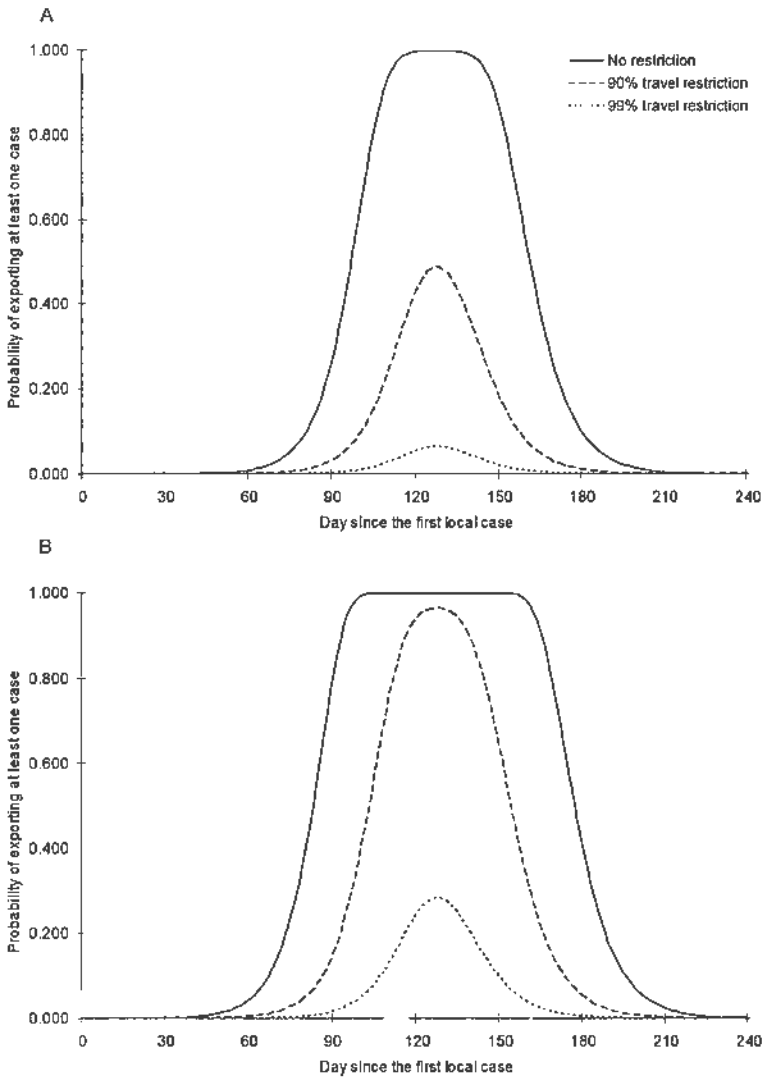
Table 4.6: Exportation days of infected cases for 0%, 90%, and 99% travel restrictions.

m	**Order of cases**	**Restriction ratio**	**LL**	**Median**	**UL**
300	First	0%	51	79	93
		90%	71	98	114
		99%	0	117	149
	Tenth	0%	91	101	107
		90%	116	125	137
		99%	0	0	0
1500	First	0%	38	65	79
		90%	57	84	99
		99%	76	104	121
	Tenth	0%	81	87	92
		90%	101	107	113
		99%	0	134	164

The medians and the confidence intervals were simulated from 50,000 iterations for each scenario. Zero day represented insufficient exported cases in the scenario.
* LL = Lower bound of a non-parametric 95% confidence interval.
* UL = Upper bound of a non-parametric 95% confidence interval.

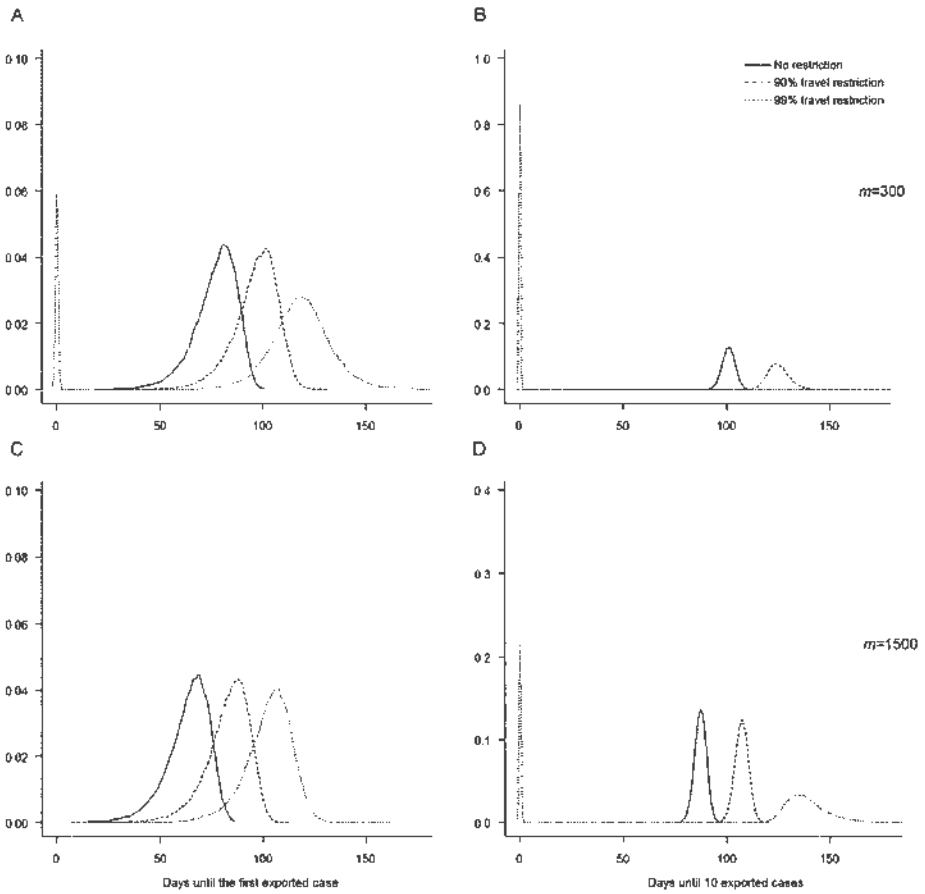
Suppose a 90% ($m = 30$) and a 99% ($m = 3$) travel restrictions were

Figure 4.10: Probability of exporting at least one case from the source country by days ($q_{t,>0}$) at baseline.



The daily rates of travel (m) were set as 300 and 1,500 in the upper figure (A) and the lower figure (B) respectively. No intervention, a 90% travel restriction, and a 99% travel restriction were illustrated in a solid line, a dashed line, and a dotted line respectively. The median reproduction number from the MCMC samples was used.

Figure 4.11: Probability distributions of time until cases exported given different daily rates of travel.



The daily rates of travel (m) were set as 300 and 1,500 in upper panel (A and B) and lower panel (C and D) respectively. The distribution of days since the first local case until the first exported case (FET) and ten exported cases were aligned on the left column (A and C) and the right column (B and D) respectively. No intervention, 90% travel restriction, and 99% travel restriction were illustrated in solid line, dashed line, and dotted line respectively. The distributions were simulated in 50,000 iterations. Zero day represented insufficient exported cases in the scenario.

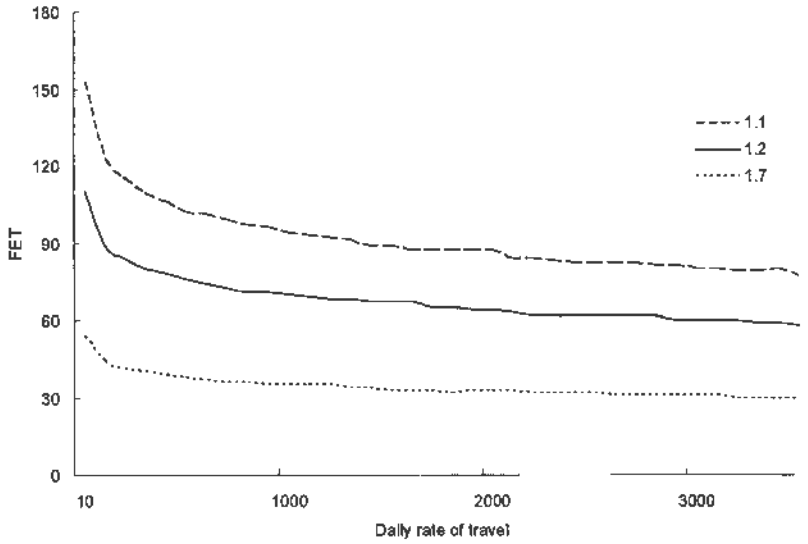
implemented between the UK and Mexico on the day that the first local case arose, almost two weeks to 1.5 months additional delay to FET was observed for the restrictions comparison with the no intervention scenario (Figure 4.11A and Table 4.6). Most of the daily probabilities of exporting cases were reduced by half when the 90% travel restrictions were imposed on air travel between the UK and Mexico (Figure 4.10A). Optimistically, once Mexico maintained 99% travel restrictions to the UK, the probability of exporting ten infected cases was near zero at the initial epidemic outbreak (Figure 4.11B). The 99% travel restrictions were able to reduce the daily probability of exporting at least one case to no more than 0.07 (Figure 4.10A).

As for Canada, 99% travel restriction ($m = 15$) still obtained about an additional 1.5 months' delay to FET in comparison with the no intervention scenario. Canada would have the first ten cases directly from Mexico on the 4.5-th month. A maximum daily exporting probability of 0.28 could be obtained (Figure 4.10B), when the 99% travel restrictions were implemented between them. However, the 90% travel restrictions ($m = 150$) were not able to stop the daily probability approaching to one during the epidemic (Figure 4.10B). Compared to the baseline situation, 90% travel restrictions mainly delayed both FET and the first ten cases for almost a further three weeks from Mexico to Canada (Figure 4.11(C, D)).

The behavior of the FET in terms of the magnitude of travel rates is summarized in Figure 4.12 and Table 4.7. In a baseline setting, the FET could be delayed for more than 100 days when the daily rate of travel was kept below 30 per day, which was almost equivalent to 99% travel restrictions between Canada and Mexico; a delay of more than three months could also be obtained, given lower than 100 daily rate of travel. In a milder scenario ($R_0 = 1.1$), an at-risk country could delay the first non-local case arrival on the fifth month once the daily rate of travel from the source region fell below ten; delay of more than three months to the FET from Mexico to another country was observed

even if there were about 3,000 travelers per day between them. When the scenario became severe ($R_0 = 1.7$), a delay of no more than two months to the FET was allowed, even if the daily rate of travel was below ten. Surprisingly, the FET did not change very much once the daily rate of travel exceeded 1,000 between Mexico and a particular country for each of the scenarios.

Figure 4.12: First exportation time (FET) of infected case given different daily rates of travel.



The median FETs were illustrated in solid line, dashed line, and dotted line for baseline scenario ($R_0 = 1.2$), mild scenario ($R_0 = 1.1$), and severe scenario ($R_0 = 1.7$) respectively. The medians were simulated from 10,000 iterations.

On the other hand, the effectiveness of travel restrictions depended greatly on the disease transmission intensity (Figure 4.13 and Table 4.7). Given a mild scenario ($R_0 = 1.1$), the FET could be delayed to the mid-fifth month for $m = 30$ from Mexico. The difference of FET between 30 and 3,000 daily rates of travel, which was equivalent to placing 99% travel restrictions, could be as long as two months between a connected country to Mexico; 90% travel restrictions

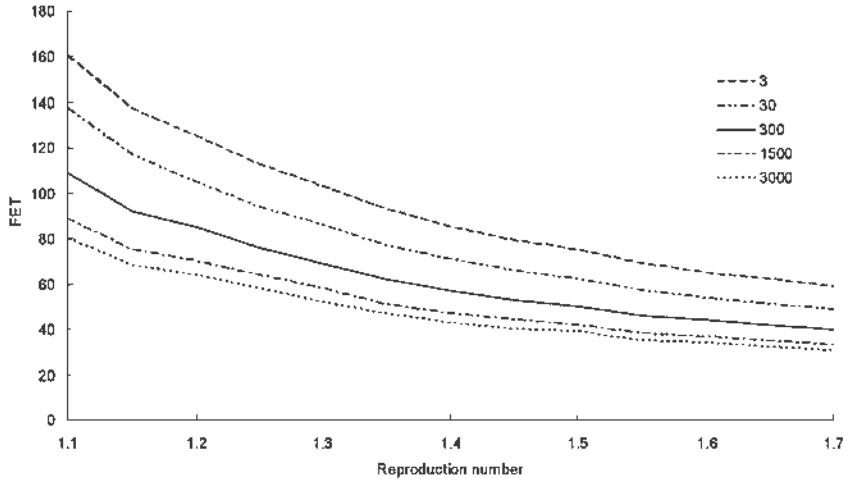
Table 4.7: First exportation time (FET) of infected case given different daily rates of travel (m) and reproduction numbers (R_0).

R_0	m	**LL**	**Median**	**UL**
1.1	3	0	162	211
	30	95	138	162
	300	69	109	131
	1500	47	89	110
	3000	39	80	101
1.2	3	0	126	160
	30	75	106	122
	300	54	84	100
	1500	40	70	85
	3000	34	63	78
1.7	3	44	59	71
	30	36	49	56
	300	26	40	47
	1500	20	33	40
	3000	17	31	38

The medians and the confidence intervals were simulated from 10,000 iterations for each scenario.
* LL = Lower bound of a non-parametric 95% confidence interval.
* UL = Upper bound of a non-parametric 95% confidence interval.

could also have a one-month delay to FET. As showed in Figure 4.13, FET mostly arrived in a country on the fifth month once the R_0 was kept around 1.1 for $m = 30$; if other interventions, such as antiviral drugs and vaccination, could had been applied to the initial growth of the epidemic (as well as reducing the local disease transmission), the containment of the epidemic would have not been impossible. Once the epidemic was uncontrollable, the 99% travel restrictions were still able to defer the FET for more than one month for $R_0 = 1.4$; the delay to FET decreased to half-a-month under 99% travel restrictions, but the delay was for no more than two weeks when 90% travel restrictions were in place and when $R_0 = 1.7$.

Figure 4.13: First exportation time (FET) of infected case given different reproduction numbers.

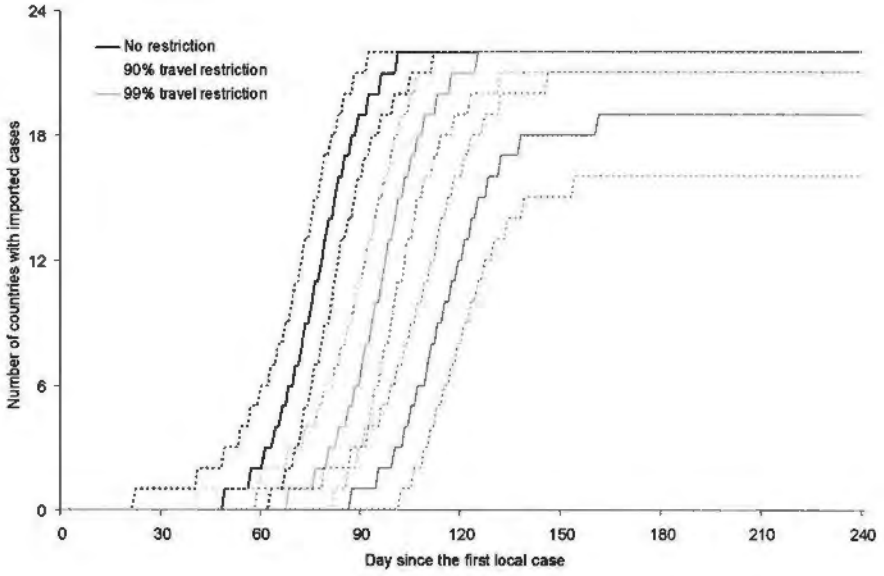


The median FETs were illustrated in dashed line, dot-dot-dash line, solid line, dot-dash line, and dotted line for daily rates of travel $m = 3$, $m = 30$, $m = 300$, $m = 1,500$, and $m = 3,000$ respectively. The medians were simulated from 10,000 iterations.

In order to have a macro view of the impact of travel restrictions, we illustrate the distributions of exporting cases from Mexico to a total of 22

countries (Figure 4.14). Without any travel restrictions, all of the 22 countries would import at least one infected case directly from the source country by the end of the fourth month after the first local case arose. Although 90% travel restrictions could delay for one further month the exporting of cases from Mexico, all of the countries would eventually import infected cases. Several countries were able to prevent the arrival of cases once 99% travel restrictions were imposed. However, more than 80% of countries still imported cases after imposing strict travel restrictions to Mexico.

Figure 4.14: Number of countries with imported case from Mexico for 0%, 90%, and 99% travel restrictions by days.



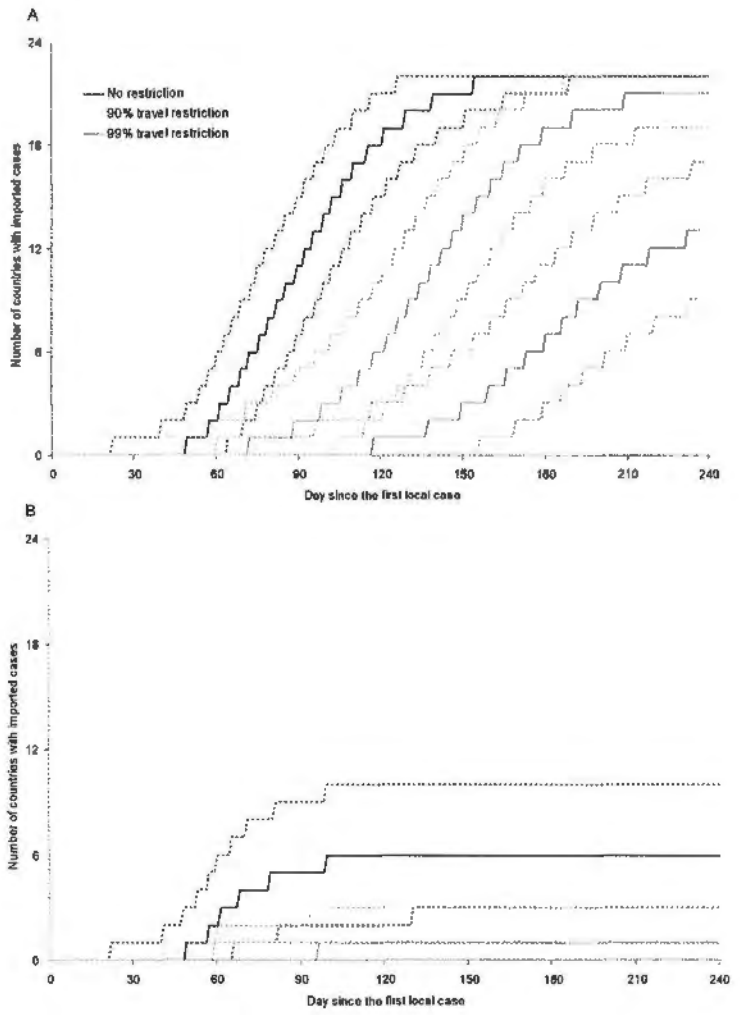
The medians were illustrated in solid lines. The upper bounds and lower bounds of non-parametric 95% confidence intervals were illustrated in dotted lines. Each scenario was simulated from 10,000 iterations.

As described above, the severity of the disease transmission was an important factor with regard to the effectiveness of the travel restrictions. If other interventions, such as antiviral drugs, have reduced 20% of the R_0 after the

60th day (i.e., the effective reproduction number ≈ 1), several at-risk countries have been able to prevent the arrival of cases even with 90% travel restrictions; 99% travel restrictions were able to decrease the number of countries by half (Figure 4.15A). Once the effectiveness of the other interventions could decrease 40% of the R_0 (i.e. the effective reproduction number ≈ 0.7), no more than three countries would received an infection from the source country when imposing 90% travel restrictions during the epidemic. And 99% travel restrictions would contain the influenza A (H1N1) epidemic at the source country (Figure 4.15B). Therefore, reducing the severity of the disease transmission during the early phase of the epidemic would greatly enhance the effectiveness of travel restriction and the possibility of containment.

In the simulation scenarios, we implemented the travel restrictions on the day of the first influenza A (H1N1) case announcement i.e., March 14, 2009, and the estimated initial start time of epidemic was at late February. Fraser, *et al.* [47] has estimated the lower bound of the initial time could be down to early November. However, if the epidemic started in November, the local incidence would have been closed to the epidemic peak when the first influenza A (H1N1) case was announced. The probabilities of exporting infected cases approached to one even though the travel restrictions were imposed on March 14, 2009 (day 134-th) immediately (Figure 4.10). The late detection would contribute to a large public health impact of the outbreak. If the epidemic had been started in early January, imposing the travel restrictions on the day of the first influenza A (H1N1) case announcement (day 73-th) would have been still effective since the probabilities of exporting infected cases had been below 0.1 before the implementation.

Figure 4.15: Number of countries with imported case from Mexico when the reproduction numbers were reduced on day 60-th.



A 20% and a 40% decrease of the reproduction numbers were demonstrated in the upper figure (A) and the lower figure (B) respectively. The medians were illustrated in solid lines. The upper bounds and lower bounds of non-parametric 95% confidence intervals were illustrated in dotted lines. Each scenario was simulated from 10,000 iterations.

4.5 Discussion

In this section, we developed a stochastic SIR model to study the distribution of the exported infections from the source country. The model incorporated the aspects of the time delay until an epidemic initiated and the under-reporting in the parameters estimation. The developed model was based on Bayesian inference which took the uncertainty into account for the model parameters along with the stochastic variation. Besides, a MCMC algorithm was developed to impute the unobserved process within the dynamic equations. In order to validate the estimation algorithm, a simulation study was done. The simulation results were satisfactory as all of the parameters converged to the acceptable values and performed well in diagnosis.

We further applied the methodology to parameters estimation in the initial outbreak period of influenza A (H1N1) epidemic at Mexico. The estimated basic reproduction number, subjected to issues of the initial delay, and the under-reporting, was about 1.233 with a credible interval [1.221 to 1.242] using the surveillance data between March 14, 2009 and April 30, 2009 (Figure 4.4). Moreover, the estimated start date of outbreak was around February 25, 2009 with a 95% credible interval [February 23, 2009 to March 1, 2009], and the reporting rate was about 8%. The estimates were consistent to other studies.

By incorporating the estimates, the impact of the travel restrictions, as well as the possibility of containing the influenza A (H1N1) epidemic, were examined in the view of the source country. In the baseline scenario ($R_0 = 1.23$), the FETs were around 2 to 2.5 months to other countries. The UK, to which a daily rate of travelers was closed to 300, import no more than ten cases from Mexico during the epidemic when 99% travel restrictions were implemented. Even though Canada had 1,500 travelers per day from Mexico, it was possible to have no more than ten imported cases after 99% travel restrictions were imposed. Nevertheless, imposing 90% travel restrictions only

delayed the time of cases exporting from Mexico for any country. In terms of the magnitude of travel rates, the FET could be deferred for more than 100 days once the daily rate of travel was restricted to below 30 persons per day. But it would not change so much once the daily rate of travel exceeded 1,000. Moreover, the effectiveness of travel restrictions increased with the reduction of the local disease transmission intensity.

Generally, most of the countries would receive the infections from the source country given the complex airline network whether or not they imposed strict travel restrictions. As described by Hollingsworth, et. al. [60], early control measures in the source region were crucial to contain the epidemic. If other interventions had reduced a certain proportion of the disease transmission intensity initially, the travel restriction would have been able to prevent the arrival of cases from other countries; as large as possible in the R_0 reduction, it would have been possible to contain the influenza A (H1N1) epidemic at the source region.

In conclusion, the sole adoption of travel restrictions would be insufficient to halt the spread of the epidemic. Because of the high incidence rates in Mexico, only strict 99% restrictions on travelling i.e., three to 15 travelers per day, could have a chance of preventing an at-risk country from importing cases from the source region. However, if the travel restriction had been implemented in combination with other interventions, such as antiviral drugs, to reduce the disease transmission locally, the containment at the source area would have been possible. Most importantly, the travel restrictions were valuable in retarding the export of cases in terms of weeks.

Chapter 5

Summary and Conclusion

5.1 Summary of Findings

In this thesis, new methodologies were developed, by performing a case study on the influenza A (H1N1) pandemic in 2009 to assess the impact of travel restriction on the spread of disease in views of the *at-risk* countries and the *source* country.

In the pandemic influenza A (H1N1), the international traffic accelerated the spread of infections across a wide geographic area. Policy-makers will be interested to learn of its impact on the disease transmission once the traffic has been re-scaled. Because the pharmaceutical interventions will not be available during the early stage of the pandemic, travel restrictions should be a high-potential intervention for including into the official containment and mitigation strategies. In some researches, the value of travel restrictions remains controversial and, more importantly, several practical and theoretical limitations have been found and were described in chapter 2. These problems largely motivated us to study the effectiveness of travel restriction on the pandemic control in both *source* country and *at-risk* countries, in order to have well planning of strategies in the future.

In chapter 3, we developed a stochastic model that incorporated air, sea, and land transportation to explore the impact of the travel restriction in view

of the *at-risk* countries. The use of antivirals and hospitalization was also incorporated in the model in order to provide a more realistic compartment on the recovery, and also to compare the effectiveness of these control measures. The modeling results showed that restrictions on air travel, the main means of transport from the Americas to Hong Kong, was the most effective of the three types of restriction in delaying the arrival of the infected cases during the early stage of the epidemic. With the use of antivirals and hospitalization, the restrictions on all means of transport could reduce the peak incidence by more than half. Also, the spread of the local epidemic was halted by these interventions when the scenario was mild. However, the effectiveness of the travel restrictions strongly relied on the use of antivirals and hospitalization when the scenario was severe. Our result also showed that if other control measures had been taken effect in the non-local countries that could decrease the disease transmission intensity, the restriction of all means of external transport would possibly have halted the local spread of the disease in seven months time, whether or not the antivirals and hospitalization had been used. Moreover, increasing the screening sensitivity at the entry border points was beneficial in slowing down the growth of the cumulative attack rates. In brief, the greatest value of travel restrictions was in their ability to slow down the spread of the epidemic. With the imposition of other interventions that can suppress the disease transmission intensity, whether locally or not, the restrictions on all external travel reduced the local attack rates, and they even halted the disease spread. According to our additional cost-effectiveness study (Appendix C), the travel restriction was cost-effective and the epidemic delay reduced a large portion of health care costs for imposing 5 months and 6.5 months before the availability of effective interventions once the disease transmission intensity was comparatively mild with 6% to 15% case-fatality rate. In general, the travel restriction was also cost-effective for a late delivery of treatments when case-fatality rate attained 25%.

In chapter 4, a new methodology was developed to evaluate the possibility of the disease containment by travel restriction in view of the *source* country. A MCMC method, which imputed the unobserved process within the dynamic equations, was also developed to estimate the reproduction number subjected to the problems of initial reporting delay and the reporting rate behind the surveillance data. The estimation algorithm has been validated by a series of simulation. The methodology was further applied to parameters estimation in the initial outbreak period of influenza A (H1N1) epidemic in Mexico. The estimated basic reproduction number was about 1.233 with a credible interval [1.221 to 1.242] and the estimated start date of outbreak was around February 25, 2009 with a 95% credible interval [February 23, 2009 to March 1, 2009]. By incorporating the estimates, the impact of the travel restriction as well as the possibility of containing the influenza A (H1N1) epidemic was examined in the view of the *source* country. Due to the high incidence rate in Mexico, only a strict 99% restriction on travelling, i.e., allowing three to fifteen travelers exported per day, could have a chance to prevent an at-risk country from importing cases from the source region. However, if the travel restriction had been implemented along with other interventions such as antiviral to reduce the disease transmission locally, the containment at the source area would have been possible. Besides, early control measures in the source region were crucial to contain the epidemic. In most of the situations, travel restriction was able to slow down the export of cases in terms of weeks.

In summary, travel restriction of either *at-risk* countries or the *source* country is valuable on slowing down the growth of epidemics. The time delay of the epidemic would offer public health experts, policy makers, and scientists more time for preparation and decision making on epidemic control especially when an unknown virus emerges to our society. Although solely imposing the travel restriction showed little benefit on reducing the final attack rate and the probability of cases exportation, it offered additional contribution on even

halting the epidemic growth once other interventions such as antiviral and hospitalization could also be utilized. Therefore, the implementation of the travel restriction should be a potential intervention to control the epidemic spread, especially when the next epidemics which could be lethal and highly intrusive.

5.2 Limitation

The methodologies have four major limitations.

First, limited data affected the model structure. In chapter 3, we focused on the local disease transmission dynamic incorporating the transportation from a local area to the others. Due to limited data, we could not construct every coupling between countries in the model so that we could not depict the result from a global point of view. On the other hand, we adopted a simple stochastic SIR model structure only with the use of the incidence count instead of a $SEIR$ model in chapter 4. In practice, border points may not be able to screen out all individuals especially for those in latent status, so the incorporation of the latent compartment into the model is required to quantify the impact of screening sensitivity. But since we could not confirm whether the exported cases who arrived in other countries were in latent status or infectious status, the adoption of SIR model would be preferred in the estimation process. Moreover, the investigated travel restrictions were rigorously enforced in chapter 4, and both latent and infectious individuals would have the same rescaling proportion. Therefore, imposing the extra latent compartment in the model does not improve our result.

Second, the characteristics of the travel pattern and the influenza A (H1N1) virus affected the model structure. In chapter 3 and chapter 4, the daily rates

of travels were assumed uniformly distributed but they may decrease gradually due to the increasing severity of disease spread, thus, fat-tailed distribution such as Log-normal and Weibull distributions would be more appropriate. Apart from that, several studies discovered that the distributions of the incubation and infectious periods were mainly right skewed like Gamma, Log-normal, and Weibull distributions [30, 29, 109, 69]. Since the length of the infectious period highly correlated to the reproduction number, the convergence was bad and it was not considered in the MCMC algorithm. On the other hand, since the variability from the binomial distributions have been adopted in the process of infections generation, the importance of adding extra variability from the periods would be relatively low. However, additional sensitivity analysis has been done for the lengths of latent and infectious durations to explore these effect on our results in chapter 3 and chapter 4. In general, it is suggested conducting household transmission studies in order to draw more realistic distributions before being employed in the models [29].

Third, the resolution level of our models may not be high enough. In chapter 3, we did not quantify the risk of infection of an inbound travel and on an aircraft [19], the local incidence may be underestimated. Additional compartments could be built to account for those effects, but it will increase the model complexity. Moreover, transports that require long traveling time do not easily allow rapid international spread of disease with a short generation time. It cannot be doubted that travels that spend several days to Hong Kong exist, but such cases must be very small proportion. For example, the proportion of overseas passenger who directly traveled from overseas by sea transport is extremely low. But instead of taking direct flight or ferry to Hong Kong, some overseas passenger prefer taking a flight to Macau or mainland China and then transfer to Hong Kong by ferry or train connections due to the economical and time concerns. The multi-leg travel would greatly reduce the required waiting time since infection among import cases. In our model,

no adjustment was made for multi-leg travel for one or more transports due to the limitation of available statistics; however, the previous results did not show a big quantitative difference between single-leg and multi-leg travels [37].

Forth, the MCMC estimation converges badly in some situations. In chapter 4, we have designed several situations for the simulation in accordance with previous influenza scenarios; but in some extreme scenarios, such as when the reproduction number is close to 1, the convergence would take a long time with poor mixing. Moreover, the MCMC method works best if the data are taken from the initial growth of the epidemic; it works much less well when the data are taken after the peak time. But we seldom apply the method beyond the initial epidemic growth because it is unrealistic that epidemics are discovered by officials after the peak times.

5.3 Future Research

In the thesis, we offered advice on the implementation of travel restrictions through the use of epidemic modeling on the influenza A (H1N1) pandemic. But in the future, new viruses could be lethal and highly invasive when antiviral drugs or vaccination are not yet ready. So in what scenario that we should implement the travel restrictions and even combination with other interventions with maximum benefit? The question motivates us to consider widely in the scenario and intervention setting.

First, our result showed that travel restriction worked better when the initial growth of the epidemic was mild; it benefited the control disease more when the new virus had a mild transmissibility but caused high mortality. The model should explore its potential applicability of various infectious disease virus such as SARS. Additionally, a more comprehensive travel network with heterogeneity in travelers should be considered. In balancing the health impacts, economic costs, and intervention efficacies, the threshold point of

travel restrictions implementation should be investigated.

Second, there are many available pharmaceutical and non-pharmaceutical interventions to control the epidemics nowadays. In the thesis, we assessed the interaction between the practice of travel restriction, antivirals, and hospitalization; other interventions such as vaccination [85] and school closure [110] are also effective in reducing the growths of epidemics. Because of limited resources and safety impacts, it is better to have considerate and optimum combination strategies for controlling epidemics effectively across wide ranges of disease transmissibility and lethality. Therefore, a large scale compartmental model or a network model should be developed to evaluate each combination strategy.

To conclude, the future research will provide a throughout guidelines on containment and mitigation of an epidemic to health policy makers.

Bibliography

[1] Writing Committee of the WHO Consultation on Clinical Aspects of Pandemic (H1N1) 2009 Influenza. Clinical aspects of pandemic 2009 influenza A (H1N1) virus infection. *New England Journal of Medicine*, 362(18):1708–1719, 2010.

[2] H. Abbey. An examination of the Reed-Frost theory of epidemics. *Human Biology*, 24(3):201–233, 1952.

[3] M. E. Alexander, S. M. Moghadas, G. Rest, and J. T. Wu. A delay differential model for pandemic influenza with antiviral treatment. *Bulletin of Mathematical Biology*, 70(2):382–397, 2008.

[4] M. an der Heiden, U. Buchholz, G. r. Krause, H. Kirchner, G. andD Claus, and W. H. Haas. Breaking the waves: modelling the potential impact of public health measures to defer the epidemic peak of novel influenza A/H1N1. *PLoS ONE*, 4(12):e8356, 2009.

[5] R. M. Anderson and R. M. May. Population biology of infectious diseases: Part I. *Nature*, 280(5721):361–367, 1979.

[6] R. M. Anderson and R. M. May. Directly transmitted infections diseases: control by vaccination. *Science*, 215(4536):1053–1060, 1982.

[7] Centers for Disease Control and Prevention. Questions and answers: Antiviral drugs, 2009-2010 flu Season. Retrieved 30 December 2009, from: http://www.cdc.gov/H1N1flu/antiviral.htm/, 2010.

[8] Reuters. Argentina confirms first H1N1 flu case. Released on 12 May 2009. Retrieved 28 January 2010, from: http://www.reuters.com/article/idUSN07423083, 2009.

[9] J. Arino, J. David, D. Hartley, R. Jordan, J. Miller, and P. van den Driessche. A multi-species epidemic model with spatial migration. *Mathematical Medicine and Biology*, 22(2):129–142, 2005.

[10] J. Arino, R. Jordan, and P. van den Driessche. Quarantine in a multi-species epidemic model with spatial dynamics. *Mathematical Biosciences*, 206(1):46–60, 2007.

[11] M. Baguelin, A. J. Van Hoek, M. Jit, S. Flasche, P. J. White, and W. J. Edmunds. Vaccination against pandemic influenza A/H1N1v in England: a real-time economic evaluation. *Vaccine*, 28(12):2370 2384, 2010.

[12] M. S. Bartlett. Some evolutionary stochastic processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 11(2):211–229, 1949.

[13] C. T. Bauch, J. O. Lloyd-Smith, M. P. Coffee, and A. P. Galvani. Dynamically modeling SARS and other newly emerging respiratory illnesses: past, present and future. *Epidemiology*, 16(6):791–801, 2005.

[14] N. G. Becker. Uses of the EM algorithm in the analysis of data on HIV/AIDS and other infectious diseases. *Statistical Methods in Medical Research*, 6(1):24 37, 1997.

[15] L. M. A. Bettencourt and R. M. Ribeiro. Real time bayesian estimation of the epidemic potential of emerging infectious diseases. *PLoS ONE*, 3(5):e2185, 2008.

[16] P. Y. Boelle, P. Bernillon, and J. C. Desenclos. A preliminary estimation of the reproduction ratio for new influenza A(H1N1) from the outbreak in Mexico, March-April 2009. *Eurosurveillance*, 14(19):pii=19205, 2009.

[17] L. Brouwers, B. Cakici, M. Camitz, A. Tegnell, and M. Boman. Economic consequences to society of pandemic H1N1 influenza 2009: preliminary results for Sweden. *Eurosurveillance*, 14(37):pii=19333, 2009.

[18] J. S. Brownstein, C. J. Wolfe, and K. D. Mandl. Empirical evidence for the effect of airline travel on inter-regional influenza spread in the United States. *PLoS Medicine*, 3(10):e401, 2006.

[19] P. Caley, N. G. Becker, and D. J. Philp. The waiting time for inter-country spread of pandemic influenza. *PLoS ONE*, 2(1):e143, 2007.

[20] G. Chowell, C. E. Ammon, N. W. Hengartner, and J. M. Hyman. Transmission dynamics of the great influenza pandemic of 1918 in Geneva, Switzerland: assessing the effects of hypothetical interventions. *Journal of Theoretical Biology*, 241(2):193–204, 2006.

[21] G. Chowell, N. W. Hengartnerb, C. Castillo-Chaveza, P. W. Fenimorea, and J. M. Hyman. The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. *Journal of Theoretical Biology*, 229(1):119–126, 2004.

[22] Centre for Health Protection, Department of Health, HKSAR. Swine and Seasonal Flu Monitor (Week 29). Retrieved 16 July 2010, from: `http://www.chp.gov.hk/en/guideline1_year/29/134/518.html`, 2010.

[23] M. L. Ciofi degli Atti, S. Merler, C. Rizzo, M. Ajelli, M. Massari, P. Manfredi, C. Furlanello, G. Scalia Tomba, and M. Iannelli. Mitigation measures for pandemic influenza in Italy: an individual based model considering different scenarios. *PLoS ONE*, 3(3):e1790, 2008.

[24] V. Colizza and A. Barrat. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences*, 103(7):2015–2020, 2006.

[25] V. Colizza, A. Barrat, M. Barthelemy, A. J. Valleron, and A. Vespignani. Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS Medicine*, 4(1):e13, 2007.

[26] Reuters AlertNet. Colombia reports flu case, Ecuador monitors border. Released on 3 May 2009. Retrieved 28 January 2010, from: `http://www.alertnet.org/thenews/newsdesk/N03367692.htm`, 2009.

[27] B. S. Cooper, R. J. Pitman, W. J. Edmunds, and N. J. Gay. Delaying the international spread of pandemic influenza. *PLoS Medicine*, 3(6):e212, 2006.

[28] B. Cowling, L. Lau, P. Wu, H. Wong, V. Fang, S. Riley, and H. Nishiura. Entry screening to delay local transmission of 2009 pandemic influenza A (H1N1). *BMC Infectious Diseases*, 10(1):82, 2010.

[29] B. J. Cowling, V. J. Fang, S. Riley, J. S. Malik Peiris, and G. M. Leung. Estimation of the serial interval of influenza. *Epidemiology*, 20(3):344–347, 2009.

[30] B. J. Cowling, M. P. Muller, I. O. L. Wong, L. M. Ho, M. Louie, A. McGeer, and G. M. Leung. Alternative methods of estimating an incubation distribution: examples from severe acute respiratory syndrome. *Epidemiology*, 18(2):253–259, 2007.

[31] U. C. De Silva, J. Warachit, S. Waicharoen, and M. Chittaganpitch. A preliminary analysis of the epidemiology of influenza A(H1N1)v virus infection in Thailand from early outbreak data, June-July 2009. *Eurosurveillance*, 14(31):pii=19292, 2009.

[32] O. Diekmann and J. A. P. Heesterbeek. *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation.* Wiley; New York, 2000.

[33] European Centre for Disease Prevention and Control. Public health use of influenza antivirals during influenza pandemics. ECDC Interim Guidance. Retrieved 1 September 2011, from: `http://ecdc.europa.eu/en/publications/Publications/0907_GUI_Public_Health_use_of_Influenza_Antivirals_during_Influenza_Pandemic.pdf`, 2009.

[34] European Centre for Disease Prevention and Control. Daily update on the 2009 influenza A (H1N1) pandemic. Retrieved 28 January 2010, from: `http://www.ecdc.europa.eu/en/healthtopics/h1n1/pages/daily_update.aspx`, 2010.

[35] B. D. Elderd, V. M. Dukic, and G. Dwyer. Uncertainty in predictions of disease spread and public health responses to bioterrorism and emerging diseases. *Proceedings of the National Academy of Sciences*, 103(42):15693–15697, 2006.

[36] Channel NewsAsia. El Salvador confirms first two cases of Influenza A(H1N1). Released on 4 May 2009. Retrieved 28 January 2010, from: `http://www.channelnewsasia.com/stories/afp_world/view/426728/1/.html`, 2009.

[37] J. M. Epstein, D. M. Goedecke, F. Yu, R. J. Morris, D. K. Wagener, and G. V. Bobashev. Controlling pandemic flu: the value of international air travel restrictions. *PLoS ONE*, 2(5):e401, 2007.

[38] Census and Statistics Department, HKSAR. Transport, Communications and Tourism Statistics. Hong Kong Resident Departures by Control Point. Retrieved 25 November 2009, from: `http://www.censtatd.gov.hk/FileManager/EN/Content_807/transport.pdf`, 2007.

[39] N. M. Ferguson, D. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, and D. Burke. Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature*, 437(7056):209–214, 2005.

[40] N. M. Ferguson, D. A. T. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and B. D. S. Strategies for mitigating an influenza pandemic. *Nature*, 442(7101):448–452, 2006.

[41] N. M. Ferguson, C. A. Donnelly, and R. M. Anderson. Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature*, 413(6855):542–548, 2001.

[42] K. A. Fitzner, K. F. Shortridge, S. M. McGhee, and A. J. Hedley. Cost-effectiveness study on influenza prevention in Hong Kong. *Health Policy*, 56(3):215–234, 2001.

[43] A. Flahault and A. J. Valleron. A method for assessing the global spread of HIV-1 infection based on air travel. *Mathematical Population Studies*, 3(3):161–171, 1992.

[44] A. Flahault, E. Vergu, and P. Y. Boelle. Potential for a global dynamic of influenza A (H1N1). *BMC Infectious Diseases*, 9(1):129, 2009.

[45] A. Flahault, E. Vergu, L. Coudeville, and R. F. Grais. Strategies for containing a global influenza pandemic. *Vaccine*, 24(44-46):6751 6755, 2006.

[46] Influenza A(H1N1)v investigation teams. Modified surveillance of influenza A(H1N1)v virus infections in France. Eurosurveillance, 14(29):pii=19276, 2009.

[47] C. Fraser, C. A. Donnelly, S. Cauchemez, W. P. Hanage, M. D. Van Kerkhove, T. D. Hollingsworth, J. Griffin, R. F. Baggaley, H. E.

Jenkins, E. J. Lyons, T. Jombart, W. R. Hinsley, N. C. Grassly, F. Balloux, A. C. Ghani, N. M. Ferguson, A. Rambaut, O. G. Pybus, H. Lopez-Gatell, C. M. Alpuche-Aranda, I. B. Chapela, E. P. Zavala, D. M. E. Guevara, F. Checchi, E. Garcia, S. Hugonnet, C. Roth, and T. W. R. P. A. Collaboration. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science*, 324(5934):1557 1561, 2009.

[48] D. Gamerman. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman & Hall, 1997.

[49] R. Gani, H. Hughes, D. Fleming, T. Griffin, J. Medlock, and S. Leach. Potential impact of antiviral drug use during influenza pandemic. *Emerging Infectious Diseases*, 11(9):1355–1362, 2005.

[50] R. Gani and S. Leach. Transmission potential of smallpox in contemporary populations. *Nature*, 414(6865):748–751, 2001.

[51] A. E. Gelfand, S. E. Hills, A. Racine-Poon, and A. F. M. Smith. Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*, 85(412):972–985, 1990.

[52] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *Journal of Applied Statistics*, 20(5):721 741, 1984.

[53] G. J. Gibson, A. Kleczkowski, and C. A. Gilligan. Bayesian analysis of botanical epidemics using stochastic compartmental models. *Proceedings of the National Academy of Sciences*, 101(33):12120–12124, 2004.

[54] G. J. Gibson and E. Renshaw. Estimating parameters in stochastic compartmental models using Markov chain methods. 15(**1**):19–40, 1998.

[55] R. F. Grais, G. Hugh, and E. Gregory. Assessing the impact of airline travel on the geographic spread of pandemic influenza. *European Journal of Epidemiology*, 18(11):1065–1072, 2003.

[56] M. Greenwood. On the statistical measure of infectiousness. *Journal of Hygiene (London)*, 31(3):336–351, 1931.

[57] W. H. Hamer. Epidemic disease in England. *The Lancet*, 1(733):739, 1906.

[58] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[59] Census and Statistics Department, HKSAR. Gross Domestic Product (GDP), implicit price deflator of GDP and per capita GDP. Retrieved 1 September 2011, from: `http://www.censtatd.gov.hk/hong_kong_statistics/statistical_tables/index.jsp?charsetID=1&tableID=030&subjectID=12`, 2009.

[60] T. D. Hollingsworth, N. Ferguson, and R. Anderson. Will travel restrictions control the international spread of pandemic influenza? *Nature Medicine*, 12(5):497–499, 2006.

[61] L. Hufnagel, D. Brockmann, and T. Geisel. Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences*, 101(42):15124–15129, 2004.

[62] Hong Kong Tourism Board. Visitor Arrival Statistics 2007. Retrieved 19 April 2010, from: `http://partnernet.hktourismboard.com/pnweb/jsp/doc/listDocL.jsp?charset=en&doc_id=107560&filename=VAS+12200`, 2007.

[63] Reuters. New flu has been around for years in pig - study. Fox, Maggie. 11 June 2009. Retrieved 19 April 2010, from: http://www.reuters.com/article/middleeastCrisis/idUSN11399103, 2009.

[64] N. Jennifer and E. J. Orav. A generalized chain binomial model with application to HIV infection. *Mathematical Biosciences*, 101(1):99 -119, 1990.

[65] H. A. Kelly, G. N. Mercer, J. E. Fielding, G. K. Dowse, K. Glass, D. Carcione, K. A. Grant, P. V. Effler, and R. A. Lester. Pandemic (H1N1) 2009 influenza community transmission was established in one Australian state when the virus was first identified in North America. *PLoS ONE*, 5(6):e11341, 2010.

[66] W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics, part I. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 115(772):700–721, 1927.

[67] N. Lee, P. K. S. Chan, G. C. Y. Lui, B. C. K. Wong, W. W. Y. Sin, K.-W. Choi, R. Y. K. Wong, E. L. Y. Lee, A. C. M. Yeung, K. L. K. Ngai, M. C. W. Chan, R. W. M. Lai, A. W. Y. Yu, and D. S. C. Hui. Complications and outcomes of pandemic 2009 influenza A (H1N1) virus infection in hospitalized adults: how do they differ from those in seasonal influenza? *Journal of Infectious Diseases*, 203(12):1739–1747, 2011.

[68] P. E. Lekone and B. F. Finkenstadt. Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics*, 62(4):1170 1177, 2006.

[69] J. Lessler, N. G. Reich, R. Brookmeyer, T. M. Perl, K. E. Nelson, and D. A. Cummings. Incubation periods of acute respiratory viral infections: a systematic review. *The Lancet Infectious Diseases*, 9(5):291 -300, 2009.

[70] M. Lipsitch, T. Cohen, B. Cooper, J. M. Robins, S. Ma, L. James, G. Gopalakrishna, S. K. Chew, C. C. Tan, M. H. Samore, D. Fisman, and M. Murray. Transmission dynamics and control of severe acute respiratory syndrome. *Science*, 300(5627):1966 -1970, 2003.

[71] M. Lipsitch, M. Lajous, J. J. O'Hagan, T. Cohen, J. C. Miller, E. Goldstein, L. Danon, J. Wallinga, S. Riley, S. F. Dowell, C. Reed, and M. McCarron. Use of cumulative incidence of novel influenza A/H1N1 in foreign travelers to estimate lower bounds on cumulative incidence in Mexico. *PLoS ONE*, 4(9):e6895, 2009.

[72] A. L. Lloyd. Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. *Proceedings of the Royal Society B: Biological Sciences*, 268(1470):985–993, 2001.

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

[74] I. M. Longini, M. E. Halloran, A. Nizam, and Y. Yang. Containing pandemic influenza with antiviral agents. *American Journal of Epidemiology*, 159(7):623 -633, 2004.

[75] I. M. Longini, A. Nizam, S. Xu, K. Ungchusak, W. Hanshaoworakul, D. A. Cummings, and E. M. Halloran. Containing pandemic influenza at the source. *Science*, 309(5737):1083–1087, 2005.

[76] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087 -1092, 1953.

[77] WHO Department of Epidemic and Pandemic Alert and Response. WHO ad hoc scientific teleconference on the current influenza A(H1N1) situation. 11 June 2009. Retrieved 19 April 2010, from: http://www.who.int/csr/resources/publications/swineflu/tc_report_2009_04_29/en/index.html, 2009.

[78] Centers for Disease Control and Prevention. Morbidity and Mortality Weekly Report, 58(Dispatch):1-3. Swine influenza A (H1N1) infection in two children, Southern California, March-April 2009. Retrieved 19 April 2010, from: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm58d0421a1.htm, 2009.

[79] Centers for Disease Control and Prevention. Morbidity and Mortality Weekly Report, 58(Dispatch):1-3. Update: swine influenza A (H1N1) infections, California and Texas, April 2009. Retrieved 19 April 2010, from: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm58d0424a1.htm, 2009.

[80] The Ministry of Health of Mexico. Current status of the epidemic, 12 October 2009. Distribution of confirmed cases according to date of onset of symptoms. Retrieved 19 April 2010, from: http://portal.salud.gob.mx/descargas/pdf/influenza/situacion_actual_epidemia_121009.pdf, 2009.

[81] National Council for Population of Mexico. Estimated population 2009. Retrieved 19 April 2010, from: http://www.conapo.gob.mx/index.php?option=com_content&view=article&id=125&Itemid=193, 2009.

[82] C. E. Mills, J. M. Robins, and M. Lipsitch. Transmissibility of 1918 pandemic influenza. *Nature*, 432(7019):904–906, 2004.

[83] Health Protection Agency West Midlands H1N1v Investigation Team. Preliminary descriptive epidemiology of a large school outbreak of influenza A(H1N1)v in the West Midlands, United Kingdom. Eurosurveillance, 14(27):pii=19264, 2009.

[84] H. Nishiura. Early efforts in modeling the incubation period of infectious diseases with an acute course of illness. *Emerging Themes in Epidemiology*, 4(1):2, 2007.

[85] M. Nuno, G. Chowell, and A. Gumel. Assessing the role of basic control measures, antivirals and vaccine in curtailing pandemic influenza: scenarios for the US, UK and the Netherlands. *Journal of the Royal Society Interface*, 4(14):505–521, 2007.

[86] P. D. O'Neill. Bayesian inference for stochastic multitype epidemics in structured populations using sample data. *Biostatistics*, 10(4):779–791, 2009.

[87] P. D. O'Neill, D. J. Balding, N. G. Becker, M. Eerola, and D. Mollison. Analyses of infectious disease data from household outbreaks by Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 49(4):517–542, 2000.

[88] P. D. O'Neill and G. O. Roberts. Bayesian inference for partially observed stochastic epidemics. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 162(1):121–129, 1999.

[89] U.S. Census Bureau. International Data base. Total midyear population by world 2009. Retrieved 20 April 2010, from: `http://www.census.gov/ipc/www/idb/region.php`, 2009.

[90] Hong Kong Information Services Department. News Archives. HKSAR Press Releases 27 June 2009. New hospital and clinic arrangements for

human swine flu patients. Retrieved 25 November 2009, from: www.info.gov.hk/gia/general/200906/27/P200906270251.htm, 2009.

[91] Hong Kong Information Services Department. News Archives. HKSAR Press Releases on 2 May 2009. Press releases on human swine flu. Retrieved 28 January 2010, from: http://www.isd.gov.hk/pr/eng/, 2009.

[92] Hong Kong Information Services Department. News Archives. HKSAR press releases from 1 May 2009 to 31 Dec 2009. Press releases on human swine flu. Retrieved 28 January 2010, from: http://www.isd.gov.hk/pr/eng/, 2009.

[93] J. Ranta, P. H. Makela, A. Takala, and E. Arjas. Predicting the course of meningococcal disease outbreaks in closed subpopulations. *Epidemiology and Infection*, 123(3):359 371, 1999.

[94] J. Reiczigel, K. Brugger, F. Rubel, N. Solymosi, and Z. Lang. Bayesian analysis of a dynamical model for the spread of the Usutu virus. *Stochastic Environmental Research and Risk Assessment*, 24(3):455–462, 2010.

[95] S. Riley. Large-scale spatial-transmission models of infectious disease. *Science*, 316(5829):1298 1301, 2007.

[96] S. Riley and N. M. Ferguson. Smallpox transmission and control: Spatial dynamics in Great Britain. *Proceedings of the National Academy of Sciences*, 103(33):12637–12642, 2006.

[97] S. Riley, J. T. Wu, and G. M. Leung. Optimizing the dose of pre-pandemic influenza vaccines to reduce the infection attack rate. *PLoS Medicine*, 4(6):e218, 2007.

[98] G. O. Roberts and J. S. Rosenthal. Optimal scaling for various Metropolis-Hastings algorithms. *Statistical Science*, 16(4):351–367, 2001.

[99] R. Ross. An application of the theory of probabilities to the study of a priori pathometry. *Proceedings of the Royal Society of London A*, 92(638):204 230, 1916.

[100] L. A. Rvachev and I. M. Longini Jr. A mathematical model for the global spread of influenza. *Mathematical Biosciences*, 75(1):3–22, 1985.

[101] Census and Statistics Department, HKSAR. Average Wage Rates for Employees up to Supervisory Level (Excluding Managerial and Professional Employees) by Broad Occupational Group by Selected Industry Section. Retrieved 1 September 2011, from: `http://www.censtatd.gov.hk/hong_kong_statistics/statistical_tables/index.jsp?charsetID=1&tableID=021&subjectID=2`, 2009.

[102] M. R. Siddiqui and W. J. Edmunds. Cost-effectiveness of antiviral stockpiling and near-patient testing for potential influenza pandemic. *Emerging Infectious Diseases*, 14(2):267–274, 2008.

[103] A. F. M. Smith and G. O. Roberts. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(1):3–23, 1993.

[104] World Health Organization. Cost-effectiveness thresholds. Choosing Interventions that are Cost Effective (WHO-CHOICE). Retrieved 1 September 2011, from: `http://www.who.int/choice/costs/CER_thresholds/en/index.html`, 2011.

[105] Census and Statistics Department, HKSAR. Tourism Value Added by Industry, 2005, 2007 and 2009. Hong Kong Monthly Digest of Statistics. Retrieved 1 September 2011, from: `http://www.statistics.gov.hk/publication/general_stat_digest/B10100022011MM03B0100.pdf`, 2011.

[106] S. Towers and Z. Feng. Pandemic H1N1 influenza: predicting the course of a pandemic and assessing the efficacy of the planned vaccination programme in the United States. *Eurosurveillance*, 14(41):pii=19358, 2009.

[107] S. M. Tracht, S. Y. Del Valle, and J. M. Hyman. Mathematical modeling of the effectiveness of facemasks in reducing the spread of novel influenza A (H1N1). *PLoS ONE*, 5(2):e9018, 2010.

[108] T. Tsutsui, N. Minami, M. Koiwai, T. Hamaoka, I. Yamane, and K. Shimura. A stochastic-modeling evaluation of the foot-and-mouth-disease survey conducted after the outbreak in Miyazaki, Japan in 2000. *Preventive Veterinary Medicine*, 61(1):45–58, 2003.

[109] A. R. Tuite, A. L. Greer, M. Whelan, A.-L. Winter, B. Lee, P. Yan, J. Wu, S. Moghadas, D. Buckeridge, B. Pourbohloul, and D. N. Fisman. Estimated epidemiologic parameters and morbidity associated with pandemic H1N1 influenza. *Canadian Medical Association Journal*, 182(2):131–136, 2010.

[110] E. Vynnycky and W. J. Edmunds. Analyses of the 1957 (Asian) influenza pandemic in the United Kingdom and the impact of school closures. *Epidemiology and Infection*, 136(2):166–179, 2008.

[111] J. Wallinga and P. Teunis. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *American Journal of Epidemiology*, 160(6):509 516, 2004.

[112] H. J. Wearing, P. Rohani, and M. J. Keeling. Appropriate models for the management of infectious diseases. *PLoS Medicine*, 2(7):e174, 2005.

[113] K. C. Chong. Surveillance data from forty-four countries. Retrieved 28 June 2011, from: `https://spreadsheets.google.com/spreadsheet/`

ccc?key=0AlvdkTczvgPBdDdBcTctUThkUTEwU2tfWEVlRUoxOUE&hl=en_US&authkey=CJ-0qdYL, 2011.

[114] World Health Organization. Weekly epidemiological record. Retrieved 1 March 2010, from: http://www.who.int/wer/2009/en/, 2009.

[115] World Health Organization. Situation updates - Pandemic (H1N1) 2009. Retrieved 28 January 2010, from: http://www.who.int/csr/disease/swineflu/updates/en/index.html, 2009.

[116] World Health Organization. Pandemic (H1N1) 2009 update 109. Retrieved 16 July 2010, from: http://www.who.int/csr/don/2010_07_16/en/index.html, 2010.

[117] J. G. Wood, N. Zamani, C. R. MacIntyre, and N. G. Becker. Effects of internal border control on spread of pandemic influenza. *Emerging Infectious Diseases*, 13(7):1038 1045, 2007.

[118] J. T. Wu, B. J. Cowling, E. H. Lau, D. K. Ip, L. M. Ho, T. Tsang, S.-K. Chuang, P.-Y. Leung, S. V. Lo, S.-H. Liu, and S. Riley. School closure and mitigation of pandemic (H1N1) 2009, Hong Kong. *Emerging Infectious Diseases*, 16(3):538–541, 2010.

[119] J. T. Wu, S. Riley, C. Fraser, and G. M. Leung. Reducing the impact of the next influenza pandemic using household-based public health interventions. *PLoS Medicine*, 3(9):e361, 2006.

[120] W. Xiaoli, Y. Peng, S. Holly, Z. Yi, D. Ying, P. Xinghuo, H. Xiong, and W. Quanyi. Estimates of the true number of cases of pandemic (H1N1) 2009, Beijing, China. *Emerging Infectious Diseases*, 16(11):1786–1788, 2010.

[121] Y. Yang, J. D. Sugimoto, M. E. Halloran, N. E. Basta, D. L. Chao, L. Matrajt, G. Potter, E. Kenah, and J. Longini, Ira M. The transmissibility and control of pandemic influenza A (H1N1) virus. *Science*, 326(5953):729 733, 2009.

Appendix A

SAS programs for studying the impact of travel restriction for *at-risk* countries

The programs were built in software SAS 9.1.3. Here are the following programs and their functions:

- READDATA: To create SAS dataset from csv data.
- TRAVELFIT: To fit the initial exponential growth and estimate the reproduction number for the travel data from non-local countries.
- EIFIT: To fit the $E(0)$s and $I(0)$s for different non-local countries.
- BASELINEFIT: To fit the baseline reproduction number.
- FPTSIM: To simulate the FPT and FHPT from random import.
- SIMMODEL: To simulate the scenario.
- SIMCI: To simulate the confidence intervals for scenarios.

READDATA

```
/*******************************************************************************
/ Program name: READDATA
/ Version: 1.0
/ Author: Marc, Biostatistician
/ Study name: H1N1 simulaton study
/ Created date: 01MAY2010
/ Purpose: To Create SAS dataset from csv data
/ Notes:
/ ============================================================
/ Amendment history:
/ |--Amended date--|--Amended by--|----------Description----------|
/
/ ============================================================
*******************************************************************************/

/* Build Up Library directory */
libname simdata 'D:\PhD Study\Thesis\data';

/* Create SAS Table population */
PROC IMPORT datafile="D:\population.csv" OUT=simdata.population;
  getnames = yes;
run;

/* Create SAS Table export */
PROC IMPORT datafile="D:\export.csv" OUT=simdata.export;
  getnames = yes;
run;

/* Create SAS Table import */
PROC IMPORT datafile="D:\import.csv" OUT=simdata.import;
  getnames = yes;
run;

/* Create SAS Table infect */
PROC IMPORT datafile="D:\infect.csv" OUT=simdata.infect;
  getnames = yes;
run;

/* Create SAS Table local */
PROC IMPORT datafile="D:\local.csv" OUT=simdata.local;
  getnames = yes;
run;

/* Create SAS Table datepattern */
PROC IMPORT datafile="D:\datepattern.csv" OUT=simdata.datepattern;
  getnames = yes;
run;

/* Sort the dataset */
proc sort data = simdata.import;
  by country;
run;

proc sort data = simdata.population;
  by country;
run;

proc sort data = simdata.infect;
  by country date;
run;
```

TRAVELFIT

```
/*******************************************************************************
/ Program name: TRAVELFIT
/ Version: 1.0
/ Author: Marc, Biostatistician
/ Study name: H1N1 simulaton study
```

```
/ Created date: 01MAY2010
/ Purpose: To fit the initial exponential growth and estimate the reproduction number
/          for the travel data from non-local countries
/ Notes:
/ ============================================================
/ Amendment history:
/ |--Amended date--|--Amended by--|----------Description-------------|
/
/ ============================================================
*******************************************************************************************/

/* Build Up Library directory */
libname simdata 'C:\Users\marc.marc-HP\Documents\My SAS Files\PLOS\data';

/* Create individual country day number */
* country with first date;
data first_dt_im (rename =(date=first_dt));
  set simdata.infect (keep = country date) ;
  by country;
  if first.country;
run;

data first_dt_ex (rename =(date=first_dt));
  set simdata.local (keep = country date) ;
  by country;
  if first.country;
run;

* calculate the day number after first onset case;
proc sql;
  create table inf_ima as
  select A.*, log(A.confirmed) as logcon, B.first_dt, (A.date - B.first_dt + 1) as iday
  from simdata.infect as A, first_dt_im as B
  where trim(A.country)=trim(B.country);
quit;

* Limit 2 months for fitting;
data inf_im;
  set inf_ima;
  if iday <61;
run;

proc sql;
  create table inf_ex as
  select A.*, B.first_dt, (A.date - B.first_dt + 1) as iday
  from simdata.local as A, first_dt_ex as B
  where trim(A.country)=trim(B.country);
quit;

/* Macro for exponential fit:
lat: length of latent period
inf: length of latent period */
%macro expfit(lat, inf);

/* Exponential growth rate estimation */
proc reg data = inf_im tableout outest=b noprint;
     model logcon = iday ;
   by country;
   run;

data mb (keep = country iday rename=(iday=r));
  set b;
  if _TYPE_ = "PARMS";
run;

data lb (keep = country iday rename=(iday=lr));
  set b;
  if _TYPE_ = "L95B";
run;

data ub (keep = country iday rename=(iday=ur));
  set b;
  if _TYPE_ = "U95B";
run;
```

```
data umlb;
  merge lb mb ub;
  by country;
run;

/* Reproduction number and beta for inf period = 2.9; lat period = 1.45 */
data epi_para;
  set umlb;
  repro =1 + &inf*&lat*(r**2 + r*((1/&inf)+(1/&lat)));
  lrepro =1 + &inf*&lat*(lr**2 + lr*((1/&inf)+(1/&lat)));
  urepro =1 + &inf*&lat*(ur**2 + ur*((1/&inf)+(1/&lat)));
  beta = repro/&inf;
run;

%mend expfit;

/* Exponential fit for inf period = 2.9; lat period = 1.45 */
%expfit(1.45, 2.9);

/* Deterministic SEIR model by country */

/* Macro for SEIR generation:
e0: initial number of latent subjects
i0: initial number of infectious subjects
lat: length of latent period
inf: length of latent period
dayno: number of days generation */
%macro seir(e0, i0, lat, inf, dayno);

/* Simulated the SEIR for each country */
%do city_id= 1 %to 44;

  proc sql;
    create table init_&city_id as
    select A.country, A.beta, A.repro, B.population
    from epi_para as A, simdata.population as B
    where trim(A.country)=trim(B.country) and B.id=&city_id;
  quit;

  data seir_&city_id;
    set init_&city_id;
    s=population;
    e=&e0;
    i=&i0;
    r=0;
    iday =1;
    output;
    do iday = 2 to &dayno;
    incident=s*(1-exp( ( (-1*beta*i) /population) ));
    infectious=e*(1-exp(-1/&lat));
    remove=i*(1-exp(-1/&inf));
    s=s - incident;
    e=e+incident-infectious;
    i=i+infectious-remove;
    r=r+remove;
    output;
    end;
  run;

    %if &city_id>1 %then %do;

      proc append base = seir_1 data = seir_&city_id;
      run;

      /* List of R */
      proc append base = init_1 data = init_&city_id;
      run;

    %end;

%end;

proc sort data = seir_1;
  by iday country;
run;
```

```
/* Merge the day list */
proc sql;
  create table daylist_im as
  select A.*, B.country
  from simdata.datepattern as A, simdata.import as B;
quit;

proc sql;
  create table daylist_ex as
  select A.*, B.country
  from simdata.datepattern as A, simdata.export as B;
quit;

proc sql;
  create table im_dt as
  select A.*, B.first_dt, (A.date - B.first_dt + 1) as iday
  from daylist_im as A, first_dt_im as B
  where trim(A.country)=trim(B.country);
quit;

proc sql;
  create table ex_dt as
  select A.*, B.first_dt, (A.date - B.first_dt + 1) as iday
  from daylist_ex as A, first_dt_ex as B
  where trim(A.country)=trim(B.country);
quit;

/* Merge the population, import and export data */
proc sql;
  create table im_trans as
  select A.*, C.population, int((B.total)/365) as total_d,
        int((B.air)/365) as air_d, int((B.sea)/365) as sea_d,
        int((B.land)/365) as land_d
  from im_dt as A, simdata.import as B, simdata.population as C
  where trim(A.country)=trim(B.country)=trim(C.country);
quit;

proc sql;
  create table ex_trans as
  select A.*, C.population,  int((B.total)/365) as total_d,
        int((B.air)/365) as air_d, int((B.sea)/365) as sea_d,
        int((B.land)/365) as land_d
  from ex_dt as A, simdata.export as B, simdata.population as C
  where trim(A.country)=trim(B.country)=trim(C.country);
quit;

proc sort data = im_trans;
  by iday country;
run;

/* Merge seir cases to day list */
data seir_im;
  merge im_trans seir_1(keep = iday country e i);
  by iday country;
  if not missing(dayno);
  if iday < 0 then iday=.;
  if missing(e) then e = 0;
  else e = int(e);
  if missing(i) then i = 0;
  else i=int(i);
  m_total=total_d/population;
  m_air=air_d/population;
  m_sea=sea_d/population;
  m_land=land_d/population;
run;

* Sort the case according to day number;
proc sort data = seir_im;
  by dayno country;
run;

/* Merge daily case to local */
data daily_ex (keep =date dayno iday confirmed m_total m_air m_sea m_land);
  merge ex_trans inf_ex(keep = iday confirmed);
```

```
  by iday;
  if not missing(dayno);
  if iday < 0 then iday=.;
  m_total=total_d/population;
  m_air=air_d/population;
  m_sea=sea_d/population;
  m_land=land_d/population;
run;

%mend seir;
```

EIFIT

```
/*****************************************************************************************
/ Program name: EIFIT
/ Version: 1.0
/ Author: Marc, Biostatistician
/ Study name: H1N1 simulaton study
/ Created date: 01MAY2010
/ Purpose: To fit the E(0)s and I(0)s for different non-local countries
/ Notes:
/ ============================================================
/ Amendment history:
/ |--Amended date--|--Amended by--|----------Description-------------|
/
/ ============================================================
*****************************************************************************************/

/* Export and explore the output file */
libname result 'C:\Baselinefit';

/* Macro for E(0), I(0) fitting:
ei_ll: lower limit of E(0), I(0)
ei_ul: upper limit of E(0), I(0)
d: the length of E(0), I(0) increase per step
sim: number of realizations per E(0), I(0)
lat: length of latent period
inf: length of infectious period
screen: the proportion of import infectious subjects
dayno: max number of days generation */

%macro ei_fit(ei_ll, ei_ul, d, sim,lat, inf, screen, dayno);

%let e_sim = &ei_ll;
%let i_sim = &ei_ll;
%let rno = 1;

/* Start E(0), I(0) realization */
%do %until (&e_sim > &ei_ul);

  /* Start per E(0), I(0) simulation */
  %do si = 1 %to &sim;

    %seir(&e_sim, &i_sim, &lat, &inf, &dayno);

    /* Day limit */
    data seir_im_dl;
      set seir_im;
      if dayno < &dayno;
    run;

    /* Generate a set of random import data up to FPT */

    %put My name is Marc. I am running E(0)= &e_sim, I(0)= &i_sim simulation &si of &sim;

    /* Daily sum of random import */
    data seir_im_rand;
      set seir_im_dl;

      * Random latent cases;
      if e = 0 or m_total = 0 then erand_total = 0;
```

```
  else erand_total = ranbin(0,e,m_total);
  if e = 0 or m_air = 0 then erand_air = 0;
  else erand_air = ranbin(0,e,m_air);
  if e = 0 or m_sea = 0 then erand_sea = 0;
  else erand_sea = ranbin(0,e,m_sea);
  if e = 0 or m_land = 0 then erand_land = 0;
  else erand_land = ranbin(0,e,m_land);

  * Random infectious cases;
  if i = 0 or m_total = 0 then irand_total = 0;
  else irand_total = int(&screen*ranbin(0,i,m_total));
  if i = 0 or m_air = 0 then irand_air = 0;
  else irand_air = int(&screen*ranbin(0,i,m_air));
  if i = 0 or m_sea = 0 then irand_sea = 0;
  else irand_sea = int(&screen*ranbin(0,i,m_sea));
  if i = 0 or m_land = 0 then irand_land = 0;
  else irand_land = int(&screen*ranbin(0,i,m_land));

run;

proc sort data = seir_im_rand;
  by date dayno country;
run;

proc sql;
  create table im_list as
  select date, dayno,
    sum(erand_total) as e_total , sum(erand_air) as e_air,
    sum(erand_sea) as e_sea, sum(erand_land) as e_land,
    sum(irand_total) as i_total , sum(irand_air) as i_air,
    sum(irand_sea) as i_sea, sum(irand_land) as i_land
  from seir_im_rand
  group by date, dayno;
quit;

/* pick up the days with imported cases */
data im_list_ind;
  set im_list;
  if e_air > 0 or e_sea  > 0 or e_land >0 or i_air > 0 or i_sea  > 0 or i_land >0
    then st_import =1;
  else st_import=0;
run;

data im_list_pick;
  set im_list_ind;
  if st_import=1;
run;

proc sort data = im_list_pick;
  by st_import dayno;
run;

/* Calculate the FPT */
%let s=_&si;
%let fname= &rno&s;

/* FPT per simulation */
data fpt_&fname;
  set im_list_pick;
  by st_import;
  if first.st_import;
  iter_set=&rno;
  iter=&si;
  eset=&e_sim;
  iset=&i_sim;
run;

/* Store the per E(0), I(0) simulation result */

%if &si > 1 or &rno ^= 1 %then %do;

  proc append base = fpt_i_1 data = fpt_&fname;
  run;

%end;
```

```
  /* End of per E(0), I(0) simulation */
  %end;

  %let e_sim = %sysevalf(&e_sim+&d);
  %let i_sim = %sysevalf(&i_sim+&d);
  %let rno = %sysevalf(&rno+1);

/* End of E(0), I(0) realization */
%end;

/* Simulation summary */
proc sql;
  create table mean_day as
  select mean(dayno) as m_day, eset, iset
  from fpt_1_1
  group by eset, iset;
quit;

proc print;run;

/* 95% CI for FPT */
proc sort data = result.eifit24032011;
  by dayno;
run;

data sortfpt;
  set result.eifit24032011;
  meanlag=int((dayno+lag(dayno))/2);
run;

data ll_fpt(rename= meanlag = lfpt) med_fpt (rename= dayno = mfpt)
      ul_fpt(rename= meanlag = ufpt);
  set sortfpt;
  by dayno;
  if _n_= 3 then output ll_fpt;
  if _n_= 50 then output med_fpt;
  if _n_= 98 then output ul_fpt;
run;

data fptci;
  merge ll_fpt(keep=lfpt) med_fpt(keep=mfpt) ul_fpt(keep=ufpt);
run;

proc print;run;

%put Hello! Finished!;

%mend ei_fit;
```

BASELINEFIT

```
/*****************************************************************************
/ Program name: BASELINEFIT
/ Version: 1.0
/ Author: Marc, Biostatistician
/ Study name: H1N1 simulaton study
/ Created date: 01MAY2010
/ Purpose: To fit the baseline reproduction number
/ Notes:
/ ===========================================================
/ Amendment history:
/ |--Amended date--|--Amended by--|----------Description-------------|
/
/ ===========================================================
*****************************************************************************/

/* Macro for R fitting:
r_ll: lower limit of the reporduction number
r_ul: upper limit of the reporduction number
d: the length of R increase per step
```

```
sim: number of realizations per R
lat: length of latent period
inf: length of latent period
screen: the proportion of import infectious subjects
dayno: number of days generation */

%macro baselinefit(r_ll, r_ul, d, sim,lat, inf, screen, dayno);

/* Initiated the countries own cases with estimated E(0), I(0) */
%seir(90, 90, 1.45, 2.9, &dayno);

/* Day limit */
data seir_im_dl;
  set seir_im;
  if dayno < &dayno;
run;

/* Start R realization */
%let repro = &r_ll;
%let rno = 1;

%do %until (&repro > &r_ul);

  /* Start per R simulation */
  %do si = 1 %to &sim;

  %put My name is Marc. I am running R0= &repro simulation &si of &sim;

  /* Daily sum of random import */
  data seir_im_rand;
    set seir_im_dl;

    * Random latent cases;
    if e = 0 or m_total = 0 then erand_total = 0;
    else erand_total = ranbin(0,e,m_total);
    if e = 0 or m_air = 0 then erand_air = 0;
    else erand_air = ranbin(0,e,m_air);
    if e = 0 or m_sea = 0 then erand_sea = 0;
    else erand_sea = ranbin(0,e,m_sea);
    if e = 0 or m_land = 0 then erand_land = 0;
    else erand_land = ranbin(0,e,m_land);

    * Random infectious cases;
    if i = 0 or m_total = 0 then irand_total = 0;
    else irand_total = int(&screen*ranbin(0,i,m_total));
    if i = 0 or m_air = 0 then irand_air = 0;
    else irand_air = int(&screen*ranbin(0,i,m_air));
    if i = 0 or m_sea = 0 then irand_sea = 0;
    else irand_sea = int(&screen*ranbin(0,i,m_sea));
    if i = 0 or m_land = 0 then irand_land = 0;
    else irand_land = int(&screen*ranbin(0,i,m_land));

  run;

  proc sort data = seir_im_rand;
    by date dayno country;
  run;

  proc sql;
    create table im_daily as
    select date, dayno,
      sum(erand_total) as e_total , sum(erand_air) as e_air,
      sum(erand_sea) as e_sea, sum(erand_land) as e_land,
      sum(irand_total) as i_total , sum(irand_air) as i_air,
      sum(irand_sea) as i_sea, sum(irand_land) as i_land
    from seir_im_rand
    group by date, dayno;
  quit;

  /* Merge export data */
  proc sql;
    create table trans_daily as
    select A.*, B.m_total, B.m_air, B.m_sea, B.m_land, B.confirmed
    from im_daily as A, daily_ex as B
    where A.date =  B.date and A.dayno =  B.dayno;
```

```
quit;

/* Transpose data for array use */
proc transpose data=trans_daily out=im_e_air (drop=_name_) prefix=im_e_air;
  var e_air;
   run;

proc transpose data=trans_daily out=im_e_sea (drop=_name_) prefix=im_e_sea;
  var e_sea;
   run;

proc transpose data=trans_daily out=im_e_land (drop=_name_) prefix=im_e_land;
  var e_land;
   run;

proc transpose data=trans_daily out=im_i_air (drop=_name_) prefix=im_i_air;
  var i_air;
   run;

proc transpose data=trans_daily out=im_i_sea (drop=_name_) prefix=im_i_sea;
  var i_sea;
   run;

proc transpose data=trans_daily out=im_i_land (drop=_name_) prefix=im_i_land;
  var i_land;
   run;

proc transpose data=trans_daily out=m_air (drop=_name_) prefix=m_air;
  var m_air;
   run;

proc transpose data=trans_daily out=m_sea (drop=_name_) prefix=m_sea;
  var m_sea;
   run;

proc transpose data=trans_daily out=m_land (drop=_name_) prefix=m_land;
  var m_land;
   run;

proc transpose data=trans_daily out=confirmed (drop=_name_) prefix=confirmed;
  var confirmed;
   run;

/* calculate the parameters */
%let dno=%sysevalf(&dayno+1);
%let beta=%sysevalf(&repro/&inf);

data daily_arr (keep=dayno s e i r confirmed cum_c b c d ex_e_air
            ex_e_sea ex_e_land ex_i_air ex_i_sea ex_i_land
            im_e_air im_e_sea im_e_land im_i_air im_i_sea im_i_land);
  merge im_e_air im_e_sea im_e_land
        im_i_air im_i_sea im_i_land
        m_air m_sea m_land confirmed;

  array im_e_air_arr(*) im_e_air1 - im_e_air&dno;
  array im_e_sea_arr(*) im_e_sea1 - im_e_sea&dno;
  array im_e_land_arr(*) im_e_land1 - im_e_land&dno;
  array im_i_air_arr(*) im_i_air1 - im_i_air&dno;
  array im_i_sea_arr(*) im_i_sea1 - im_i_sea&dno;
  array im_i_land_arr(*) im_i_land1 - im_i_land&dno;

  array m_air_arr(*) m_air1 - m_air&dno;
  array m_sea_arr(*) m_sea1 - m_sea&dno;
  array m_land_arr(*) m_land1 - m_land&dno;
  array confirmed_arr(*) confirmed1 - confirmed&dno;

  array s_arr(&dayno);
  array e_arr(&dayno) ;
  array i_arr(&dayno) ;
  array r_arr(&dayno) ;
  array dayno_arr(&dayno) ;

  do k = 1 to &dayno;

    if k = 1 then do;
```

```
  dayno_arr(k)=k;
  s_arr(k) =7055071;
  e_arr(k)=0;
  i_arr(k)=0;
  r_arr(k)=0;

end;

else do ;
  dayno_arr(k)=k;

  * Binomial parameters;
  binb=1-exp(-1*((&beta*i_arr(k-1))/7055051));
  binc=1-exp(-1*(1/&lat));
  bind=1-exp(-1*(1/&inf));

  if s_arr(k-1) <= 0 or i_arr(k-1) <= 0 then b=0;
  else b=ranbin(0,s_arr(k-1) ,binb);
  if e_arr(k-1) <= 0 then c=0;
  else c=ranbin(0,e_arr(k-1) ,binc);
  if i_arr(k-1) <= 0 then d=0;
  else d=ranbin(0,i_arr(k-1) ,bind);

  * Random export;
  if e_arr(k-1) <= 0 then do;
    ex_e_air=0;
    ex_e_sea=0;
    ex_e_land=0;
  end;
  else do;
    ex_e_air=ranbin(0,e_arr(k-1),m_air_arr(k));
    ex_e_sea=ranbin(0,e_arr(k-1),m_sea_arr(k));
    ex_e_land= ranbin(0,e_arr(k-1),m_land_arr(k));
  end;

  if i_arr(k-1) <= 0 then do;
    ex_i_air=0;
    ex_i_sea=0;
    ex_i_land=0;
  end;
  else do;
    ex_i_air=ranbin(0,i_arr(k-1),m_air_arr(k));
    ex_i_sea=ranbin(0,i_arr(k-1),m_sea_arr(k));
    ex_i_land=ranbin(0,i_arr(k-1),m_land_arr(k));
  end;

  * Random import;
  im_e_air=im_e_air_arr(k);
  im_e_sea=im_e_sea_arr(k);
  im_e_land=im_e_land_arr(k);
  im_i_air=im_i_air_arr(k);
  im_i_sea=im_i_sea_arr(k);
  im_i_land=im_i_land_arr(k);

  s_arr(k) = s_arr(k-1) - b;
  e_arr(k) = e_arr(k-1) + b - c
          + im_e_air + im_e_sea + im_e_land
          - ex_e_air - ex_e_sea - ex_e_land;
  i_arr(k) = i_arr(k-1) + c - d
          + im_i_air + im_i_sea + im_i_land
          - ex_i_air - ex_i_sea - ex_i_land;
  r_arr(k) = r_arr(k-1) + d;

end;

s=s_arr(k);
e=e_arr(k);
i=i_arr(k);
r=r_arr(k);
dayno=dayno_arr(k);
confirmed=confirmed_arr(k);
cum_c=c+lag(c);
output;
```

```
    end;

  run;

  /* Calculate the MSE */

  %let s=_&si;
  %let st=_1;
  %let fname= &rno&s;
  %let fnamest= &rno&st;

  proc sql;
    create table daily_se as
      select sum(((confirmed-cum_c)**2))/count(confirmed) as mse
      from daily_arr;
  quit;

  data daily_se_&fname;
    set daily_se;
    repro = &repro;
    sim=&si;
  run;

  /* Store the simulation result */
  %if &si > 1 or &rno ^= 1 %then %do;

    proc append base = daily_se_1_1 data = daily_se_&fname;
    run;

  %end;

  /* End of per R simulation */
  %end;

  %let repro= %sysevalf(&repro+&d);
  %let rno = %sysevalf(&rno+1);

/* End of R realization */
%end;

/* Export summary of mean square error for R0 */
proc sql;
  create table mean_mse as
  select mean(mse) as m_mse, repro
  from daily_se_1_1
  group by repro;
quit;

proc print;run;

%put Hello! Finished!;

%mend baselinefit;
```

FPTSIM

```
/*****************************************************************************
/ Program name: FPTSIM
/ Version: 1.0
/ Author: Marc, Biostatistician
/ Study name: H1N1 simulaton study
/ Created date: 01MAY2010
/ Purpose: To simulate the FPT and FHPT from random import
/ Notes:
/ ===========================================================
/ Amendment history:
/ |--Amended date--|--Amended by--|----------Description-----------|
/
/ ===========================================================
*****************************************************************************/
```

```
/* Macro for model simulation:
dir: the result directory
sim: number of realizations
lat: length of latent period
inf: length of infectious period
screen: the proportion of import infectious subjects
gdayres: day number start for travel restriction
imfair: Import restriction fraction for air
imfsea: Import restriction fraction for sea
imfland: Import restriction fraction for land
dayno: number of days generation */

%macro fptsim(dir, sim,lat, inf, screen, gdayres, imfair, imfsea, imfland,dayno);

libname result &dir;

/* Non-local cases with IO=90 and EO=90 */
%seir(90, 90, &lat, &inf, &dayno);

/* Day limit */
data seir_im_dl;
  set seir_im;
  if dayno < &dayno;
run;

/* Start per R simulation */
%do si = 1 %to &sim;

  %put Simulation:  air &imfair sea &imfsea land &imfland for day &gdayres;
  %put Im Marc. The simulation &si is starting;

  /* Daily sum of random import */
  data seir_im_rand;
    set seir_im_dl;

    * Random latent cases;
    if e = 0 or m_total = 0 then erand_total = 0;
    else erand_total = ranbin(0,e,m_total);
    if e = 0 or m_air = 0 then erand_air = 0;
    else erand_air = ranbin(0,e,m_air);
    if e = 0 or m_sea = 0 then erand_sea = 0;
    else erand_sea = ranbin(0,e,m_sea);
    if e = 0 or m_land = 0 then erand_land = 0;
    else erand_land = ranbin(0,e,m_land);

    * Random infectious cases;
    if i = 0 or m_total = 0 then irand_total = 0;
    else irand_total = ranbin(0,i,m_total);
    if i = 0 or m_air = 0 then irand_air = 0;
    else irand_air = ranbin(0,i,m_air);
    if i = 0 or m_sea = 0 then irand_sea = 0;
    else irand_sea = ranbin(0,i,m_sea);
    if i = 0 or m_land = 0 then irand_land = 0;
    else irand_land = ranbin(0,i,m_land);

  run;

  proc sort data = seir_im_rand;
    by date dayno country;
  run;

  proc sql;
    create table sumim as
    select date, dayno,
      sum(erand_total) as e_total , sum(erand_air) as e_air,
      sum(erand_sea) as e_sea, sum(erand_land) as e_land,
      sum(irand_total) as i_total , sum(irand_air) as i_air,
      sum(irand_sea) as i_sea, sum(irand_land) as i_land
    from seir_im_rand
    group by date, dayno;
  quit;

  %put Simulation:  air &imfair sea &imfsea land &imfland for day &gdayres screen &screen;
  %put Im Marc. The simulation &si is starting;
```

```
  /* Calculated the cumulative imported cases */
  data im_list_&si;
    set sumim;
    if 1 < dayno < &gdayres then total_im_inf=e_air + e_sea + e_land
         + i_air + i_sea + i_land;
    if dayno >= &gdayres then
      total_im_inf=int((1-&imfair)*e_air) + int((1-&imfsea)*e_sea)
               + int((1-&imfland)*e_land)+int(&screen*((1-&imfair)*i_air
               + (1-&imfsea)*i_sea + (1-&imfland)*i_land));
    if dayno in (0,1) then cum_im_inf =0;
    else cum_im_inf+total_im_inf;
  run;

  /* Day of FPT and FHPT*/
  data fptcut_&si (keep=dayno cum_im_inf fpt_flag);
    set im_list_&si;
    if cum_im_inf>=1;
    fpt_flag=1;
  run;

  data fpt_&si (keep=fpt);
    set fptcut_&si;
    by fpt_flag;
    if first.fpt_flag;
    fpt=dayno;
  run;

  data fhptcut_&si (keep=dayno cum_im_inf fhpt_flag);
    set im_list_&si;
    if cum_im_inf>=100;
    fhpt_flag=1;
  run;

  data fhpt_&si (keep=fhpt);
    set fhptcut_&si;
    by fhpt_flag;
    if first.fhpt_flag;
    fhpt=dayno;
  run;

  %if &si>1 %then %do;

    proc append base = fpt_1 data = fpt_&si;
    run;

    proc append base = fhpt_1 data = fhpt_&si;
    run;

  %end;

%end;

data result.fpt;
  merge fpt_1 fhpt_1;
run;

%put Hello! Finished!;

%mend fptsim;
```

SIMMODEL

```
/*****************************************************************************
/ Program name: SIMMODEL
/ Version: 1.0
/ Author: Marc, Biostatistician
/ Study name: H1N1 simulaton study
/ Created date: 01MAY2010
/ Purpose: To simulate the scenarios
/ Notes:
/ ============================================================
```

```
/ Amendment history:
/ |--Amended date--|--Amended by--|----------Description-------------|
/
/ ============================================================
*******************************************************************************************/

/* Macro for model simulation:
dir: the result directory
repro: the reporduction number
sim: number of realizations
lat: length of latent period
inf: length of infectious period
screen: the proportion of import infectious subjects
dayres: day number start for travel restriction
dayintervem: day number start for antivirals and hospitalization
pt: proportions of infectious subjects selected for treatment
ph: proportions of infectious subjects selected for hospitalization
pu: proportions of untreated infectious subjects
imfair: Import restriction fraction for air
imfsea: Import restriction fraction for sea
imfland: Import restriction fraction for land
exfair: Emport restriction fraction for air
exfsea: Emport restriction fraction for sea
exfland: Emport restriction fraction for land
gam_t: transition rates from treatment state to removed state
gam_h: transition rates from hospitalization state to removed state
infre: infectiousness reduction for receiving antivirals
dayno: number of days generation */

%macro simmodel(dir, repro, sim,lat, inf, screen, dayres, dayinterven, pt, ph,
                pu,imfair, imfsea, imfland,exfair, exfsea, exfland, gam_t, gam_h,
                infre, dayno);

libname result &dir;

/* Non-local cases with I0=90 and E0=90 */
%seir(90, 90, &lat, &inf, &dayno);

/* Day limit */
data seir_im_dl;
  set seir_im;
  if dayno < &dayno;
run;

/* Start per R simulation */
%do si = 1 %to &sim;

  /* Daily sum of random import */
  data seir_im_rand;
    set seir_im_dl;

    * Random latent cases;
    if e = 0 or m_total = 0 then erand_total = 0;
    else erand_total = ranbin(0,e,m_total);
    if e = 0 or m_air = 0 then erand_air = 0;
    else erand_air = ranbin(0,e,m_air);
    if e = 0 or m_sea = 0 then erand_sea = 0;
    else erand_sea = ranbin(0,e,m_sea);
    if e = 0 or m_land = 0 then erand_land = 0;
    else erand_land = ranbin(0,e,m_land);

    * Random infectious cases;
    if i = 0 or m_total = 0 then irand_total = 0;
    else irand_total = ranbin(0,i,m_total);
    if i = 0 or m_air = 0 then irand_air = 0;
    else irand_air = ranbin(0,i,m_air);
    if i = 0 or m_sea = 0 then irand_sea = 0;
    else irand_sea = ranbin(0,i,m_sea);
    if i = 0 or m_land = 0 then irand_land = 0;
    else irand_land = ranbin(0,i,m_land);

  run;

  proc sort data = seir_im_rand;
    by date dayno country;
```

```
run;

proc sql;
  create table sumim as
  select date, dayno,
    sum(erand_total) as e_total , sum(erand_air) as e_air,
    sum(erand_sea) as e_sea, sum(erand_land) as e_land,
    sum(irand_total) as i_total , sum(irand_air) as i_air,
    sum(irand_sea) as i_sea, sum(irand_land) as i_land
  from seir_im_rand
  group by date, dayno;
quit;

%put Simulation: dayres &dayres  dayinterven &dayinterven air &imfair sea
  &imfsea land &imfland with pt &pt ph &ph pu &pu screen &screen infre &infre;
%put Im Marc. The simulation &si is starting;

/* Merge export data */
proc sql;
  create table trans_daily as
  select A.*, B.m_total, B.m_air, B.m_sea, B.m_land
  from sumim as A, daily_ex as B
  where A.date =  B.date and A.dayno =  B.dayno;
quit;

/* Transpose data for array use */
proc transpose data=trans_daily out=im_e_air (drop=_name_) prefix=im_e_air;
  var e_air;
   run;

proc transpose data=trans_daily out=im_e_sea (drop=_name_) prefix=im_e_sea;
  var e_sea;
   run;

proc transpose data=trans_daily out=im_e_land (drop=_name_) prefix=im_e_land;
  var e_land;
   run;

proc transpose data=trans_daily out=im_i_air (drop=_name_) prefix=im_i_air;
  var i_air;
   run;

proc transpose data=trans_daily out=im_i_sea (drop=_name_) prefix=im_i_sea;
  var i_sea;
   run;

proc transpose data=trans_daily out=im_i_land (drop=_name_) prefix=im_i_land;
  var i_land;
   run;

proc transpose data=trans_daily out=m_air (drop=_name_) prefix=m_air;
  var m_air;
   run;

proc transpose data=trans_daily out=m_sea (drop=_name_) prefix=m_sea;
  var m_sea;
   run;

proc transpose data=trans_daily out=m_land (drop=_name_) prefix=m_land;
  var m_land;
   run;

/* calculate the parameters */
%let dno=%sysevalf(&dayno+1);
%let beta=%sysevalf(&repro/&inf);

data daily_arr (keep=dayno s e i t h r b c d m n p q ex_e_air ex_e_sea
            ex_e_land ex_i_air ex_i_sea ex_i_land im_e_air im_e_sea im_e_land im_i_air
            im_i_sea im_i_land total_im_inf cum_im_inf imfmair imfmsea
            imfmland exfmair exfmsea exfmland
            transediff transidiff binb binc bind binp binq);
  merge im_e_air im_e_sea im_e_land
        im_i_air im_i_sea im_i_land
        m_air m_sea m_land;
```

```
array im_e_air_arr(*) im_e_air1 - im_e_air&dno;
array im_e_sea_arr(*) im_e_sea1 - im_e_sea&dno;
array im_e_land_arr(*) im_e_land1 - im_e_land&dno;
array im_i_air_arr(*) im_i_air1 - im_i_air&dno;
array im_i_sea_arr(*) im_i_sea1 - im_i_sea&dno;
array im_i_land_arr(*) im_i_land1 - im_i_land&dno;

array m_air_arr(*) m_air1 - m_air&dno;
array m_sea_arr(*) m_sea1 - m_sea&dno;
array m_land_arr(*) m_land1 - m_land&dno;

array s_arr(&dayno);
array e_arr(&dayno) ;
array i_arr(&dayno) ;
array r_arr(&dayno) ;
array t_arr(&dayno) ;
array h_arr(&dayno) ;
array dayno_arr(&dayno) ;

/* start loop of day */
do k = 1 to &dayno;

  /* day 0 */
  if k =1 then do;

    dayno_arr(k)=k;
    s_arr(k) =7055071;
    e_arr(k)=0;
    i_arr(k)=0;
    r_arr(k)=0;
    t_arr(k)=0;
    h_arr(k)=0;

  end;

  /* day>0 */
  else do;

    /* days before travel restriction and control measure */
    if k < &dayres and k < &dayinterven then do;

      dayno_arr(k)=k;

      * Binomial parameters;
      binb=1-exp(-1*((&beta*i_arr(k-1))/7055051));
      binc=1-exp(-1*(1/&lat));
      bind=1-exp(-1*(1/&inf));

      if s_arr(k-1) <= 0 or i_arr(k-1) <= 0 then b=0;
      else b=ranbin(0,s_arr(k-1) ,binb);
      if e_arr(k-1) <= 0 then c=0;
      else c=ranbin(0,e_arr(k-1) ,binc);
      if i_arr(k-1) <= 0 then d=0;
      else d=ranbin(0,i_arr(k-1) ,bind);

      * Random export;
      if e_arr(k-1) <= 0 then do;
        ex_e_air=0;
        ex_e_sea=0;
        ex_e_land=0;
      end;
      else do;
        ex_e_air=ranbin(0,e_arr(k-1),m_air_arr(k));
        ex_e_sea=ranbin(0,e_arr(k-1),m_sea_arr(k));
        ex_e_land= ranbin(0,e_arr(k-1),m_land_arr(k));
      end;

      if i_arr(k-1) <= 0 then do;
        ex_i_air=0;
        ex_i_sea=0;
        ex_i_land=0;
      end;
      else do;
        ex_i_air=ranbin(0,i_arr(k-1),m_air_arr(k));
        ex_i_sea=ranbin(0,i_arr(k-1),m_sea_arr(k));
```

```
    ex_i_land=ranbin(0,i_arr(k-1),m_land_arr(k));
  end;

  * Random import;
  im_e_air=im_e_air_arr(k);
  im_e_sea=im_e_sea_arr(k);
  im_e_land= im_e_land_arr(k);
  im_i_air=im_i_air_arr(k);
  im_i_sea=im_i_sea_arr(k);
  im_i_land=im_i_land_arr(k);

  * Transportation operator;
  transediff=im_e_air + im_e_sea + im_e_land
        - ex_e_air - ex_e_sea - ex_e_land;
  transidiff=int(&screen*(im_i_air + im_i_sea + im_i_land))
          - ex_i_air - ex_i_sea - ex_i_land;

  s_arr(k) = s_arr(k-1) - b;
  e_arr(k) = e_arr(k-1) + b - c + transediff;
  i_arr(k) = i_arr(k-1) + c - d
           + transidiff;
  r_arr(k) = r_arr(k-1) + d;
  t_arr(k)=0;
  h_arr(k)=0;

/* Endif of days before travel restriction and control measure */
end;

/* days after travel restriction but before control measure */
if k >= &dayres and k < &dayinterven then do;
  dayno_arr(k)=k;

  imfmair=&imfair;
  imfmsea=&imfsea;
  imfmland=&imfland;
  exfmair=&exfair;
  exfmsea=&exfsea;
  exfmland=&exfland;

  * Binomial parameters;
  binb=1-exp(-1*((&beta*i_arr(k-1))/7055051));
  binc=1-exp(-1*(1/&lat));
  bind=1-exp(-1*(1/&inf));

  if s_arr(k-1) <= 0 or i_arr(k-1) <= 0 then b=0;
  else b=ranbin(0,s_arr(k-1) ,binb);
  if e_arr(k-1) <= 0 then c=0;
  else c=ranbin(0,e_arr(k-1) ,binc);
  if i_arr(k-1) <= 0 then d=0;
  else d=ranbin(0,i_arr(k-1) ,bind);

  * Random export;
  if e_arr(k-1) <= 0 then do;
    ex_e_air=0;
    ex_e_sea=0;
    ex_e_land=0;
  end;
  else do;
    ex_e_air=int((1-exfmair)*ranbin(0,e_arr(k-1),m_air_arr(k)));
    ex_e_sea=int((1-exfmsea)*ranbin(0,e_arr(k-1),m_sea_arr(k)));
    ex_e_land= int((1-exfmland)*ranbin(0,e_arr(k-1),m_land_arr(k)));
  end;

  if i_arr(k-1) <= 0 then do;
    ex_i_air=0;
    ex_i_sea=0;
    ex_i_land=0;
  end;
  else do;
    ex_i_air=int((1-exfmair)*ranbin(0,i_arr(k-1),m_air_arr(k)));
    ex_i_sea=int((1-exfmsea)*ranbin(0,i_arr(k-1),m_sea_arr(k)));
    ex_i_land= int((1-exfmland)*ranbin(0,i_arr(k-1),m_land_arr(k)));
  end;

  * Random import;
```

```
    im_e_air=int((1-imfmair)*im_e_air_arr(k));
    im_e_sea=int((1-imfmsea)*im_e_sea_arr(k));
    im_e_land= int((1-imfmland)*im_e_land_arr(k));
    im_i_air=int((1-imfmair)*im_i_air_arr(k));
    im_i_sea=int((1-imfmsea)*im_i_sea_arr(k));
    im_i_land= int((1-imfmland)*im_i_land_arr(k));

    * Transportation operator;
    transediff=im_e_air + im_e_sea + im_e_land
          - ex_e_air - ex_e_sea - ex_e_land;
    transidiff=int(&screen*(im_i_air + im_i_sea + im_i_land))
           - ex_i_air - ex_i_sea - ex_i_land;

    s_arr(k) = s_arr(k-1) - b;
    e_arr(k) = e_arr(k-1) + b - c + transediff;
    i_arr(k) = i_arr(k-1) + c - d
           + transidiff;
    r_arr(k) = r_arr(k-1) + d;
    t_arr(k)=0;
    h_arr(k)=0;

/* Endif of days after travel restriction but before control measure */
end;

/* days before travel restriction but after control measure */
if k < &dayres and k >= &dayinterven then do;

  imfmair=0;
  imfmsea=0;
  imfmland=0;
  exfmair=0;
  exfmsea=0;
  exfmland=0;

  dayno_arr(k)=k;

  * Binomial parameters;
  binb=1-exp(-1*((&beta*(i_arr(k-1)+(1-&infre)*t_arr(k-1)+h_arr(k-1)))/7055051));
  binc=1-exp(-1*(1/&lat));
  bind=&pu*(1-exp(-1*(1/&inf)));
  binp=1-exp(-1*&gam_t);
  binq=1-exp(-1*&gam_h);

  if s_arr(k-1) <= 0 or i_arr(k-1) <= 0 then b=0;
  else b=ranbin(0,s_arr(k-1) ,binb);
  if e_arr(k-1) <= 0 then c=0;
  else c=ranbin(0,e_arr(k-1) ,binc);

  if i_arr(k-1) <= 0 then d=0;
  else d=ranbin(0,i_arr(k-1) ,bind);
  if i_arr(k-1) <= 0 or &pt <= 0 then m=0;
  else m=ranbin(0,i_arr(k-1), &pt);
  if i_arr(k-1) <= 0 or &ph <= 0 then n=0;
  else n=ranbin(0,i_arr(k-1), &ph);

  if t_arr(k-1) <= 0 or binp <= 0 then p=0;
  else p=ranbin(0,t_arr(k-1),binp);
  if h_arr(k-1) <= 0  or binq <= 0 then q=0;
  else q=ranbin(0,h_arr(k-1),binq);

  * Random export;
  if e_arr(k-1) <= 0 then do;
    ex_e_air=0;
    ex_e_sea=0;
    ex_e_land=0;
  end;
  else do;
    ex_e_air=int((1-exfmair)*ranbin(0,e_arr(k-1),m_air_arr(k)));
    ex_e_sea=int((1-exfmsea)*ranbin(0,e_arr(k-1),m_sea_arr(k)));
    ex_e_land= int((1-exfmland)*ranbin(0,e_arr(k-1),m_land_arr(k)));
  end;

  if i_arr(k-1) <= 0 then do;
    ex_i_air=0;
    ex_i_sea=0;
```

```
    ex_i_land=0;
  end;
  else do;
    ex_i_air=int((1-exfmair)*ranbin(0,i_arr(k-1),m_air_arr(k)));
    ex_i_sea=int((1-exfmsea)*ranbin(0,i_arr(k-1),m_sea_arr(k)));
    ex_i_land= int((1-exfmland)*ranbin(0,i_arr(k-1),m_land_arr(k)));
  end;

  * Random import;
  im_e_air=int((1-imfmair)*im_e_air_arr(k));
  im_e_sea=int((1-imfmsea)*im_e_sea_arr(k));
  im_e_land= int((1-imfmland)*im_e_land_arr(k));
  im_i_air=int((1-imfmair)*im_i_air_arr(k));
  im_i_sea=int((1-imfmsea)*im_i_sea_arr(k));
  im_i_land= int((1-imfmland)*im_i_land_arr(k));

  * Transportation operator;
  transediff=im_e_air + im_e_sea + im_e_land
        - ex_e_air - ex_e_sea - ex_e_land;
  transidiff=int(&screen*(im_i_air + im_i_sea + im_i_land))
          - ex_i_air - ex_i_sea - ex_i_land;

  s_arr(k) = s_arr(k-1) - b;
  e_arr(k) = e_arr(k-1) + b - c + transediff;
  i_arr(k) = i_arr(k-1) + c - d
          + transidiff
          - m - n;
  t_arr(k) = t_arr(k-1) + m - p;
  h_arr(k) = h_arr(k-1) + n - q;
  r_arr(k) = r_arr(k-1) + d + p + q;

/* Endif days before travel restriction but after control measure */
end;

/* days after travel restriction and control measure */
if k >= &dayres and k >= &dayinterven then do;

  imfmair=&imfair;
  imfmsea=&imfsea;
  imfmland=&imfland;
  exfmair=&exfair;
  exfmsea=&exfsea;
  exfmland=&exfland;

  dayno_arr(k)=k;

  * Binomial parameters;
  binb=1-exp(-1*((&beta*(i_arr(k-1)+(1-&infre)*t_arr(k-1)+h_arr(k-1)))/7055051));
  binc=1-exp(-1*(1/&lat));
  bind=&pu*(1-exp(-1*(1/&inf)));
  binp=1-exp(-1*&gam_t);
  binq=1-exp(-1*&gam_h);

  if s_arr(k-1) <= 0 or i_arr(k-1) <= 0 then b=0;
  else b=ranbin(0,s_arr(k-1) ,binb);
  if e_arr(k-1) <= 0 then c=0;
  else c=ranbin(0,e_arr(k-1) ,binc);

  if i_arr(k-1) <= 0 then d=0;
  else d=ranbin(0,i_arr(k-1) ,bind);
  if i_arr(k-1) <= 0 or &pt <= 0 then m=0;
  else m=ranbin(0,i_arr(k-1), &pt);
  if i_arr(k-1) <= 0 or &ph <= 0 then n=0;
  else n=ranbin(0,i_arr(k-1), &ph);

  if t_arr(k-1) <= 0 or binp <= 0 then p=0;
  else p=ranbin(0,t_arr(k-1),binp);
  if h_arr(k-1) <= 0  or binq <= 0 then q=0;
  else q=ranbin(0,h_arr(k-1),binq);

  * Random export;
  if e_arr(k-1) <= 0 then do;
    ex_e_air=0;
    ex_e_sea=0;
    ex_e_land=0;
```

```
        end;
        else do;
          ex_e_air=int((1-exfmair)*ranbin(0,e_arr(k-1),m_air_arr(k)));
          ex_e_sea=int((1-exfmsea)*ranbin(0,e_arr(k-1),m_sea_arr(k)));
          ex_e_land= int((1-exfmland)*ranbin(0,e_arr(k-1),m_land_arr(k)));
        end;

        if i_arr(k-1) <= 0 then do;
          ex_i_air=0;
          ex_i_sea=0;
          ex_i_land=0;
        end;
        else do;
          ex_i_air=int((1-exfmair)*ranbin(0,i_arr(k-1),m_air_arr(k)));
          ex_i_sea=int((1-exfmsea)*ranbin(0,i_arr(k-1),m_sea_arr(k)));
          ex_i_land= int((1-exfmland)*ranbin(0,i_arr(k-1),m_land_arr(k)));
        end;

        * Random import;
        im_e_air=int((1-imfmair)*im_e_air_arr(k));
        im_e_sea=int((1-imfmsea)*im_e_sea_arr(k));
        im_e_land= int((1-imfmland)*im_e_land_arr(k));
        im_i_air=int((1-imfmair)*im_i_air_arr(k));
        im_i_sea=int((1-imfmsea)*im_i_sea_arr(k));
        im_i_land= int((1-imfmland)*im_i_land_arr(k));

        * Transportation operator;
        transediff=im_e_air + im_e_sea + im_e_land
              - ex_e_air - ex_e_sea - ex_e_land;
        transidiff=int(&screen*(im_i_air + im_i_sea + im_i_land))
                - ex_i_air - ex_i_sea - ex_i_land;

        s_arr(k) = s_arr(k-1) - b;
        e_arr(k) = e_arr(k-1) + b - c + transediff;
        i_arr(k) = i_arr(k-1) + c - d
                + transidiff
                - m - n;
        t_arr(k) = t_arr(k-1) + m - p;
        h_arr(k) = h_arr(k-1) + n - q;
        r_arr(k) = r_arr(k-1) + d + p + q;

      /* Endif days after travel restriction and control measure */
      end;

    /* Enddo days>0 */
    end;

    s=s_arr(k);
    e=e_arr(k);
    i=i_arr(k);
    t=t_arr(k);
    h=h_arr(k);
    r=r_arr(k);
    total_im_inf=im_e_air + im_e_sea + im_e_land + im_i_air + im_i_sea + im_i_land;
    if k in (0,1) then cum_im_inf =0;
    else cum_im_inf+total_im_inf;
    dayno=dayno_arr(k);
    output;

  /* End of loop day */
  end;

run;

data result.sim_&si;
  set daily_arr;
run;

/* End of per R simulation */
%end;

%put Hello! Finished!;

%mend simmodel;
```

```
/* Macro for resulting statistics:
dir: the result directory
sim: The simulated times */
%macro simresult(dir, sim);

libname result &dir;

/* Start per R simulation */
%do si = 1 %to &sim;

  * Day of FPT;
  data firstcut_&si (keep=dayno cum_im_inf fpt_flag);
    set result.sim_&si;
    if cum_im_inf>=1;
    fpt_flag=1;
  run;

  data fpt_&si (keep=arr1day);
    set firstcut_&si;
    by fpt_flag;
    if first.fpt_flag;
    arr1day=dayno;
  run;

  * Day of FHPT;
  data first100cut_&si (keep=dayno cum_im_inf fhpt_flag);
    set result.sim_&si;
    if cum_im_inf>=100;
    fhpt_flag=1;
  run;

  data fhpt_&si (keep=arr100day);
    set first100cut_&si;
    by fhpt_flag;
    if first.fhpt_flag;
    arr100day=dayno;
  run;

  * Peak attack rate (%);
  proc sql;
    create table pi_&si as
    select round((100*max(b))/7055051,0.01) as peak_rate
    from result.sim_&si;
  quit;

  proc sort data = pi_&si nodup;
    by peak_rate;
  run;

  * Peak times (weeks);
  proc sql;
    create table pt_&si as
    select round((dayno)/7,0.1) as peak_time
    from result.sim_&si
    where b = (select max(b) from result.sim_&si);
  quit;

  proc sort data = pt_&si nodup;
    by peak_time;
  run;

  * Cumumlative attack rate (%) by months;
  proc sql;
    create table ci4_&si as
    select round((100*sum(b))/7055051,0.1) as cum_rate
    from result.sim_&si
    where dayno <=120;
  quit;

  proc sql;
    create table ci5_&si as
    select round((100*sum(b))/7055051,0.1) as cum_rate
    from result.sim_&si
    where dayno <=150;
  quit;
```

```
proc sql;
  create table ci6_&si as
  select round((100*sum(b))/7055051,0.1) as cum_rate
  from result.sim_&si
  where dayno <=180;
quit;

proc sql;
  create table ci7_&si as
  select round((100*sum(b))/7055051,0.1) as cum_rate
  from result.sim_&si
  where dayno <=210;
quit;

proc sql;
  create table ci8_&si as
  select round((100*sum(b))/7055051,0.1) as cum_rate
  from result.sim_&si
  where dayno <=240;
quit;

proc sql;
  create table ci9_&si as
  select round((100*sum(b))/7055051,0.1) as cum_rate
  from result.sim_&si
  where dayno <=270;
quit;

proc sql;
  create table ci10_&si as
  select round((100*sum(b))/7055051,0.1) as cum_rate
  from result.sim_&si
  where dayno <=300;
quit;

proc sql;
  create table ci11_&si as
  select round((100*sum(b))/7055051,0.1) as cum_rate
  from result.sim_&si
  where dayno <=330;
quit;

proc sql;
  create table ci12_&si as
  select round((100*sum(b))/7055051,0.1) as cum_rate
  from result.sim_&si
  where dayno <=360;
quit;

proc sql;
  create table ciall_&si as
  select round((100*sum(b))/7055051,0.1) as cum_rate
  from result.sim_&si
  where dayno <=600;
quit;

%if &si>1 %then %do;

  proc append base = fpt_1 data = fpt_&si;
  run;

  proc append base = fhpt_1 data = fhpt_&si;
  run;

  proc append base = pi_1 data = pi_&si;
  run;

  proc append base = pt_1 data = pt_&si;
  run;

  proc append base = ci4_1 data = ci4_&si;
  run;

  proc append base = ci5_1 data = ci5_&si;
```

```
    run;

    proc append base = ci6_1 data = ci6_&si;
    run;

    proc append base = ci7_1 data = ci7_&si;
    run;

    proc append base = ci8_1 data = ci8_&si;
    run;

    proc append base = ci9_1 data = ci9_&si;
    run;

    proc append base = ci10_1 data = ci10_&si;
    run;

    proc append base = ci11_1 data = ci11_&si;
    run;

    proc append base = ci12_1 data = ci12_&si;
    run;

    proc append base = ciall_1 data = ciall_&si;
    run;


  %end;

%end;

/* Summary of results */
/* FPT */
proc means data=fpt_1 noprint;
   var arr1day;
   output out = fpt_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data fpt_print (keep = name N mean p1 p50 pu);
  set fpt_print;
  p1=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = 'FPT                          ';
run;

/* FHPT */
proc means data=fhpt_1 noprint;
   var arr100day;
   output out = fhpt_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data fhpt_print (keep = name N mean p1 p50 pu);
  set fhpt_print;
  p1=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = 'FHPT';
run;

/* Peak AR */
proc means data=pi_1 noprint;
   var peak_rate;
   output out = pi_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data pi_print (keep = name N mean p1 p50 pu);
  set pi_print;
  p1=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = 'Peak AR (%)';
run;

/* Peak time */
proc means data=pt_1 noprint;
   var peak_time;
   output out = pt_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
```

```
run;

data pt_print (keep = name N mean pl p50 pu);
  set pt_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = 'Peak times (weeks)';
run;

/* Cum AR by months */
proc means data=ci4_1 noprint;
   var cum_rate;
   output out = ci4_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data ci4_print (keep = name N mean pl p50 pu);
  set ci4_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = '4 months cum AR (%)';
run;

proc means data=ci5_1 noprint;
   var cum_rate;
   output out = ci5_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data ci5_print (keep = name N mean pl p50 pu);
  set ci5_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = '5 months cum AR (%)';
run;

proc means data=ci6_1 noprint;
   var cum_rate;
   output out = ci6_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data ci6_print (keep = name N mean pl p50 pu);
  set ci6_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = '6 months cum AR (%)';
run;

proc means data=ci7_1 noprint;
   var cum_rate;
   output out = ci7_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data ci7_print (keep = name N mean pl p50 pu);
  set ci7_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = '7 months cum AR (%)';
run;

proc means data=ci8_1 noprint;
   var cum_rate;
   output out = ci8_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data ci8_print (keep = name N mean pl p50 pu);
  set ci8_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = '8 months cum AR (%)';
run;

proc means data=ci9_1 noprint;
   var cum_rate;
   output out = ci9_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;
```

```
data ci9_print (keep = name N mean pl p50 pu);
  set ci9_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = '9 months cum AR (%)';
run;

proc means data=ci10_1 noprint;
   var cum_rate;
   output out = ci10_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data ci10_print (keep = name N mean pl p50 pu);
  set ci10_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = '10 months cum AR (%)';
run;

proc means data=ci11_1 noprint;
   var cum_rate;
   output out = ci11_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data ci11_print (keep = name N mean pl p50 pu);
  set ci11_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = '11 months cum AR (%)';
run;

proc means data=ci12_1 noprint;
   var cum_rate;
   output out = ci12_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data ci12_print (keep = name N mean pl p50 pu);
  set ci12_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = '12 months cum AR (%)';
run;

proc means data=ciall_1 noprint;
   var cum_rate;
   output out = ciall_print N=N mean=mean p1=p1 p5=p5 p50=p50 p95=p95 p99=p99;
run;

data ciall_print (keep = name N mean pl p50 pu);
  set ciall_print;
  pl=(p1+p5)/2;
  pu=(p95+p99)/2;
  name = 'End of epidemic  cum AR (%)';
run;

data summary;
  set fpt_print fhpt_print pi_print pt_print ci4_print ci5_print ci6_print
      ci7_print ci8_print ci9_print ci10_print ci11_print ci12_print ciall_print;
run;

proc print ;run;

%mend simresult;
```

SIMCI

```
/*****************************************************************************
/ Program name: SIMCI
/ Version: 1.0
/ Author: Marc, Biostatistician
/ Study name: H1N1 simulaton study
```

```
/ Created date: 01MAY2010
/ Purpose: To simulate the confidence intervals for scenerios
/ Notes:
/ ============================================================
/ Amendment history:
/ |--Amended date--|--Amended by--|----------Description----------|
/
/ ============================================================
*******************************************************************************************/

/* Suppress the notes of log window */
options nonotes;

/* Macro for finding the confident interval by day:
idir: the dataset directory with 100 epidemic series
dir: the result directory
fname: name of the file
repro: reproduction number shown in log window */
%macro simci(idir, dir, fname, repro);

libname init &idir;
libname result &dir;

/* Start per R simulation */
%do day = 1 %to 600;

    %let st = _1;
    %let sidst=&day&st;

  %do si = 1 %to 100;

    %let sid = _&si;
    %let sidno=&day&sid;

    data resultd_&sidno (keep = dayno b);
      set init.sim_&si;
      if dayno = &day;
    run;

    %if &si>1 %then %do;

      proc append base = resultd_&day&st data = resultd_&sidno;
      run;

    %end;

  %end;

  proc sort data = resultd_&day&st;
    by b;
  run;

  data resultlag_&day&st;
    set resultd_&day&st;
    lagb=int((b+lag(b))/2);
  run;

  data ll_&day(keep=dayno lagb rename= lagb = lb)
       ul_&day(keep=dayno lagb rename= lagb = ub);
    set resultlag_&day&st;
    by b;
    if _n_= 3 then output ll_&day;
    if _n_ = 98 then output ul_&day;
  run;

  proc sql;
  create table ml_&day as
  select int(mean(b)) as mb, dayno
  from resultd_&day&st
  group by dayno;
  quit;


  %if &day>1 %then %do;
```

```
      proc append base = ll_1 data = ll_&day;
      run;

      proc append base = ul_1 data = ul_&day;
      run;

      proc append base = ml_1 data = ml_&day;
      run;

  %end;

%PUT write &repro file name &fname day &day;

%end;

data result.&fname;
  merge ul_1 ml_1 ll_1;
  by dayno;
run;

%PUT Finish &fname;

%mend simc1;
```

Appendix B

R programs for studying the impact of travel restriction for the *source* country

The programs were built in software R 2.12.1. Here are the following programs and their functions:

- COREFILE: A core file to call all sub-files.
- MCMCfunction: A function for MCMC metropolis random walk
- MCMCplot: A function to generate the time series trace plot, autocorrelation plot, and probability density plot from simulated data
- OBSDIFF: A function to calculated the absolute difference between exported days
- SIMULATION: A simulation study to test the MCMC metropolis random walk
- SIROBS: Application of the MCMC method on the epidemic data and a sensitivity analysis

- SIMINCIDENCE: To simulate incidence curve
- SIMFETPDF: To simulate the probability distributions of FET and F10ET
- SIMEPFET: To simulate expected daily probability distributions of FET
- SIMFETVSM: To simulate FET against different daily rates of travel
- SIMFETVSR0: To simulate FET against different R0
- SIMCOUNTRY: To simulate number of countries received infected cases by day

COREFILE

```
################################################################################
# Program name: COREFILE
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: A core file for the study (Chapter 4)
# Notes:
################################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description-------------|
#
################################################################################

#############################
### Function of the study ###
#############################

### Function of MCMC ###
source("C:\\MCMCfunction")

### Function of MCMC plots ###
source("C:\\MCMCplot")

### Function to validate estimated values ###
source("C:\\OBSDIFF")

########################
### Simulation study ###
########################

### Simulation testing study ###
source("C:\\SIMULATION")

##################
### Case study ###
##################

### MCMC from epidemic data and sensitivity analysis ###
source("C:\\SIROBS")

# Simulate incidence curve
source("C:\\SIMINCIDENCE")

# Simulate the probability distributions of FET and F10ET
source("C:\\SIMFETPDF")

# Simulate expected daily probability distributions of FET
source("C:\\SIMEPFET")

# Simulate FET against different daily rates of travel
source("C:\\SIMFETVSM")

# Simulate FET against different R0
source("C:\\SIMFETVSR0")

# Simulate number of countries received infected cases by day
source("C:\\SIMCOUNTRY")
```

MCMCfunction

```
################################################################################
# Program name: MCMCfunction
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: A function for MCMC metropolis random walk to obtain the
```

```
#   posterior distribution
# Notes:
#------------------------------------------------------------------------------
# Amendment history:
# |--Amended date--|--Amended by--|----------Description------------|
#
#------------------------------------------------------------------------------

#########################################################
### Log likelihood for the Stochastic dynamic model ###
#########################################################
logpost=function(u,N,R0ad,TIp,tau,rate){

  # stochastic SIR model #

  # adapt tau
  I0=7
  datalength=length(u[,1])
  sirp=matrix(0,tau+1,7)
  sirp[1,]=c(0,N,I0,0,0,0,0)
  for (i in 1:tau){
    pi=1-exp(-1*(R0ad*sirp[i,3])/N)
    x=round(pi*sirp[i,2], 0)
    S=max(0,sirp[i,2]-x)
    y=round(sirp[i,3]*TIp,0)
    I=max(0,sirp[i,3]+x-y)
    R=max(0,sirp[i,4]+y)
    sirp[i+1,]=c(i,S,I,R,0,x,y)
  }
  Stau=sirp[tau+1,2]
  Itau=sirp[tau+1,3]

  sir =matrix(0,datalength+1,7)
  sir[1,]=c(tau,Stau,Itau,0,0,0,0)

  #likelihood matrix
  for (i in 1:datalength){
    pi=1-exp(-1*(R0ad*sir[i,3])/N)
    x=trunc(u[i,2]/rate)
    S=max(0,sir[i,2]-x)
    y=round(sir[i,3]*TIp,0)
    I=max(0,sir[i,3]+x-y)
    R=max(0,sir[i,4]+y)

    # likelihood
    loglikeli=-log(dbinom(x,sir[i,2],pi,0))
    sir[i+1,]=c(tau+i,S,I,R,loglikeli,x,y)
  }
  sumlikesir=sum(sir[2:(datalength+1),5])
  return (sumlikesir)
}

MCMCepi=function(u,N,burnin,M){

  # Set prior information
  # Initial R0, TI, rate, a0
  R0j=1.5
  TI=3

  R0adj=R0j/TI
  LR0ad=1/TI
  UR0ad=3/TI

  TIpj=1-exp(-1*(1/TI))

  ratej=0.2
  Lrate=0.001
  Urate=0.8

  tauj=20
  Ltau=1
  Utau=120

  draws=c(0,R0adj,TIpj,tauj,ratej,0,0,0,0)
```

```
  # time record
  system.time(
    for (iter in 1:(burnin+M-1)){

      # random walk single move #
      # rules: accept with prob min(1, A) #

      # Initiate acceptance rate
      R0adacpt=0
      TIpacpt=0
      tauacpt=0
      rateacpt=0

      # random step size
      VR0ad=(runif(1)/30)*(UR0ad-LR0ad)
      Vtau=(runif(1)/5)*(Utau-Ltau)

      # update tau, Stau, Itau
      taunew=round(rnorm(1,tauj,Vtau),0)
      if (taunew>Ltau & taunew<Utau){
        tauA=logpost(u,N,R0adj,TIpj,taunew,ratej)
        if (is.finite(tauA)){
          A=exp(-1*(tauA-logpost(u,N,R0adj,TIpj,tauj,ratej)))
          if (runif(1) < min(1,A)) {
            tauj=taunew
            tauacpt=1
          }
        }
      }

      # update R0
      R0adnew=rnorm(1,R0adj,VR0ad)
      if (R0adnew>LR0ad & R0adnew<UR0ad){
        R0adA=logpost(u,N,R0adnew,TIpj,tauj,ratej)
        if (is.finite(R0adA)){
          A=exp(-1*(R0adA-logpost(u,N,R0adj,TIpj,tauj,ratej)))
          if (runif(1) < min(1,A)) {
            R0adj=R0adnew
            R0adacpt=1
          }
        }
      }

      # store the simulations #
      draws = rbind(draws,c(iter,R0adj,TIpj,tauj,ratej,
      R0adacpt,TIpacpt,tauacpt,rateacpt))
    }
  )
  return(draws)
}
```

MCMCplot

```
##############################################################################
# Program name: MCMCplot
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: A function to generate the time series trace plot,
#   autocorrelation plot, and probability density plot
#   from simulated data
# Notes:
##############################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description-------------|
#
##############################################################################

MCMCplot=function(drawsdat){
```

```
# 3 by 3 plots
par(mfrow=c(4,2))

# R0
plot(drawsdat[,2],xlab="",ylab="",main="R0",type="l",axes=F);axis(1);axis(2)
acf(drawsdat[,2],xlab="lag",ylab="",main="R0",axes=F);axis(1);axis(2)
hist(drawsdat[,2],freq = FALSE,breaks =300, main="R0",xlab="",ylab="")
plot(density(drawsdat[,2]),main="R0",xlab="",ylab="")

# tau
plot(drawsdat[,4],xlab="",ylab="",main="tau",type="l",axes=F);axis(1);axis(2)
acf(drawsdat[,4],xlab="lag",ylab="",main="tau",axes=F);axis(1);axis(2)
hist(drawsdat[,4],freq = FALSE,breaks =300,main="tau",xlab="",ylab="")
plot(density(drawsdat[,4],bw=0.4),main="tau",xlab="",ylab="")
}
```

OBSDIFF

```
##############################################################################
# Program name: OBSDIFF
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: A function to calculated the absolute difference between
#   exported days
# Notes:
##############################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description-------------|
#
##############################################################################

##############################################################################
### Function to calculated the absolute difference between exported days ###
##############################################################################
obsdiff=function(firstdata,m,N,R0ad,TIp,tau){

  I0=7
  datalength=200
  sir =matrix(0,datalength+1,6)
  sir[1,]=c(0,N,I0,0,0,0)

  # SIR matrix
  for (i in 1:datalength){
    pi=1-exp(-1*(R0ad*sir[i,3])/N)
    x=round(pi*sir[i,2], 0)
    S=max(0,sir[i,2]-x)
    y=round(sir[i,3]*TIp,0)
    I=max(0,sir[i,3]+x-y)
    R=max(0,sir[i,4]+y)
    sir[i+1,]=c(i,S,I,R,x,y)
  }

  mlength=length(firstdata)
  itern=200
  arrdaydrawk=NULL

  for (iter in 1:itern){
    arrdayk=NULL
    for (i in 1:mlength){
      for (j in 1:datalength){
        #pm=1-exp((-R0ad*m[i]*sir[j+1,3])/(N))
        pm=1-exp((-R0ad*sir[j+1,3])/(N))
        #first=rbinom(1, 1,pm)
        first=rbinom(1,m[i],pm)
        if (first>0) {
          arrday=sir[j+1,1]-tau
          break
        }
```

```
        arrday=999
      }
      arrdayk = rbind(arrdayk,arrday)
    }
    arrdaydrawk=cbind(arrdaydrawk,arrdayk)
  }
  arrdaymed=round(apply(arrdaydrawk,1,median),0)

  diff=abs(arrdaymed-obsfirstdata)
  absdiff = sum(diff)
  return (absdiff)
}

# Notes: geometric dist can be changed to binomial dist (result is similar)
```

SIMULATION

```
################################################################################
# Program name: SIMULATION
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: A simulation to test the MCMC metropolis random walk
# Notes:
################################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description------------|
#
################################################################################

###################################
### Prepare simulation dataset ###
###################################
# stochastic SIR model function to obtain the simulation data
SIRsim=function(N,TI,R0,tau,rate,datalength,bufferperiod){

  I0=7
  sir =matrix(0,tau+datalength+bufferperiod+1,6)

  # day 0
  sir[1,]=c(0,N,I0,0,0,0)

  for (i in 1:(tau+datalength+bufferperiod)){
    pi=1-exp((-R0*sir[i,3])/(TI*N))

    # expected incidence
    x=round(pi*sir[i,2],0)
    S=max(0,sir[i,2]-x)
    pr=1-exp(-1*(1/TI))
    y=round(pr*sir[i,3],0)
    I=max(0,sir[i,3]+x-y)
    R=max(0,sir[i,4]+y)
    sir[i+1,]=c(i,S,I,R,x,y)
  }

  # function to generate misreported incidence data
  u=seq(1,datalength,by=1)
  for (i in 1:datalength){
    u[i]=round(sir[tau+i+1,5]*rate,0)
  }
  ku=seq(1,datalength,by=1)
  udata=cbind(ku,u)
  return(udata)
}

### Simulation settings: P1, P2, P3 ###
# Scenario P1 #
P1datalength=30
P1bufferperiod=60
P1N=1000000
```

```
P1R0=1.2
P1TI=3
P1tau=28
P1rate=0.3
P1u=SIRsim(P1N,P1TI,P1R0,P1tau,P1rate,P1datalength,P1bufferperiod)

# Scenario P2 #
P2datalength=30
P2bufferperiod=60
P2N=1000000
P2R0=1.5
P2TI=3
P2tau=16
P2rate=0.15
P2u=SIRsim(P2N,P2TI,P2R0,P2tau,P2rate,P2datalength,P2bufferperiod)

# Scenario P3 #
P3datalength=30
P3bufferperiod=60
P3N=1000000
P3R0=1.8
P3TI=3
P3tau=7
P3rate=0.05
P3u=SIRsim(P3N,P3TI,P3R0,P3tau,P3rate,P3datalength,P3bufferperiod)

##########################
### Simulation testing ###
##########################

### P1 ###
# Set up burn-in period and iteration numbers M
P1burnin=10000
P1M=100000
# 1min for 6000 iterations
system.time({
P1para=MCMCepi(P1u,1000000,P1burnin,P1M)
})

# Acceptance rate
acptrates=matrix(0,1,4)
for (i in 1:4){
acptrates[,i]=sum(P1para[,(5+i)])/(P1burnin+P1M)
}
acptrates

# Eliminate burn-in period
P1parab=P1para[(P1burnin+1):(P1burnin+P1M),]

# MCMC statistics
MCMCstat=matrix(0,4,5)
for (i in 1:4){
  mean=mean(P1parab[,i+1])
  median=median(P1parab[,i+1])
  a=table(round(P1parab[,i+1],4))
  mode=as.numeric(names(a)[a==max(a)])
  sd=sqrt(var(P1parab[,i+1]))
  L=quantile(P1parab[,i+1],0.025)
  U=quantile(P1parab[,i+1],0.975)
  MCMCstat[i,]=c(mean,median,sd,L,U)
}
MCMCstat[1,]=round(MCMCstat[1,],3)
MCMCstat[2,]=round(MCMCstat[2,],2)
MCMCstat[3,]=round(MCMCstat[3,],0)
MCMCstat[4,]=round(MCMCstat[4,],4)
MCMCstat

# The MCMC diagnostic plots
MCMCplot(P1parab)

# Scatter plot
pairs(~P1parab[,2]+P1parab[,4],labels=c("R0","tau"),
  main="Scatterplot matrix")

# Save MCMC data
```

```
save(P1para,
file = "C:\\P1.RData")


### P2 ###
# Set up burn-in period and iteration numbers M
P2burnin=10000
P2M=100000
# 1min for 6000 iterations
system.time({
P2para=MCMCepi(P2u,1000000,P2burnin,P2M)
})

# Acceptance rate
acptrates=matrix(0,1,4)
for (i in 1:4){
acptrates[,i]=sum(P2para[,(5+i)])/(P2burnin+P2M)
}
acptrates

# Eliminate burn-in period
P2parab=P2para[(P2burnin+1):(P2burnin+P2M),]

# MCMC statistics
MCMCstat=matrix(0,4,5)
for (i in 1:4){
  mean=mean(P2parab[,i+1])
  median=median(P2parab[,i+1])
  a=table(round(P2parab[,i+1],4))
  mode=as.numeric(names(a)[a==max(a)])
  sd=sqrt(var(P2parab[,i+1]))
  L=quantile(P2parab[,i+1],0.025)
  U=quantile(P2parab[,i+1],0.975)
  MCMCstat[i,]=c(mean,median,sd,L,U)
}
MCMCstat[1,]=round(MCMCstat[1,],3)
MCMCstat[2,]=round(MCMCstat[2,],2)
MCMCstat[3,]=round(MCMCstat[3,],0)
MCMCstat[4,]=round(MCMCstat[4,],4)
MCMCstat

# The MCMC diagnostic plots
MCMCplot(P2parab)

# Scatter plot
pairs(~P2parab[,2]+P2parab[,4],labels=c("R0","tau"),
  main="Scatterplot matrix")

# Save MCMC data
save(P2para,
file = "C:\\P2.RData")

### P3 ###
# Set up burn-in period and iteration numbers M
P3burnin=10000
P3M=100000
# 1min for 6000 iterations
system.time({
P3para=MCMCepi(P3u,1000000,P3burnin,P3M)
})

# Acceptance rate
acptrates=matrix(0,1,4)
for (i in 1:4){
acptrates[,i]=sum(P3para[,(5+i)])/(P3burnin+P3M)
}
acptrates

# Eliminate burn-in period
P3parab=P3para[(P3burnin+1):(P3burnin+P3M),]

# MCMC statistics
MCMCstat=matrix(0,4,5)
for (i in 1:4){
  mean=mean(P3parab[,i+1])
```

```
  median=median(P3parab[,i+1])
  a=table(round(P3parab[,i+1],4))
  mode=as.numeric(names(a)[a==max(a)])
  sd=sqrt(var(P3parab[,i+1]))
  L=quantile(P3parab[,i+1],0.025)
  U=quantile(P3parab[,i+1],0.975)
  MCMCstat[i,]=c(mean,median,sd,L,U)
}
MCMCstat[1,]=round(MCMCstat[1,],3)
MCMCstat[2,]=round(MCMCstat[2,],2)
MCMCstat[3,]=round(MCMCstat[3,],0)
MCMCstat[4,]=round(MCMCstat[4,],4)
MCMCstat

# The MCMC diagnostic plots
MCMCplot(P3parab)

# Scatter plot
pairs(~P3parab[,2]+P3parab[,4],labels=c("R0","tau"),
  main="Scatterplot matrix")

# Save MCMC data
save(P3para,
file = "C:\\P3.RData")
```

SIROBS

```
##############################################################################
# Program name: SIROBS
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: MCMC from epidemic data and sensitivity analysis
# Notes:
##############################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description-------------|
#
##############################################################################

########################
### Epidemic dataset ###
########################

# Mexi gov 14/3-30/4 (48)
u=c(2,1,3,1,2,3,3,4,4,5,7,2,1,2,5,7,4,10,10,9,4,4,11,5,7,4,4,4,11,17,
  26,20,12,19,26,33,44,107,114,154,226,280,318,398,411,304,280,227)

datalength=length(u)

# corresponding day
ku=seq(1,datalength,by=1)

udata=cbind(ku,u)


####################
### Arrival data ###
####################

# day 1 = 14/3
# First case arrival day
obsfirstdata=c(46,46,46,47,48,48,49,51,52,55)

# Travel rate
mi=c(101313,65724,20513,16950,35772,27640,61960,24535,15090,24609)
obsmdaily=trunc(mi/61)


########################
```

```
### MCMC estimation ###
#######################

# Set up burn-in period and iteration numbers M
burnin=5000
M=20000
system.time({
obspara=MCMCepi(udata,106682518,burnin,M)
})

# Acceptance rate
acptrates=matrix(0,1,4)
for (i in 1:4){
acptrates[,i]=sum(obspara[,(5+i)])/(burnin+M)
}
acptrates

# Eliminate burn-in period
obsparab=obspara[(burnin+1):(burnin+M),]

# MCMC statistics
MCMCstat=matrix(0,4,5)
for (i in 1:4){
  mean=mean(obsparab[,i+1])
  median=median(obsparab[,i+1])
  a=table(round(obsparab[,i+1],4))
  mode=as.numeric(names(a)[a==max(a)])
  sd=sqrt(var(obsparab[,i+1]))
  L=quantile(obsparab[,i+1],0.025)
  U=quantile(obsparab[,i+1],0.975)
  MCMCstat[i,]=c(mean,median,sd,L,U)
}
MCMCstat[1,]=round(MCMCstat[1,],3)
MCMCstat[2,]=round(MCMCstat[2,],2)
MCMCstat[3,]=round(MCMCstat[3,],0)
MCMCstat[4,]=round(MCMCstat[4,],4)
MCMCstat

# The MCMC diagnostic plots
MCMCplot(obsparab)

# Save MCMC data
save(obspara,
file = "C:\\obs.RData")

# Validate the estimated parameters
R0mcmc=4
TIpmcmc=1-exp(-1*(1/3))
taumcmc=20
obsdiff(obsfirstdata,obsmdaily,106682518,R0mcmc,TIpmcmc,taumcmc)

############################
### Sensitivity analysis ###
############################
##########
# TI=2.5 #
##########
# Set up burn-in period and iteration numbers M
sa1burnin=10000
sa1M=100000
system.time({
sa1para=MCMCepi(udata,106682518,sa1burnin,sa1M)
})

# Acceptance rate
acptrates=matrix(0,1,4)
for (i in 1:4){
acptrates[,i]=sum(sa1para[,(5+i)])/(sa1burnin+sa1M)
}
acptrates

# Eliminate burn-in period
sa1parab=sa1para[(sa1burnin+1):(sa1burnin+sa1M),]

# MCMC statistics
```

```
MCMCstat=matrix(0,4,5)
for (i in 1:4){
  mean=mean(sa1parab[,i+1])
  median=median(sa1parab[,i+1])
  a=table(round(sa1parab[,i+1],4))
  mode=as.numeric(names(a)[a==max(a)])
  sd=sqrt(var(sa1parab[,i+1]))
  L=quantile(sa1parab[,i+1],0.025)
  U=quantile(sa1parab[,i+1],0.975)
  MCMCstat[i,]=c(mean,median,sd,L,U)
}
MCMCstat[1,]=round(MCMCstat[1,],3)
MCMCstat[2,]=round(MCMCstat[2,],2)
MCMCstat[3,]=round(MCMCstat[3,],0)
MCMCstat[4,]=round(MCMCstat[4,],4)
MCMCstat

# The MCMC diagnostic plots
MCMCplot(sa1parab)

# Save MCMC data
save(sa1para,
file = "C:\\sa1.RData")

##########
# TI=3.5 #
##########
# Set up burn-in period and iteration numbers M
sa2burnin=10000
sa2M=100000
system.time({
sa2para=MCMCepi(udata,106682518,sa2burnin,sa2M)
})

# Acceptance rate
acptrates=matrix(0,1,4)
for (i in 1:4){
acptrates[,i]=sum(sa2para[,(5+i)])/(sa2burnin+sa2M)
}
acptrates

# Eliminate burn-in period
sa2parab=sa2para[(sa2burnin+1):(sa2burnin+sa2M),]

# MCMC statistics
MCMCstat=matrix(0,4,5)
for (i in 1:4){
  mean=mean(sa2parab[,i+1])
  median=median(sa2parab[,i+1])
  a=table(round(sa2parab[,i+1],4))
  mode=as.numeric(names(a)[a==max(a)])
  sd=sqrt(var(sa2parab[,i+1]))
  L=quantile(sa2parab[,i+1],0.025)
  U=quantile(sa2parab[,i+1],0.975)
  MCMCstat[i,]=c(mean,median,sd,L,U)
}
MCMCstat[1,]=round(MCMCstat[1,],3)
MCMCstat[2,]=round(MCMCstat[2,],2)
MCMCstat[3,]=round(MCMCstat[3,],0)
MCMCstat[4,]=round(MCMCstat[4,],4)
MCMCstat

# The MCMC diagnostic plots
MCMCplot(sa2parab)

# Save MCMC data
save(sa2para,
file = "C:\\sa2.RData")
```

SIMINCIDENCE

```
################################################################################
# Program name: SIMINCIDENCE
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: To simulate incidence curve
# Notes:
################################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description------------|
#
################################################################################

######################
# Load MCMC samples #
######################
load("C:\\obs.RData")

# Eliminate burn-in period
burnin=10000
M=100000
obsparab=obspara[(burnin+1):(burnin+M),]

# Epidemic settings
N=106682518
I0=7

###################
# Incidence curve #
###################
# Create store file for incidence
datalength=50
buffer=30
# number of simulations
itern=1000

xstore=matrix(0,buffer+datalength+1,itern)

# Simulate incidence
for (j in 1:itern){

  # Random select R0 and tau
  randno1=trunc(runif(1,1,(M+1)))
  R0adrand=obsparab[randno1,2]
  TI=3
  TIprand=1-exp(-1*(1/TI))
  randno2=trunc(runif(1,1,(M+1)))
  taurand=obsparab[randno2,4]

  # stochastic SIR model
  sirsim=matrix(0,buffer+datalength+1,6)
  sirsim[1,]=c(0,N,I0,0,0,0)

  for (i in 1:(buffer+datalength)){
    if (i<taurand){
      sirsim[i+1,]=c(i,N,I0,0,0,0)
    }
    else {
      pi=1-exp(-1*(R0adrand*sirsim[i,3])/N)
      #x=rbinom(1,sirsim[i,2],pi)
      x=round(pi*sirsim[i,2], 0)
      S=max(0,sirsim[i,2]-x)
      #y=rbinom(1,sirsim[i,3],TIprand)
      y=round(TIprand*sirsim[i,3], 0)
      I=max(0,sirsim[i,3]+x-y)
      R=max(0,sirsim[i,4]+y)
      sirsim[i+1,]=c(i,S,I,R,x,y)
    }
  }

  # store the incidence
  xstore[,j]=sirsim[,5]
```

```
}

# incidence statistics
xdraw = NULL
for (k in 1:(buffer+datalength+1)){
  x_ll=quantile(xstore[k,],0.025)
  x_ml=median(xstore[k,])
  x_ul=U=quantile(xstore[k,],0.975)
  xdraw=rbind(xdraw,c(x_ll,x_ml,x_ul))
}

xdraw

# File save
write.csv(xdraw,
file="C:\\xdraw.csv",row.names = FALSE)
```

SIMFETPDF

```
################################################################################
# Program name: SIMFETPDF
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: To simulate the probability distributions of FET and F10ET
# Notes:
################################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description-------------|
#
################################################################################

######################
# Load MCMC samples #
######################
load("C:\\obs.RData")

# Eliminate burn-in period
burnin=10000
M=100000
obsparab=obspara[(burnin+1):(burnin+M),]

# Epidemic settings
N=106682518
I0=7

###################################
# Distribution of exported cases #
###################################
simdaydist=function(itern,setm,firstcut){

  datalength=240
  arrdayk=NULL

  # SIR model
  for (iter in 1:itern){

    # Random select R0
    randno1=trunc(runif(1,1,(M+1)))
    R0adrand=obsparab[randno1,2]
    TI=3
    TIprand=1-exp(-1*(1/TI))

    sir =matrix(0,datalength+1,7)
    sir[1,]=c(0,N,I0,0,0,0,0)

    # SIR matrix
    for (i in 1:datalength){
      pi=1-exp(-1*(R0adrand*sir[i,3])/N)
```

```
      x=round(pi*sir[i,2], 0)
      S=max(0,sir[i,2]-x)
      y=round(sir[i,3]*TIprand,0)
      I=max(0,sir[i,3]+x-y)
      R=max(0,sir[i,4]+y)
      newcase=rbinom(1,setm,pi)
      cumcase=max(0,sir[i,7]+newcase)
      sir[i+1,]=c(i,S,I,R,x,y,cumcase)
      if (cumcase>firstcut) {
        arrday=sir[i+1,1]
        break
      }
      arrday=0
    }
  # Store the arrival day
  arrdayk = rbind(arrdayk,arrday)
  }

  # export the simulated exported day
  return (arrdayk)
}

# distribution of m=300, first case, 10000 iterations #
system.time({
first1m300=simdaydist(50000,300,0)
first1m300r90=simdaydist(50000,30,0)
first1m300r99=simdaydist(50000,3,0)
})
FET300=cbind(first1m300,first1m300r90,first1m300r99)

# statistics
FETstat=matrix(0,3,5)
for (k in 1:3){
mean=mean(FET300[,k])
median=median(FET300[,k])
sd=sqrt(var(FET300[,k]))
L=quantile(FET300[,k],0.025)
U=quantile(FET300[,k],0.975)
FETstat[k,]=c(mean,median,sd,L,U)
}
FETstat

# pdf plot
hist(FET300[,1],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

hist(FET300[,2],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

hist(FET300[,3],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

plot(density(FET300[,1],bw=0.8),xlim=c(0,180),ylim=c(0,0.1),main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()
lines(density(FET300[,2],bw=0.8),lty=2)
lines(density(FET300[,3],bw=0.6),lty=3)

# distribution of m=1500, first case, 1000 iterations #
system.time({
first1m1500=simdaydist(50000,1500,0)
first1m1500r90=simdaydist(50000,150,0)
first1m1500r99=simdaydist(50000,15,0)
})
FET1500=cbind(first1m1500,first1m1500r90,first1m1500r99)

# statistics
FETstat=matrix(0,3,5)
```

```
for (k in 1:3){
mean=mean(FET1500[,k])
median=median(FET1500[,k])
sd=sqrt(var(FET1500[,k]))
L=quantile(FET1500[,k],0.025)
U=quantile(FET1500[,k],0.975)
FETstat[k,]=c(mean,median,sd,L,U)
}
FETstat

# pdf plot
hist(FET1500[,1],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

hist(FET1500[,2],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

hist(FET1500[,3],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

plot(density(FET1500[,1],bw=0.8),xlim=c(0,180),ylim=c(0,0.1),main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()
lines(density(FET1500[,2],bw=0.8),lty=2)
lines(density(FET1500[,3],bw=0.8),lty=3)

# distribution of m=300, first 10 cases, 50000 iterations #
system.time({
first10m300=simdaydist(50000,300,9)
first10m300r90=simdaydist(50000,30,9)
first10m300r99=simdaydist(50000,3,9)
})
F10ET300=cbind(first10m300,first10m300r90,first10m300r99)

# statistics
F10ETstat=matrix(0,3,5)
for (k in 1:3){
mean=mean(F10ET300[,k])
median=median(F10ET300[,k])
sd=sqrt(var(F10ET300[,k]))
L=quantile(F10ET300[,k],0.025)
U=quantile(F10ET300[,k],0.975)
F10ETstat[k,]=c(mean,median,sd,L,U)
}
F10ETstat

# pdf plot
hist(F10ET300[,1],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

hist(F10ET300[,2],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

hist(F10ET300[,3],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

plot(density(F10ET300[,1],bw=0.5),xlim=c(0,180),ylim=c(0,1),
     main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()
lines(density(F10ET300[,2],bw=0.5),lty=2)
```

```
lines(density(F10ET300[,3],bw=0.4),lty=3)

# distribution of m=1500, first 10 cases, 50000 iterations #
system.time({
first10m1500=simdaydist(50000,1500,9)
first10m1500r90=simdaydist(50000,150,9)
first10m1500r99=simdaydist(50000,15,9)
})
F10ET1500=cbind(first10m1500,first10m1500r90,first10m1500r99)

# statistics
F10ETstat=matrix(0,3,5)
for (k in 1:3){
mean=mean(F10ET1500[,k])
median=median(F10ET1500[,k])
sd=sqrt(var(F10ET1500[,k]))
L=quantile(F10ET1500[,k],0.025)
U=quantile(F10ET1500[,k],0.975)
F10ETstat[k,]=c(mean,median,sd,L,U)
}
F10ETstat

# pdf plot
hist(F10ET1500[,1],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

hist(F10ET1500[,2],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

hist(F10ET1500[,3],breaks =300, main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()

plot(density(F10ET1500[,1],bw=0.5),xlim=c(0,180),ylim=c(0,0.4),
     main="",xlab="",ylab="",axes=F)
axis(1)
axis(2,las=1)
box()
lines(density(F10ET1500[,2],bw=0.5),lty=2)
lines(density(F10ET1500[,3],bw=0.45),lty=3)
```

SIMEPFET

```
##############################################################################
# Program name: SIMEPFET
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: To simulate expected daily probability distributions of FET
# Notes:
##############################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description-------------|
#
##############################################################################

# Epidemic settings
N=106682518
I0=7

########################################
# Expected PDF of export > 0 case by day #
########################################
Epexp=function(setm){
```

```
  datalength=240
  arrdayk=NULL

  # expected R0
  TI=3
  TIprand=1-exp(-1*(1/TI))
  R0ad=0.411

  # SIR model
  sir =matrix(0,datalength+1,8)
  pbyday =matrix(0,datalength+1,3)
  sir[1,]=c(0,N,I0,0,0,0,0,0)

  # SIR matrix
  for (i in 1:datalength){
    pi=1-exp(-1*(R0ad*sir[i,3])/N)
    x=round(pi*sir[i,2], 0)
    S=max(0,sir[i,2]-x)
    y=round(sir[i,3]*TIprand,0)
    I=max(0,sir[i,3]+x-y)
    R=max(0,sir[i,4]+y)
    p0=dbinom(0,setm,pi)
    p1=1-dbinom(0,setm,pi)
    sir[i+1,]=c(i,S,I,R,x,y,p0,p1)
    pbyday[i+1,]=c(i,p0,p1)
  }

  # export the probability by day
  return (pbyday)
}

# pdf of export by day for m=300 with restrictions
Ep300r0=Epexp(300)
Ep300r90=Epexp(30)
Ep300r99=Epexp(3)

Ep300=cbind(Ep300r0,Ep300r90,Ep300r99)

# File save
write.csv(Ep300,
file="C:\\Ep300.csv",row.names = FALSE)

# pdf of export by day for m=1500 with restrictions
Ep1500r0=Epexp(1500)
Ep1500r90=Epexp(150)
Ep1500r99=Epexp(15)

Ep1500=cbind(Ep1500r0,Ep1500r90,Ep1500r99)

# File save
write.csv(Ep1500,
file="C:\\Ep1500.csv",row.names = FALSE)
```

SIMFETVSM

```
################################################################################
# Program name: SIMFETVSM
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: To simulate FET against different daily rates of travel
# Notes:
################################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description-------------|
#
################################################################################

# Epidemic settings
N=106682518
```

```
I0=7

##############################
# First export day against m #
##############################
dayvsm=function(itern,lm,um,firstcut,R0){

  datalength=200

  # Select R0
  I0=7
  TI=3
  TIprand=1-exp(-1*(1/TI))
  R0adrand=R0/TI

  mdayiter=NULL
  # SIR model
  for (iter in 1:itern){
    mdaydrawk=NULL
    for (mlevel in seq(lm,um,by=10)){
      sir =matrix(0,datalength+1,7)
      sir[1,]=c(0,N,I0,0,0,0,0)

      # SIR matrix
      for (i in 1:datalength){
        pi=1-exp(-1*(R0adrand*sir[i,3])/N)
        x=round(pi*sir[i,2], 0)
        S=max(0,sir[i,2]-x)
        y=round(sir[i,3]*TIprand,0)
        I=max(0,sir[i,3]+x-y)
        R=max(0,sir[i,4]+y)
        newcase=rbinom(1,mlevel,pi)
        cumcase=max(0,sir[i,7]+newcase)
        sir[i+1,]=c(i,S,I,R,x,y,cumcase)
        if (cumcase>firstcut) {
          mdaydraw=sir[i+1,1]
          break
        }
        mdaydraw=0
      }

    # Store the arrival day per simulations
    mdaydrawk = rbind(mdaydrawk,mdaydraw)
    }

  mdayiter=cbind(mdayiter,mdaydrawk)
  }

  # export the simulated exported day
  return (mdayiter)
}

# 1000 iterations m against first export day in baseline R0
system.time({
dayvsmdraw=dayvsm(1000,10,5000,0,1.23)
})

# export days statistics
dmdraw = NULL
drawlength=length(dayvsmdraw[,1])
for (k in 1:(drawlength)){
  dm_ll=quantile(dayvsmdraw[k,],0.025)
  dm_ml=median(dayvsmdraw[k,])
  dm_ul=U=quantile(dayvsmdraw[k,],0.975)
  dmdraw=rbind(dmdraw,c(dm_ll,dm_ml,dm_ul))
}
dmdata=cbind(seq(10,5000,10),dmdraw)

# File save
write.csv(dmdata,
file="C:\\dmdata.csv",row.names = FALSE)
```

SIMFETVSR0

```
################################################################################
# Program name: SIMFETVSR0
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: To simulate FET against different R0
# Notes:
################################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description-------------|
#
################################################################################

# Epidemic settings
N=106682518
I0=7

##############################
# First export day against R #
##############################
R0vsFET=function(itern,R0,firstcut){

  datalength=240

  # Random select R0
  I0=7
  TI=3
  TIprand=1-exp(-1*(1/TI))
  R0adrand=R0/TI

  mdayiter=NULL
  # SIR model
  for (iter in 1:itern){
    mdaydrawk=NULL
    for (mlevel in c(3,30,300,1500,3000)){
      sir =matrix(0,datalength+1,7)
      sir[1,]=c(0,N,I0,0,0,0,0)

      # SIR matrix
      for (i in 1:datalength){
        pi=1-exp(-1*(R0adrand*sir[i,3])/N)
        x=round(pi*sir[i,2], 0)
        S=max(0,sir[i,2]-x)
        y=round(sir[i,3]*TIprand,0)
        I=max(0,sir[i,3]+x-y)
        R=max(0,sir[i,4]+y)
        newcase=rbinom(1,mlevel,pi)
        cumcase=max(0,sir[i,7]+newcase)
        sir[i+1,]=c(i,S,I,R,x,y,cumcase)
        if (cumcase>firstcut) {
          mdaydraw=sir[i+1,1]
          break
        }
        mdaydraw=0
      }

    # Store the arrival day per simulations
    mdaydrawk = rbind(mdaydrawk,mdaydraw)
    }

  mdayiter=cbind(mdayiter,mdaydrawk)
  }

  # export the simulated exported day
  return (mdayiter)
}

# 1000 iterations m against first export day in baseline R0
system.time({
R0vsFETdraw=R0vsFET(10000,1.23,0)
})
```

```
# export days statistics
rfdraw = NULL
drawlength=length(R0vsFETdraw[,1])
for (k in 1:(drawlength)){
  rf_ll=quantile(R0vsFETdraw[k,],0.025)
  rf_ml=median(R0vsFETdraw[k,])
  rf_ul=U=quantile(R0vsFETdraw[k,],0.975)
  rfdraw=rbind(rfdraw,c(rf_ll,rf_ml,rf_ul))
}
cbind(c(3,30,300,1500,3000),rfdraw)
```

SIMCOUNTRY

```
##############################################################################
# Program name: SIMCOUNTRY
# Version: 1.0
# Author: Marc, Biostatistician
# Study name: H1N1 simulaton study
# Created date: 01NOV2010
# Purpose: To simulate number of countries received infected cases by day
# Notes:
##############################################################################
# Amendment history:
# |--Amended date--|--Amended by--|----------Description-------------|
#
##############################################################################

######################
# Load MCMC samples #
######################
load("C:\\obs.RData")

# Eliminate burn-in period
burnin=10000
M=100000
obsparab=obspara[(burnin+1):(burnin+M),]

# Epidemic settings
N=106682518
I0=7

#########################################################
# Number of countries received infected cases by day #
#########################################################
# Travel rates by 22 countries

msim=c(24609,5240,38749,101313,18535,24535,16950,42802,
  15090,61960,35772,39460,2340,12060,4675,27640,3101,
  48717,15478,65724,20513,9150)
msimdaily=trunc(msim/61)

countrycount=function(itern,pm,firstcut,pR,pRtime){

  datalength=240
  kcountdraw=NULL
  mlength=length(msimdaily)
  msimres=round((1-pm)*msimdaily, 0)

  # SIR model
  for (iter in 1:itern){

    # Random select R0
    randno1=trunc(runif(1,1,(M+1)))
    R0adrand=obsparab[randno1,2]
    I0=7
    TI=3
    TIprand=1-exp(-1*(1/TI))

    marrive =matrix(0,mlength,1)
    sir =matrix(0,datalength+1,7)
    kcount=matrix(0,datalength+1,1)
```

```
    sir[1,]=c(0,N,I0,0,0,0,0)
    kcount[1,]=0

    # SIR matrix
    for (i in 1:datalength){
      if (i>pRtime){
        R0adinterven=pR*R0adrand
      }
      else{
        R0adinterven=R0adrand
      }
      pi=1-exp(-1*(R0adinterven*sir[i,3])/N)
      x=round(pi*sir[i,2], 0)
      S=max(0,sir[i,2]-x)
      y=round(sir[i,3]*TIprand,0)
      I=max(0,sir[i,3]+x-y)
      R=max(0,sir[i,4]+y)

      # simulate the arrive time
      for (j in 1:mlength){
        newcase=rbinom(1,msimres[j],pi)
        if (newcase>firstcut & marrive[j,1]==0){
          marrive[j,1]=1
        }
      }

      cumcase=sum(marrive)
      sir[i+1,]=c(i,S,I,R,x,y,cumcase)
      kcount[i+1,]=cumcase
    }

    # Store the country count per simulations
    kcountdraw=cbind(kcountdraw,kcount)
  }

  # export the simulated exported day
  return (kcountdraw)
}

# 10000 iterations m against first export day in baseline R0
system.time({
kcountr0=countrycount(10000,0,0,1,60)
kcountr90=countrycount(10000,0.9,0,1,60)
kcountr99=countrycount(10000,0.99,0,1,60)
})

# export days statistics
kdrawr0 = NULL
drawlength=length(kcountr0[,1])
for (k in 1:(drawlength)){
  k_ll=quantile(kcountr0[k,],0.025)
  k_ml=median(kcountr0[k,])
  k_ul=U=quantile(kcountr0[k,],0.975)
  kdrawr0=rbind(kdrawr0,c(k_ll,k_ml,k_ul))
}

# export days statistics
kdrawr90 = NULL
drawlength=length(kcountr90[,1])
for (k in 1:(drawlength)){
  k_ll=quantile(kcountr90[k,],0.025)
  k_ml=median(kcountr90[k,])
  k_ul=U=quantile(kcountr90[k,],0.975)
  kdrawr90=rbind(kdrawr90,c(k_ll,k_ml,k_ul))
}

# export days statistics
kdrawr99 = NULL
drawlength=length(kcountr99[,1])
for (k in 1:(drawlength)){
  k_ll=quantile(kcountr99[k,],0.025)
  k_ml=median(kcountr99[k,])
  k_ul=U=quantile(kcountr99[k,],0.975)
  kdrawr99=rbind(kdrawr99,c(k_ll,k_ml,k_ul))
}
```

```
# File save
kdraw=cbind(kdrawr0,kdrawr90,kdrawr99)

write.csv(kdraw,
file="C:\\kdraw.csv",row.names = FALSE)

# 10000 iterations m against first export day in 0.8*R0
system.time({
kcountr0R8=countrycount(10000,0,0,0.8,60)
kcountr90R8=countrycount(10000,0.9,0,0.8,60)
kcountr99R8=countrycount(10000,0.99,0,0.8,60)
})

# export days statistics
kdrawr0R8 = NULL
drawlength=length(kcountr0R8[,1])
for (k in 1:(drawlength)){
  k_ll=quantile(kcountr0R8[k,],0.025)
  k_ml=median(kcountr0R8[k,])
  k_ul=U=quantile(kcountr0R8[k,],0.975)
  kdrawr0R8=rbind(kdrawr0R8,c(k_ll,k_ml,k_ul))
}

# export days statistics
kdrawr90R8 = NULL
drawlength=length(kcountr90R8[,1])
for (k in 1:(drawlength)){
  k_ll=quantile(kcountr90R8[k,],0.025)
  k_ml=median(kcountr90R8[k,])
  k_ul=U=quantile(kcountr90R8[k,],0.975)
  kdrawr90R8=rbind(kdrawr90R8,c(k_ll,k_ml,k_ul))
}

# export days statistics
kdrawr99R8 = NULL
drawlength=length(kcountr99R8[,1])
for (k in 1:(drawlength)){
  k_ll=quantile(kcountr99R8[k,],0.025)
  k_ml=median(kcountr99R8[k,])
  k_ul=U=quantile(kcountr99R8[k,],0.975)
  kdrawr99R8=rbind(kdrawr99R8,c(k_ll,k_ml,k_ul))
}

# File save
kdrawR8=cbind(kdrawr0R8,kdrawr90R8,kdrawr99R8)

write.csv(kdrawR8,
file="C:\\kdrawR8.csv",row.names = FALSE)

# 10000 iterations m against first export day in 0.6*R0
system.time({
kcountr0R6=countrycount(10000,0,0,0.6,60)
kcountr90R6=countrycount(10000,0.9,0,0.6,60)
kcountr99R6=countrycount(10000,0.99,0,0.6,60)
})

# export days statistics
kdrawr0R6 = NULL
drawlength=length(kcountr0R6[,1])
for (k in 1:(drawlength)){
  k_ll=quantile(kcountr0R6[k,],0.025)
  k_ml=median(kcountr0R6[k,])
  k_ul=U=quantile(kcountr0R6[k,],0.975)
  kdrawr0R6=rbind(kdrawr0R6,c(k_ll,k_ml,k_ul))
}

# export days statistics
kdrawr90R6 = NULL
drawlength=length(kcountr90R6[,1])
for (k in 1:(drawlength)){
  k_ll=quantile(kcountr90R6[k,],0.025)
  k_ml=median(kcountr90R6[k,])
  k_ul=U=quantile(kcountr90R6[k,],0.975)
  kdrawr90R6=rbind(kdrawr90R6,c(k_ll,k_ml,k_ul))
```

```
}

# export days statistics
kdrawr99R6 = NULL
drawlength=length(kcountr99R6[,1])
for (k in 1:(drawlength)){
  k_ll=quantile(kcountr99R6[k,],0.025)
  k_ml=median(kcountr99R6[k,])
  k_ul=U=quantile(kcountr99R6[k,],0.975)
  kdrawr99R6=rbind(kdrawr99R6,c(k_ll,k_ml,k_ul))
}

# File save
kdrawR6=cbind(kdrawr0R6,kdrawr90R6,kdrawr99R6)

write.csv(kdrawR6,
file="C:\\kdrawR6.csv",row.names = FALSE)
```

Appendix C

Cost-effectiveness study of travel restrictions

Methods

Modeling the Cost-effectiveness Analysis

In the cost-effectiveness analysis (CEA), we compared the costs and the quality-adjusted life year (QALY) with strategies of no available intervention and imposing 99% travel restrictions before the time of the availability of antiviral and hospitalization in a year. The total cost (C_{Total}) of each strategy was calculated by

$$C_{Total} = C_I + C_H + C_P \tag{C.1}$$

where the cost of travel restrictions, health care, and production were represented by C_I, C_H, and C_P respectively. The costs and benefits were discounted at rate of 3% per annum. All costs are in Hong Kong dollars (HKD$) (2009).

The incremental cost-effectiveness ratio (ICER) was the main endpoint of the study. The numerator of ICER was the incremental cost which was the difference of the costs, and the denominators was the difference of QALYs from imposing travel restrictions. An intervention is 'cost-saving' if it can reduce costs and also can raise the QALYs. An intervention is 'cost-effective' if its ICER is less than a determined threshold with increases in QALYs [104, 102]. Median ICERs were obtained from 100 realizations that were simulated from the epidemic model. The CEA was analyzed using Microsoft Excel, version 2003.

Scenario Design

We analyzed the cost-effectiveness by selecting the time of the availability of antiviral and hospitalization to 3.5 months, 5 months, and 6.5 months after the first global onset case. We assumed the virus influenza A-like with the same lengths of the latent period and infectious period. The R_0 was considered in range of 1.5 to 8 for all countries in the model. The 99% travel restriction was imposed from the time of first global case onset to the time of antiviral and hospitalization being available. The case-fatality rate was considered in range of 0.5% to 50%.

Oseltamivir and Zanamivir were selected as the standard antiviral treatment during the pandemic influenza A (H1N1) [33]. Each patient would receive one course of antiviral treatment. Severe patients who were hospitalized would stay in hospital and were absent from work for 7.7 days on average [67]. We also assumed 30%, 50% and 20% of the untreated infections as in asymptomatic, mild, and moderate status respectively. The mild and moderate cases would seek medical care as well as seeing doctors and taking prescribed medication.

Moderate cases and the patients who received antiviral would be absent from work for 1.5 days.

Costs

Costs of Travel Restrictions (C_I)

The direct economic impact from the travel restrictions is the tourism industry in Hong Kong. According to the statistics of Hong Kong government [105], the tourism industry contributed 2.6% (HKD$40,264 million) to Hong Kong's Gross Domestic Product (GDP) in 2009. The tourism-related activities such as accommodation services, retail trade, transport services, and food and beverage services were included into the GDP's calculation. We assumed the restricted length of time and scale of travel was directly proportional to the local GDP i.e. the 99% travel restrictions brought a monthly HKD$3,324.75 million loss in Hong Kong's GDP in 2009.

Health Care Costs (C_H)

Table C.1 presents the cost variables utilized in the CEA. Oseltamivir and Zanamivir were $222.7/case for a single course of treatment which included a brief medical consultation. Individual who sought medical care was required to spend $335 in total for the medical consultation and the prescribed medication. Severe subject who was hospitalized spent $16,170 for staying in a Hospital Authority (HA) hospital. We assumed zero travel costs for seeking treatments. The size of the antiviral stockpile was considered as sufficient in one year period.

Table C.1: CEA parameters and the corresponding values

Parameter	Value	Remark	Reference
Hospitalization	$16,170/case	$2,100/day and stay in hospital for 7.7 days	[42]
Oseltamivir	$222.7/case	EUR21.624 for single treatment course; one EUR converted to HKD$10.3	[33]
Zanamivir	$222.7/case	EUR21.624 for single treatment course; one EUR converted to HKD$10.3	[33]
Medical care	$335/case	$127 for medical consultation and $100 for prescribed medication with 5% inflation each year	[42]
Lost productivity	$2,825.9/hospitalized case; $550.5 per moderate case or case who received antiviral	Average monthly wage was $11,000 which divided by 30 days; hospitalized cases absented from work for 7.7 days; moderate cases and patients who received antiviral absented from work for 1.5 days	[101]
QALY lost	0.008 for an untreated infection; 0.004 for an individual who received antiviral; 0.017 for a hospitalized individual		[11] and assumption

Production Costs (C_P)

The monthly salary was assumed to be \$11,000 on average for a Hong Kong resident [101]. Severe subjects would lose their productivity when staying in hospitals. It costed \$2,825.9 for a hospitalized case. Moderate case or case who received antiviral lost \$550.5 for a 1.5 days absented period. Table C.1 summarizes the production costs in the CEA.

Quality-adjusted Life Year Lost

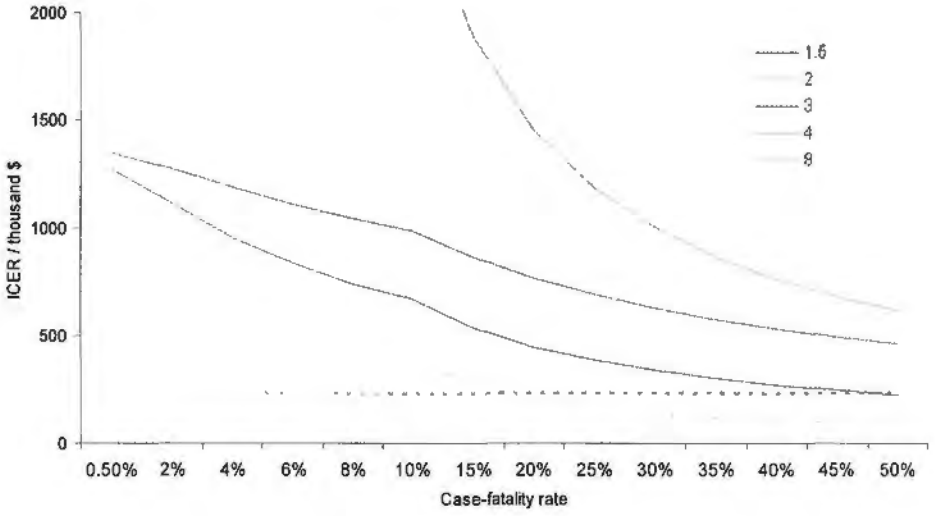
A QALY is a measure of an individual's physical health as well as a measure of disease burden. In our study, previously published QALYs losts of pandemic influenza A (H1N1) were used [11]. Individual who received antiviral was assumed in having half of QALYs lost as an untreated subject (Table C.1).

Results

In our model, the cost-effectiveness was dependent on the interplay between the case-fatality rate, the transmission intensity, and the implementation time of effective interventions. According to the modeling results, the total cost would be \$7,265 million if only antiviral and hospitalization imposed on 3.5 months after the first global onset case, and the case-fatality rate was assumed in 0.5% that similar to that of pandemic (H1N1) 2009 [1]. The cost of a 99% travel restriction of all means of transports would be \$11,636 million if it was imposed for 3.5 months. The program would cost \$363 million for each QALY gained in this situation. In views of CEA, it was not recommended including travel restrictions in the plan of pandemic (H1N1) 2009 (Figure C.1). By employing per capita GDP \$231,600 of Hong Kong in 2009 as the cost-effective threshold

[104, 59], the travel restriction was cost-effective only if the R_0 increased to 8 and the case-fatality rate > 15%.

Figure C.1: Incremental cost-effectiveness ratio (ICER) for 99% travel restrictions of all means of transports with different reproduction numbers (R_0) and case-fatality rates when the antiviral and hospitalization were available on 3.5 months after the first global onset case.

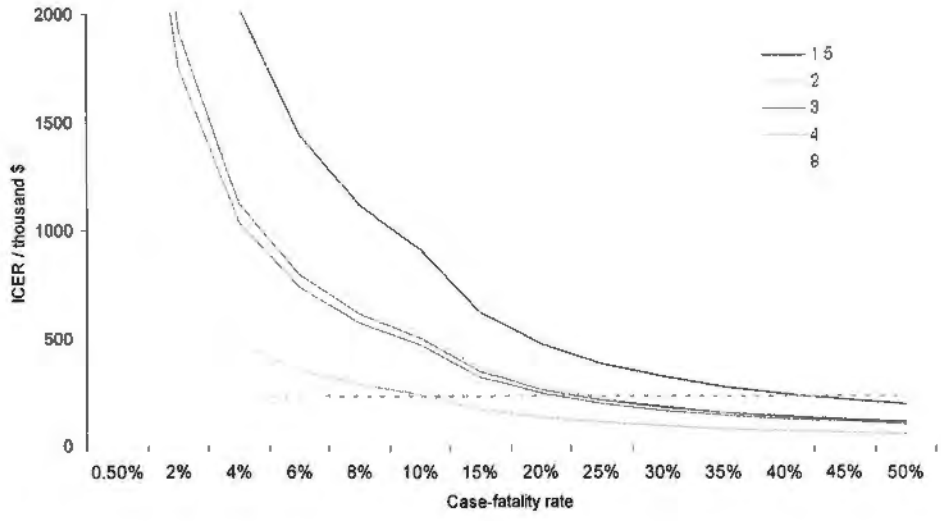


The solid lines of R_0s are illustrated in different colors showed in top-right corner. The red dotted line represents the cost-effective threshold i.e. $231,600. The solid lines of $R_0 = 1.5$ cannot be showed in the image due to its ICER > 2 million within the range of case-fatality rates.

Although a 99% travel restriction of all means of transports would cost $16,624 million for five months, it was cost-effective in some situations once the antiviral and hospitalization were available at the end of the fifth month after the first global onset case. Imposing travel restriction allowed the gain of about 154 thousand QALYs and the average ICER was about $166,400 in one year period when the $R_0 = 2$ with 15% case-fatality rate (Figure C.2). The travel restriction was also cost-effective once the case-fatality rate attained 25%

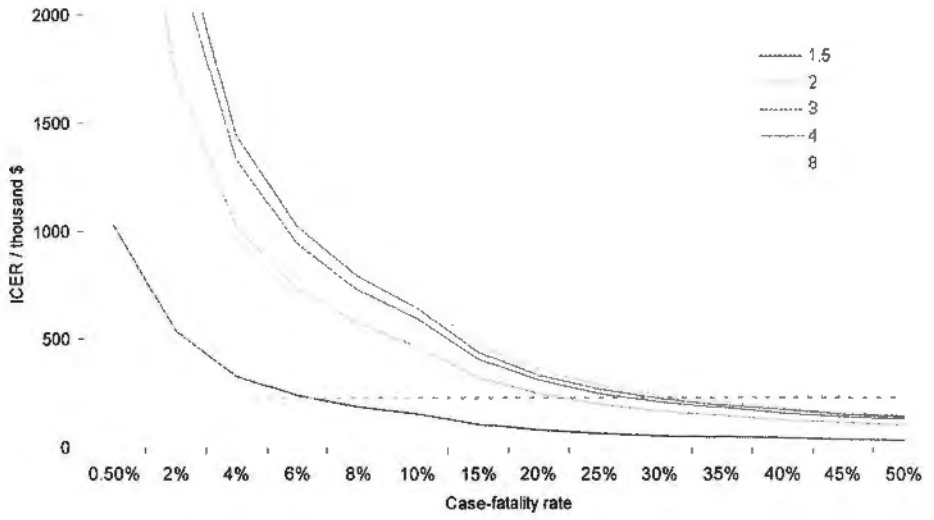
with high disease transmission intensity. If the antiviral and hospitalization were only available at 6.5 month after the first global onset case, the travel restrictions delayed the epidemic and was cost-effective when the $R_0 = 1.5$ with about 6% case-fatality rate (Figure C.3); more than 100 thousand QALYs were gained by paying $21,611 million for travel restrictions in total.

Figure C.2: Incremental cost-effectiveness ratio (ICER) for 99% travel restrictions of all means of transports with different reproduction numbers (R_0) and case-fatality rates when the antiviral and hospitalization were available on 5 months after the first global onset case.



The solid lines of R_0s are illustrated in different colors showed in top-right corner. The red dotted line represents the cost-effective threshold i.e. $231,600.

Figure C.3: Incremental cost-effectiveness ratio (ICER) for 99% travel restrictions of all means of transports with different reproduction numbers (R_0) and case-fatality rates when the antiviral and hospitalization were available on 6.5 months after the first global onset case.



The solid lines of R_0s are illustrated in different colors showed in top-right corner. The red dotted line represents the cost-effective threshold i.e. $231,600.