Modeling and optimization of the allocation of a new dengue vaccine

---

A Master's Project
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Masters of Science
Mathematics

---

by
Abhishek Pandey
May 2011

Accepted by:
Dr. Jan Medlock, Committee Chair
Dr. Elena Dimitrova
Dr. Jo Hoffacker
Dr. Pete Keisler

Abstract

Dengue is one of the most rapidly spreading mosquito-borne viral diseases in the world and inflicts significant health, economic and social burden on populations. Various mathematical models have appeared in literature to investigate dengue epidemiology. Due to significant progress made in dengue vaccine development lately, a licensed vaccine is expected to be available in less than 10 years. It is imperative to choose target age groups for the vaccine that most efficiently reduce the disease burden. A generic dengue mathematical model based on previous dengue models is formulated and a Bayesian approach that uses Monte Carlo Markov Chain (MCMC) simulation to estimate the unknown parameters of the model is outlined. We propose to use Akaike Information Criterion and Bayesian Information Criterion on the results from Bayesian MCMC on our models with clinical data from Thailand to select most parsimonious model for dengue transmission. This model selection could be significant contribution to understanding of dengue transmission. The dengue model will then be extended to include age structure in the population and an optimization framework is presented that will be used to examine the optimal vaccine allocation.

Acknowledgments

The project was made possible due to help and support of the faculty members, friends and colleagues.

I must begin by thanking my advisor Dr. Jan Medlock, who graciously took me as his pupil. Dr. Medlock's encouragement, words of wisdom, and insightful suggestions helped me in carrying research and completing this project.

I would like to thank Dr. Elena Dimitrova, Dr. Jo Hoffacker and Dr. Pete Keisler for being in my committee and I am also thankful to my friends: Vivek, Navamita and Min, who never hesitated to help me out, when this manuscript was in progress.

Table of Contents

List of Figures

Chapter 1

Introduction

Humankind has always been afflicted by infectious diseases, but epidemics were comparatively rare before the advent of human civilization. Once people stopped leading a nomadic life and began living in villages, then in towns and cities, pathogens that cause infectious diseases started spreading easily. People get exposed to these pathogens either by direct contact (through air, polluted water and food) or by indirect contact (through bloodsucking insect carriers of disease such as mosquitoes, fleas and lice).

Dengue is one of the most rapidly spreading mosquito-borne viral diseases in the world and inflicts significant health, economic and social burden on populations. Worldwide, an estimated 2.5 billion people live in areas where dengue is an epidemic, out of which approximately 975 million live in urban areas in tropical and sub-tropical countries in Southeast Asia, the Pacific and the Americas (Figure 1.1). Dengue has been recognized in over 100 countries and an estimated 50–100 million dengue infections occur annually (Guzmán and Kouri, 2002). Moreover, the global estimation of number of disability-adjusted life years (DALYs) lost to dengue in 2001 was 528000.

Dengue virus (DEN) is a small, spherical and single-stranded RNA virus in the genus *Flavivirus* in the family Flaviviridae, which also includes yellow fever virus and West Nile Virus. Dengue viruses are divided into four distinct classes, known as serotypes, referred to as DEN-1, DEN-2, DEN-3 and DEN-4. Infection with one serotype provides immunity to other viruses in that same serotype, but no long-term immunity to the other serotypes. Within each serotype, distinct genotypes have been identified which indicates the extensive genetic variability within the dengue serotypes. An individual infected with one of the four serotypes such that the individual had no prior experience

Figure 1.1: Areas at risk for dengue transmission, 2008 (Organization et al., 2009).

of infection with any of the serotype is said to be infected with Primary infection. Similarly, an infected individual is said to be infected with Secondary infection if the individual have had prior infection with any one of the dengue serotypes. Persons living in areas where dengue is an endemic can be infected with three and probably four dengue serotypes during their lifetime (Gubler, 1998a).

The various serotypes of the dengue virus are transmitted to humans through the bites of *Aedes* mosquitoes. *Aedes aegypti* mosquitoes are the predominant vectors for dengue infection, however *Aedes albopictus* and other *Aedes* species are also able to transmit dengue with varying degree of efficiency. The mosquitoes acquire the virus when they bite an infected human. The mosquitoes are capable of transmitting dengue if they bite another human immediately or after incubating the infection for eight to twelve days, which is known as the extrinsic incubation period. The mosquitoes remain infected for rest of their lives. Vertical transmission of virus from mother to offspring is thought to be rare in both mosquitoes and humans (Siler et al., 1926; Sabin, 1952; Gubler, 1998b; Guzmán and Kouri, 2002; Luz et al., 2003; Kyle and Harris, 2008) .

Once the dengue virus is inoculated into a human host, it incubates for a period of 4–10 days, the intrinsic incubation period. Following incubation, an infected person enters the acute phase of infection for about 5 days. The host recovers from the infection usually within 7–10 days. Infection with one type of dengue serotype provides lifelong protective immunity to the infecting serotype, and possibly partial and short-lived protection from infection with other dengue serotypes (Sabin, 1952). The symptoms of disease vary greatly from mild fever, high fever with severe headache and

joint pain and internal hemorrhaging, to circulatory failure and death. The cases are classified, in order of increasing severity as dengue fever (DF), dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS) (Guzmán and Kouri, 2002; Halstead, 2007). The severity of disease in an individual is determined by several factors such as age, ethnicity and, possibly the presence of some chronic diseases and previous dengue infection (Guzmán and Kouri, 2002). Young children can be especially susceptible to dengue (Guzmán et al., 2002). Most patients who develop severe forms of dengue (DHF or DSS) have had prior infections with one or more dengue serotypes (Halstead, 2007). One explanation for this phenomenon is antibody-dependent enhancement (ADE), where the presence of antibodies to one dengue serotype enhance the replication of viruses from other serotypes, perhaps leading to increased susceptibility to infection or transmission once infected, in addition to increasing the risk of severe disease (Halstead, 2007). Certain dengue genotypes, particularly those of DEN-2, are thought to be more virulent than others, since more number of cases of DHF have been associated with DEN-2 than with the other serotypes (Rico-Hesse, 2003).

With the expanding geographic distribution and increased disease incidence in past several decades, the prevention and control of dengue infection has become very important. Unfortunately, tools available to prevent dengue infection are very limited. For many years, some viral diseases have been controlled using vaccines, however a dengue vaccine is yet to be made available. Therefore, in order to reduce or prevent dengue virus transmission, there is currently no alternative to vector control. The dengue vector control programmes of most endemic countries have been frequently found insufficient, ineffective or both. The low success rate of vector control, continuing spread of dengue and increasing incidences of dengue, calls for a safe, effective and affordable vaccine. The ideal dengue vaccine should be affordable, free of important reactogenicity, and should induce life-long protection against infection with any of the four dengue serotypes (i.e. tetravalent) (Guzman et al., 2010).

Recently, significant progress has been made in the development of vaccine candidates. Two vaccines, one from Sanofi Pasteur is in its early Phase III trials and the other developed jointly by Walter Reed Army Institute of Research and GlaxoSmithKline is in advanced phase II trials, and several others are in phase I trials or advanced preclinical evaluation. The vaccine from Sanofi Pasteur has now been given safely to more than 5000 volunteers, including children and adults from dengue-endemic and non-endemic areas. The vaccine candidate has been shown to induce almost full seroconversion against all four serotypes in both adults and children after two to three injections

(Guy et al., 2011). A licensed vaccine for dengue is expected to be available in less than 10 years (Guzman et al., 2010).

The chapters hereafter are organized as follows: In chapter 2, we present three mathematical models of dengue by Cummings et al. (2005), Wearing and Rohani (2006) and Nagao and Koelle (2008) used in past by researchers to study the dynamics of dengue transmission. In chapter 3, a generic dengue model is introduced which encompasses the three models discussed in chapter 2 and then the methods to be used for the estimation of model parameters and model selection to determine the model that most parsimoniously fits the data. Finally, in chapter 4, we extend the generic dengue model introduced in chapter 3 to an age-structured model and present the approach for optimizing the allocation of a new dengue vaccine.

Chapter 2

Mathematical Models of Dengue

Studying the incidence and distribution of an infectious disease in a population provides crucial information about which factors most contribute to infection and transmission and allows for exploring possible interventions to control the disease. Numerous approaches have been used to understand the epidemiology of dengue fever (DF) and dengue hemorrhagic fever (DHF). There are several specific ecological characteristics of the dengue virus (DENV) which have often been explored using mathematical models. Various mathematical models for dengue infection have appeared so far, and successfully helped us to understand the different aspects of the disease. The different mathematical models formulated by dengue researchers differ from each other in the approaches taken and the underlying assumptions. For example, a series of studies by Focks et al. (1993a,b, 1995, 2000) investigated the quantitative value of models using epidemiological data and simulations, while others focused purely on qualitative patterns of equation system motivated by ecological interests (Ferguson et al., 1999; Esteva and Vargas, 2000; Esteva and Mo Yang, 2005). Despite different purposes, these studies share a common methodology for the underlying structure of population dynamics.

One of the most intriguing ecological feature of dengue is co-circulation of its four closely related dengue serotypes and the pathogenesis of DHF associated with secondary infection by a heterologous strain. Ecologically, it is of particular interest to understand co-circulation of dengue serotypes, which is a peculiar phenomena due to the competition (or interference) between strains. Usually, in the case of co-circulation of more than one strain of a disease, there is competition between strains due to cross-protective immunity between strains. However, dengue is different from

the majority of other diseases in that infection with a second, heterologous serotype enhances the severity (Nishiura, 2006). Ferguson et al. (1999) found that this fact supports persistent cyclical behavior. Moreover, the specific pathogenesis characterized by antibody-dependent enhancement (ADE) appears to be a factor permitting co-existence (Ferguson et al., 1999; Kawaguchi et al., 2003). After a person recovers from first dengue infection, they develop immune response to that dengue serotype. However, after a period of transient cross-protection, a second experience with a heterotypic virus may result in enhanced viral replication and lead to increase in severity of clinical manifestations of dengue virus infection. This situation is referred to as antibody-dependent enhancement (ADE). There are different possible hypothesized consequences of ADE. ADE is thought to increase the susceptibility of host or result in increased transmissibility of secondary infections. An alternative hypothesis is that ADE can increase the mortality and shorten the effective infectious period (Kawaguchi et al., 2003)..

We introduce three main mathematical dengue models by Cummings et al. (2005), Wearing and Rohani (2006) and Nagao and Koelle (2008). These models incorporate the co-circulation of multiple dengue serotypes, ADE and seasonality. Cummings et al. formulated a deterministic compartmental model with ADE and their analysis suggests that enhancement would decrease the transmission, except at high levels of enhancement, where the probability of extinction of the disease increases. Wearing and Rohani used a hybrid stochastic compartmental model with ADE and serotypes in the Thailand population and showed that a combination of seasonal variation in vector demography and crucially, a short-lived period of cross-immunity sufficiently explains periodicity in DHF. In their paper, Nagao and Koelle showed that decreases in dengue transmission may act to increase the incidence of DHF. This was explained by a new mathematical model with a short post-infection period in which individuals can have asymptomatic infections with new serotypes and gain immunity, which they called clinical cross-protection. This clinical cross-protection was shown to be able to account for the non-monotonic relationship between transmissibility of dengue infection and DHF incidence.

2.1 Model I

Ferguson et al. (1999) examined the effect of ADE on the dynamics of dengue transmission in a 2-serotype model, in which they assumed that ADE increases the viral load of individuals who

are experiencing a second, heterotypic infection. Thus, secondary infections act to increase the force of infection. Cummings et al. (2005) generalized the model used by Ferguson et al. to develop a model that incorporates the co-circulation of four dengue serotypes and more generally, n serotypes to study the effect of ADE in the systems of varying numbers of co-circulating serotypes. They used this model to determine the impact of ADE on the competition between serotypes in a single population. The dynamics of n serotypes circulating in a single population of constant size are given by

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \mu - S\sum_{i=1}^{n}\lambda_i(t) - \mu S, \tag{2.1a}$$

$$\frac{\mathrm{d}I_i}{\mathrm{d}t} = S\lambda_i(t) - (\sigma+\mu)I_i, \tag{2.1b}$$

$$\frac{\mathrm{d}R_i}{\mathrm{d}t} = \sigma I_i - R_i\sum_{\substack{j=1\\ j\neq i}}^{n}\lambda_j(t) - \mu R_i, \tag{2.1c}$$

$$\frac{\mathrm{d}I_{ij}}{\mathrm{d}t} = R_i\lambda_j(t) - \sigma I_{ij} - \mu I_{ij}, \tag{2.1d}$$

$$\frac{\mathrm{d}R_{**}}{\mathrm{d}t} = \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\sigma I_{ij} - \mu R_{**}, \tag{2.1e}$$

for $i, j = 1, 2, \ldots, n$ with $j \neq i$, where $S(t)$ is the fraction of population susceptible to all n serotypes, $I_i(t)$ is the fraction currently infectious with serotype i, $R_i(t)$ is the fraction recovered from infection with serotype i, $I_{ij}(t)$ is the fraction currently infectious with serotype j having previously been infected by serotype i, and $R_{**}(t)$ is the fraction who have recovered from two infections of dengue. Individuals are assumed to be immune to all serotypes of dengue after two sequential infections because infections from more than two serotypes are reported very rarely (Nisalak et al., 2003). It is assumed that primary infection with any single dengue serotype provides lifelong homotypic immunity. However for simplicity, short-term cross-immunity between serotypes and seasonality are not included in the model. ADE is assumed to increase the transmissibility of secondary infections of serotype i by a factor ϕ_i, labeled as ADE factor. For model simplicity, the transmission coefficient β is assumed to be equal for each serotype. The force of infection for serotype i, which is the per

capita rate of acquisition of infection with serotype i,

$$\lambda_i(t) = \beta \left(I_i(t) + \phi_i \sum_{\substack{k=1 \\ k \neq i}}^{n} I_{ki}(t) \right), \tag{2.1f}$$

is assumed to be proportional to the total fraction infectious with serotype i but with secondary infections weighted by the enhancement factor, ϕ_i (Ferguson et al., 1999). An ADE factor of 1 corresponds to no enhancement, whereas an ADE factor of 2 means that secondary infections contribute twice as much as primary infections to the force of infection. The population is assumed constant with a birth rate μ, exactly equal to death rate for model simplicity. Due to the low case fatality rate of dengue infection in Thailand (approximately 1 death per 10,000 infections) (Shepard et al., 2004), mortality rates in the population are assumed to be independent of disease status. The rate of recovery, given by σ, is assumed to be equal for both primary and secondary infections.

Using this model, Cummings et al. determined the behavior of the systems while changing the ADE factor for all serotypes at once (denoted by ϕ_{all}) and for changing the value of the ADE factor for just one serotype (without loss of generality choosing ϕ_1 to vary) while holding the ADE factor equal to 1 for all other serotypes. The behavior of solutions changed dramatically at different values of ϕ_{all} (Figure 2.1) and ϕ_1. Although equal base transmission rates for each of the serotypes was assumed, the prevalence of each serotype was not equal at all times. A general trend found was increasingly large oscillations with increasing enhancement.

The model was used to quantitatively measure the effect of ADE on the prevalence of a dengue serotype undergoing enhancement by comparing the serotype's equilibrium prevalence at various levels of enhancement in a system of n ($n = 2$ to 6) circulating serotypes and it was observed that the solutions oscillate for large values of ADE factor. However, for reasonably small changes in the ADE factor of the enhanced serotype ($\phi_1 < 2$), the equilibrium point of the system was stable. Moreover, equilibrium prevalence of the enhanced serotype increased as the ADE factor increased, and for every unit increase in ADE factor, the increase in prevalence that occurred was greater in systems with greater number of co-circulating serotypes.

The work done by Cummings et al. suggests that ADE initially provides an advantage to those serotypes that undergo enhancement compared to those serotypes that do not, and that this advantage increases with increasing numbers of co-circulating serotypes. However, as enhancement

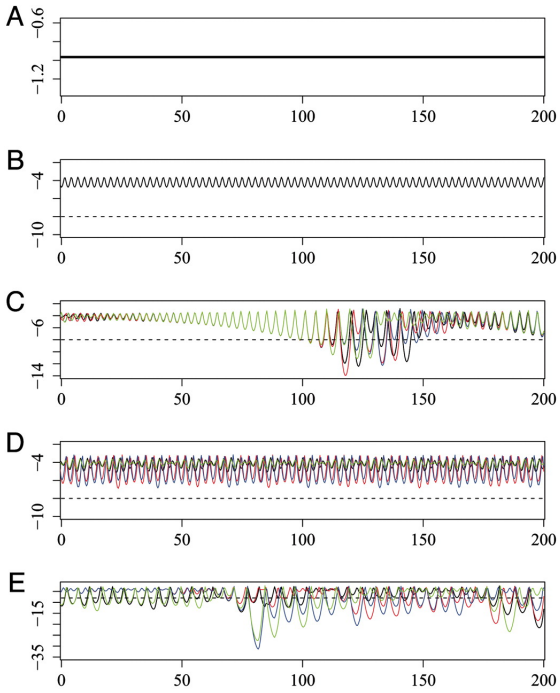

Figure 2.1: Example time series from simulation with ϕ_i equal for all i. The x axis is time in years and the y axis is log of number of infections. The different colors represent the different serotypes.

provides substantial advantage, the enhanced strains overcompete with other serotypes and exhaust the pool of individuals susceptible to secondary infections. The loss of this resource causes large amplitude cycles for the enhanced serotype and eliminates the advantage gained by ADE (because there are very few secondary infections).

2.2 Model II

To understand the dynamics of dengue epidemics, Wearing and Rohani (2006) adopted a theoretical approach that combined both ecological and immunological mechanisms. They formulated a general model for a vector-borne multi-strain pathogen and applied it to dengue to determine whether population-level data are consistent with ADE, variation in virulence or the hypothesis of temporary heterologous cross-immunity. The model is based on traditional susceptible-exposed-infectious-recovered classes for the human host and susceptible-exposed-infectious classes for mosquito host (Anderson and May, 1991). The cross-immunity and ADE are assumed to be related to cross-reactive antibody levels that wane over time, so that any period of heterotypic

cross-immunity or ADE is temporary. The model also includes an important ecological ingredient that is often ignored: seasonal variation in mosquito recruitment, which gives rise to temporal variation in transmission. ADE is incorporated in the model based on assumptions about immunological consequences of ADE: that it can result in increased transmission, by either increased susceptibility to infection after a primary infection or increased infectiousness with a secondary infection. As with ADE, variation in virulence is assumed to increase transmission or increase mortality after infection with a particular serotype.

The model with 4 serotypes is given by the differential equations

$$\frac{\mathrm{d}S_0}{\mathrm{d}t} = (N_H - S_0)\mu_H - \sum_{i=1}^{4} \lambda_{Vi} S_0, \tag{2.2a}$$

$$\frac{\mathrm{d}E_i}{\mathrm{d}t} = \lambda_{Vi} S_0 - \sum_{\substack{j=1 \\ j\neq i}}^{4} \phi_j \lambda_{Vj} E_i - (\sigma_H + \mu_H) E_i, \tag{2.2b}$$

$$\frac{\mathrm{d}I_i}{\mathrm{d}t} = \sigma_H E_i - \sum_{\substack{j=1 \\ j\neq i}}^{4} \phi_j \lambda_{Vj} I_i - (\gamma_i + \mu_H) I_i, \tag{2.2c}$$

$$\frac{\mathrm{d}C_i}{\mathrm{d}t} = \gamma_i I_i - \sum_{\substack{j=1 \\ j\neq i}}^{4} \xi_j \lambda_{Vj} C_i - (\delta_i + \mu_H) C_i, \tag{2.2d}$$

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} = (1-\rho_i)\delta_i C_i - \sum_{\substack{j=1 \\ j\neq i}}^{4} \chi_j \lambda_{Vj} S_i - \mu_H S_i, \tag{2.2e}$$

$$\frac{\mathrm{d}V_{Si}}{\mathrm{d}t} = \left(kN_H\left[1 - a\cos(2\pi t)\right] - V_{Si}\right)\mu_V - \sum_{\substack{j=1 \\ j\neq i}}^{4} \lambda_{Hj} V_{Si}, \tag{2.2f}$$

for $i = 1, 2, 3, 4$. Newborns are fully susceptible to either serotype and enter the class of immunologically naive individuals, S_0. The background mortality rate, μ_H is assumed to be equal to the birth rate and total host population is assumed to be of constant size, N_H. The variables λ_{Vi} and λ_{Hi} are the serotype-specific forces of infection exerted by the vectors and humans, respectively. Once susceptible individuals get primary infection with serotype i, they enter the exposed (infected but not yet infectious) class, E_i, and have a relative probability of contracting an infection with another serotype simultaneously, modulated by the coinfection parameter, ϕ_j. After this latent period (average length given by $1/\sigma_H$), if the individuals have not been coinfected they enter the class I_i, where they continue to have same chance ($\phi_j\lambda_j/N_H$) of becoming coinfected with another

serotype. Immediately after the infectious period (average length given by $1/\gamma_i$), individuals who have been exposed to only a single infection enter the refractory class, C_i, and are temporarily immune to the other serotype. The partial immunity is represented by $0 < \xi_j < 1$, while $\xi_j = 0$ gives complete immunity. Disease-induced mortality is included by discounting a proportion of those leaving refractory phase. This is represented by the per capita infection-induced mortality probabilities, ρ_i. To incorporate one of the possible consequence of antibody-dependent enhancement (ADE) in the model, a parameter χ_j is included to explore the effects of increased susceptibility to infection with the second serotype (by defining $\chi_j > 1$). There is no strong evidence to suggest that vertical transmission (mother to child) of dengue virus is important to the transmission cycle between humans and mosquitoes. Therefore the mosquito larvae are assumed to emerge as fully susceptible adults (V_{Si}). In the absence of seasonality ($a = 0$), recruitment to the susceptible vector class is proportional to the human population size, so that vector population can be related to the average number of mosquitoes per person, k. The dynamics of the forces of latency, ϵ_{Hi}, and force of infection, λ_{Hi}, must defined. In addition, a compartment for all individuals who are no longer susceptible to either infection (S_{**}) is also included which may include those individuals who are still exposed or infectious with either serotype (i.e. also included within ϵ_{Hi} or λ_{Hi}). The equations for these compartments are given by

$$\frac{\mathrm{d}\epsilon_{Hi}}{\mathrm{d}t} = \lambda_{Vi} S_0 + \eta_i \lambda_{Vi} \left[\sum_{\substack{j=1 \\ j \neq i}}^{4} \phi_i \left(E_j + I_j \right) + \xi_i C_j + \xi_i S_j \right] - (\sigma_H + \mu_h)\epsilon_{Hi}, \tag{2.2g}$$

$$\frac{\mathrm{d}\lambda_{Hi}}{\mathrm{d}t} = \frac{\beta_i \sigma_H \epsilon_{Hi}}{N_H} - (\gamma_i + \mu_H)\lambda_{Hi}, \tag{2.2h}$$

$$\frac{\mathrm{d}S_{**}}{\mathrm{d}t} = \sum_{i=1}^{4} \sum_{\substack{j=1 \\ j \neq i}}^{4} (1 - \rho_j)(1 - \rho_x)\lambda_{Vj} \left[\phi_j \left(E_i + I_i \right) + \xi_j C_i + \chi_j S_i \right] - \mu_H S_{**}, \tag{2.2i}$$

for $i = 1, 2, 3, 4$. The transmission rate from humans to mosquitoes is β_i. The parameter η_i is used to explore the potential of ADE to increase infectiousness of, rather than susceptibility to, secondary infections. Only one of the ADE factors is explored at a time: either η_i is varied with $\chi_i = 1$ or χ_i is varied with $\eta_i = 1$. It is assumed in the model that disease-induced mortality occurs at the end of the infection so that the effective infectious period remains unchanged by mortality as a result of infection and there is no trade off between mortality and duration of transmission. The consequence of this assumption for ADE is that increased mortality after a secondary infection ($\rho_x > 0$) has no

discernible dynamical impact when there are only two serotypes present. However, the alternate assumption of increased susceptibility after a primary infection modulated by the parameter $\chi_i > 1$ does induce qualitative dynamical changes. Additionally, the mosquito vectors can acquire multiple infections and progress through latent stage, ϵ_{Vi}, before they become infectious. Unlike human hosts, the vectors do not recover, so their force of infection (λ_{Vi}) is only depleted by mosquito mortality:

$$\frac{\mathrm{d}\epsilon_{Vi}}{\mathrm{d}t} = \lambda_{Hi} V_{Sj} - (\sigma_V + \mu_V)\epsilon_{Vi}, \tag{2.2j}$$

$$\frac{\mathrm{d}\lambda_{Vi}}{\mathrm{d}t} = \frac{\alpha_i \sigma_V \epsilon_{Vi}}{N_H} - \mu_V \lambda_{Vi}, \tag{2.2k}$$

where α_i denotes the transmission rate.

For simplicity, the model presented above assumes that any enhancement after a period of temporary ADE is permanent, but the model can be extended to include temporary ADE by adjusting the differential equations for the S_i and ϵ_H and adding 4 new compartments (A_i) with additional parameters, ω_i, which represent the rate at which cross-enhancing antibody levels from serotype i wane to neither protective nor enhancing levels. These modifications are given by

$$\frac{\mathrm{d}A_i}{\mathrm{d}t} = (1-\rho_i)\delta_i C_i - \chi_j \lambda_{Vj} A_i - (\omega_i + \mu_H)A_i \tag{2.3a}$$

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} = \omega_i A_i - \lambda_{Vj} S_i - \mu_H S_i \tag{2.3b}$$

$$\frac{\mathrm{d}S_{**}}{\mathrm{d}t} = \sum_{i=1}^{4} \sum_{\substack{j=1 \\ j\neq i}}^{4} (1-\rho_j)\lambda_{Vj} \left\{(1-\rho_x)\left[\xi_j(E_i + I_i) + \phi_j C_i + \chi_j A_i\right] + S_i\right\} - \mu_H S_{**} \tag{2.3c}$$

$$\frac{\mathrm{d}\epsilon_{Hi}}{\mathrm{d}t} = \lambda_{Vi}\left[S_0 + \xi_i(E_j + I_j) + \phi_i C_j + \chi_i A_j + S_j\right] - (\rho_H + \mu_H)\epsilon_{H1}, \tag{2.3d}$$

for $i = 1, 2, 3, 4$.

This model was used to explore the dynamical consequences of different assumptions regarding dengue infection and transmission for the simpler case of two serotypes. The model dynamics were compared with serological time series from long-term study of dengue in Thai children together with total dengue cases reported to the Thai Ministry of Public Health. The aggregate data reveal fluctuations of between 3 and 4 years in dengue case reports, whereas individual serotypes generally cycle in and out of phase with longer periods than the aggregate data (Wearing and Rohani, 2006). Seasonality was found to be necessary to explain intraannual variation in monthly dengue

incidences, but has a lesser impact on interannual dynamics. Therefore the immunological mechanisms are investigated with a fixed amplitude of seasonality. The model examines how a combination of temporary cross-immunity and the strength of ADE (Figure 2.2 A, D and G), duration of ADE (Figure 2.2 B, E and H) or asymmetry in virulence (Figure 2.2 C, F and I) influences individual serotype outbreaks (Figure 2.2 A–C), aggregate dynamics (Figure 2.2 D–F) and the correlation between serotypes (Figure 2.2 G–I).

For very short periods of cross-immunity (< 2 months), annual cycles in serotype incidence were driven by annual variation in vector recruitment, unless ADE was significantly increased, which leads to multiannual synchronized serotype cycles. For more realistic periods of cross-immunity, increasing ADE reduces the intrinsic epidemiological period for individual serotypes, and the aggregate data have an even shorter periodic signature because individual serotype epidemics do no coincide. This observation from (Figure 2.2 A, D and G) suggests that a minimum period of cross-immunity is necessary to obtain the empirically observed periods of 3 years. Moreover, ADE alone gives rise to much longer cycles. Same qualitative patterns emerged if ADE was a transient process (Figure 2.2 B, E and H). While considering differential serotype virulence in model, main focus is on increased mortality due to a primary or secondary infection with a particular serotype. It is observed that, without a period of temporary cross-immunity, increases in mortality do not result in the observed cyclic incidence of dengue (Figure 2.2 C, F and I) whereas with temporary cross-immunity, cycles of 4 years are observed, which is consistent with the dengue mortality statistics.

In summary, Wearing and Rohani observed that patterns generated solely by ADE or heterogeneity in virus virulence are not consistent with serotype-specific notification data in important ways. Furthermore, a short-lived period of cross-immunity induced by infection with another serotype seems to be sufficient to match the epidemiological observations.

2.3 Model III

Despite public health measures aimed at dengue vector control in many countries in Asia and South America, DHF incidence rates do not appear to be declining, which leads to the effectiveness of vector control in reducing dengue transmissibility being questioned. So, Nagao and Koelle (2008) used epidemiological data for dengue from Thailand to explore the reasons behind this observation. They found that dengue transmission has indeed fallen, however, the decline in transmission is

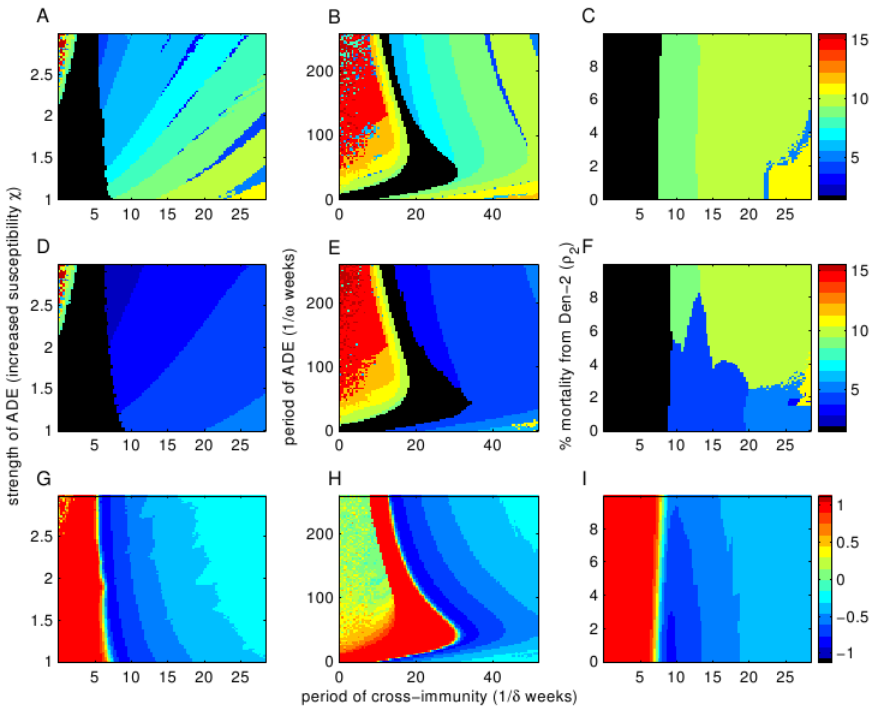

Figure 2.2: ADE or variation in virulence; the dominant period and correlation of two dengue serotypes from deterministic model simulations. A, D, and G explore the effects of permanent ADE, B, E, H explore the effects of temporary period of ADE. C, F, and I show the effects of increasing mortality due to infection with only one of the serotypes.

not associated with the decrease in DHF incidence. Instead, it was found that decrease in dengue transmission may act to increase the incidence of DHF. To understand this dynamic, Nagao and Koelle formulated three mathematical models of dengue. The models differ in their assumptions of transient between-serotype cross-protection. The first two models are used to show that ADE alone or ADE together with short-term cross-protection can not reproduce the empirical observations of higher DHF incidence rates at lower transmission rates. The previously unconsidered third model is used to show ADE together with short-term clinical cross-protection, which prevents clinical illness but allows serotype conversion, captures the empirical pattern. Finally it is shown that the deterministic simulations of the clinical cross-protection model reproduces the patterns of DHF interannual variability as well as observed patterns of interannual serotype fluctuations.

In their model, Nagao and Koelle compartmentalized the hosts according to their infection histories. Transmission of the virus occurs directly from infected individuals to susceptible hosts, without explicitly modeling mosquitoes. The model assumes that primary infection with any single dengue serotype provides lifelong homotypic immunity. The four-dengue serotype model has $2^4 = 16$ classes of uninfected hosts $S_{\mathcal{J}}$, with $\mathcal{J}$ denoting the set of strains a host has previously been exposed to and so is immune to. Therefore, the $S_{\mathcal{J}}$ includes the completely naive hosts $S_{\emptyset}$, as well as hosts who have been exposed to one or more of the four dengue serotypes. There are $32 = 4 \cdot 2^{4-1}$ classes of infected individuals, with $I^i_{\mathcal{J}}$ representing the number of individuals currently infected with strain i and having previously been exposed to strains in set $\mathcal{J}$. The differential equations are

$$\frac{\mathrm{d}S_{\emptyset}}{\mathrm{d}t} = \mu N - \sum_{j \in \mathcal{K}} \lambda^j S_{\emptyset} - \mu S_{\emptyset} \tag{2.4a}$$

$$\frac{\mathrm{d}S_{\mathcal{J}}}{\mathrm{d}t} = \sum_{j \in \mathcal{J}} \nu I^j_{\mathcal{J}\setminus\{j\}} - \sum_{i \notin \mathcal{J}} \sigma^i_{\mathcal{J}} \lambda^i S_{\mathcal{J}} - \mu S_{\mathcal{J}}, \qquad \text{for all } \mathcal{J} \neq \emptyset, \tag{2.4b}$$

$$\frac{\mathrm{d}I^i_{\mathcal{J}}}{\mathrm{d}t} = \sigma^i_{\mathcal{J}} \lambda^i S_{\mathcal{J}} - (\mu + \nu) I^i_{\mathcal{J}} \tag{2.4c}$$

where the force of infection for strain i is

$$\lambda^i = \sum_{\mathcal{J} \subset \mathcal{K}\setminus\{i\}} \frac{\beta^i_{\mathcal{J}} I^i_{\mathcal{J}}}{N} \tag{2.4d}$$

$1/\mu$ is the average host life span; $1/\nu$ is the average duration of infectiousness; $\mathcal{K} = \{1, 2, 3, 4\}$ is the set of dengue serotypes, DEN-1 to DEN-4; $\sigma^i_{\mathcal{J}}$ is the susceptibility-reduction factor; and $\beta^i_{\mathcal{J}}$ is

the transmission rate. The model incorporates the effects of ADE using the susceptibility factor σ_J^i and the transmission rate $\beta_{\mathcal{J}}^i$. In their first model, which assumes no period of heterologous cross-protection, $\sigma_J^i = 1$ for all i and $\mathcal{J}$.

To keep track of the dynamics of DHF cases (a subset of infected individuals) separately, another set of $32 = 4 \cdot 2^{4-1}$ compartments are added:

$$\frac{\mathrm{d}C_{\mathcal{J}}^i}{\mathrm{d}t} = p_{\mathcal{J}}^i \sigma_{\mathcal{J}}^i \lambda^i S_{\mathcal{J}} - (\mu + \nu) C_{\mathcal{J}}^i \tag{2.4e}$$

where $p_{\mathcal{J}}^i$ is the probability that an individual previously infected with strains $\mathcal{J}$ and becoming infected with strain i develops DHF. For the models with a transient period of heterologous cross-protection, the above equations are modified by adding $16 = 2^4$ temporarily immune classes, $T_{\mathcal{J}}$, which is indexed in the same way as the susceptible classes, $S_{\mathcal{J}}$. To include temporary immunity, the dynamics of the uninfected hosts, $S_{\mathcal{J}}$, are modified to

$$\frac{\mathrm{d}S_{\mathcal{J}}}{\mathrm{d}t} = \delta T_{\mathcal{J}} - \sum_{i \notin \mathcal{J}} \sigma_{\mathcal{J}}^i \lambda^i S_{\mathcal{J}} - \mu S_{\mathcal{J}} \tag{2.5}$$

where $1/\delta$ is the average duration of heterologous cross immunity. The dynamics of the $T_{\mathcal{J}}$ classes are dependent on the type of heterologous cross-protection. For the model with clinical cross-protection, the equations for the $T_{\mathcal{J}}$ classes are

$$\frac{\mathrm{d}T_{\mathcal{J}}}{\mathrm{d}t} = \sum_{i \in \mathcal{J}} \nu I_{\mathcal{J} \setminus i}^i - (\mu + \delta) T_{\mathcal{J}} \tag{2.6}$$

For the model with clinical cross-protection, the equations for the $T_{\mathcal{J}}$ classes are instead

$$\frac{\mathrm{d}T_{\mathcal{J}}}{\mathrm{d}t} = \sum_{i \in \mathcal{J}} \nu I_{\mathcal{J} \setminus i}^i - \sum_{i \notin \mathcal{J}} \lambda^i T_{\mathcal{J}} + \sum_{i \in \mathcal{J}} \lambda^i T_{\mathcal{J} \setminus i} - (\mu + \delta) T_{\mathcal{J}}. \tag{2.7}$$

The term $\sum_{i \notin \mathcal{J}} \lambda^i T_{\mathcal{J}}$ is the loss of individuals from the class $T_{\mathcal{J}}$ through sub-clinical infection with a new serotype. The term $\sum_{i \in \mathcal{J}} \lambda^i T_{\mathcal{J} \setminus i}$ is the gain of individuals to class $T_{\mathcal{J}}$ through sub-clinical infection from another temporarily immune class. In both models, $T_\emptyset$ and $\frac{\mathrm{d}T_\emptyset}{\mathrm{d}t}$ are zero because individuals only gain temporary immunity following an infection. For the simulations, a small immigration rate m and a small level of seasonal forcing ϵ is included. These addition result

in a minor change for the definition of the forces of infection, which are now

$$\lambda^i = [1 - \epsilon \cos(2\pi t)] \left(\sum_{\mathcal{J} \subset \mathcal{K} \setminus i} \frac{\beta^i_{\mathcal{J}} I^i_{\mathcal{J}}}{N} + m \right), \tag{2.8}$$

where t is given in years and m is the per-host immigration rate.

The first model considered by Nagao and Koelle represents the class of models such as Ferguson et al. (1999) in which ADE sets in immediately after recovery from infection. The second model includes a transient period of heterologous cross-protection against reinfection and captures the immunological assumptions present in dengue model by Wearing and Rohani (2006). The third model also includes a transient period of heterologous cross-protection. However, instead of modeling cross-protection classically, this model assumes clinical cross-protection: A challenge with a previously unexperienced serotype during the cross-protected period does not result in clinical manifestation of DHF nor in transmissible infection; however, the challenge does result in a gain of immunity toward the challenging serotype.

Figure 2.3 shows the results of simulating these three models deterministically over a wide range of basic reproductive number, R_0, which is the mean number of secondary cases a typical single infected case will cause in a population with no immunity to the disease and in the absence of interventions to control the infection. The first two models generated a monotonic relationship between R_0 and DHF incidence at higher R_0 which means at higher R_0, the fraction of infected individuals in the population is higher and there are more secondary infections leading to higher DHF incidence. In contrast to the first two models, the third model with clinical cross-protection, reproduces the non-monotonic relationship between R_0 and DHF incidence and is consistent with empirical observation. A long, slightly decreasing plateau of DHF incidence is observed at high R_0. The pattern is evident for any duration of cross-protection, although the negative relationship between R_0 and DHF incidence is more pronounced for longer durations of protection. The negative relationship between R_0 and DHF incidence at high R_0 is attributed to the increasing number of individuals that gain heterologous serotypes while being transiently cross-protected from becoming infectious and from manifesting DHF clinically.

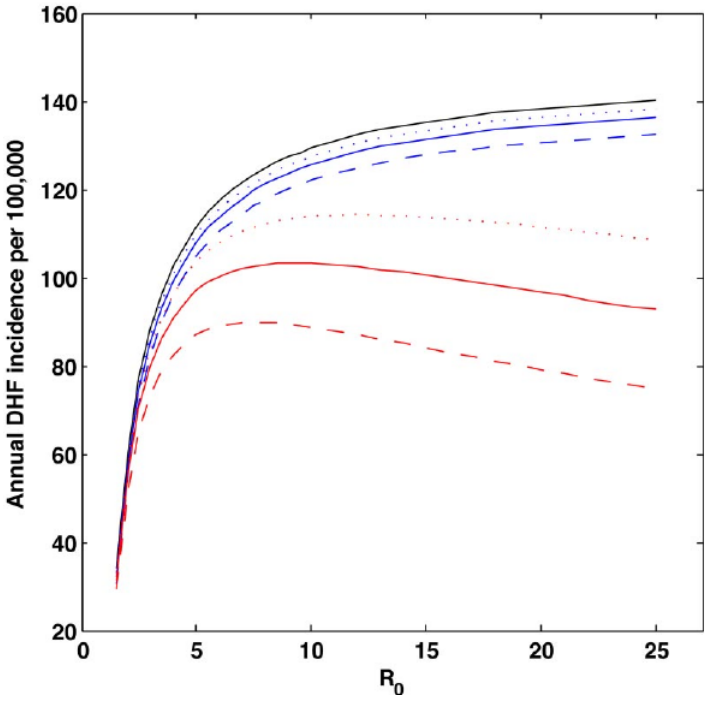

Figure 2.3: DHF incidence rates as a function of the basic reproduction number, R_0. Results from the model without temporary cross-protection are shown in black, results from the model with classical cross-protection are shown in blue, and results from the model with clinical cross-protection are shown in red. For the two models with transient cross-protection, three durations of cross-protection were considered: $1/\delta = 1/2$ year (dotted),$1/\delta = 1$ year (solid) and $1/\delta = 2$ year (dashed).

Chapter 3

Parameter Estimation and Model Selection

On the basis of the dengue models of Cummings et al. (2005), Wearing and Rohani (2006) and Nagao and Koelle (2008) discussed in the previous chapter, we formulated a generic mathematical model which encompasses all the three models and their hypotheses about dengue epidemiology, allowing us to explore all the dynamical aspects of dengue explained by these models. In general, this model has many parameters which we will estimate by fitting the clinical data on dengue from Thailand (Nisalak et al., 2003). As in the previous models, this model allows for the exclusion of the different hypotheses about dengue epidemiology by setting the associated parameter to a default value and thereby reducing the number of parameters which must be estimated. We will systematically explore whether the data are consistent with the epidemiological hypotheses by determining the goodness of fit of the model to the data for each combination of hypotheses and use the techniques of model selection to find the model that most parsimoniously agrees with the data.

3.1 Dengue Mathematical Model

We formulated this new model by modifying the model of Nagao and Koelle. The model defines the classes of the uninfected people $S_{\mathcal{J}}$, the classes of infected people $I^i_{\mathcal{J}}$ and the classes of temporarily immune people $T_{\mathcal{J}}$ in the same way as in the Nagao and Koelle's model. The new

dengue mathematical model is given by

$$\frac{\mathrm{d}S_\emptyset}{\mathrm{d}t} = \mu N - \sum_{i\in\mathcal{K}} \lambda^i(t) S_\emptyset - \mu S_\emptyset, \tag{3.1a}$$

$$\frac{\mathrm{d}S_\mathcal{J}}{\mathrm{d}t} = \delta T_\mathcal{J} - \sum_{i\notin\mathcal{J}} \sigma^i_\mathcal{J} \lambda^i(t) S_\mathcal{J} - \mu S_\mathcal{J}, \qquad \text{for all } \mathcal{J} \neq \emptyset, \tag{3.1b}$$

$$\frac{\mathrm{d}I^i_\mathcal{J}}{\mathrm{d}t} = \sigma^i_\mathcal{J} \lambda^i(t) S_\mathcal{J} - (\mu + \gamma^i_\mathcal{J} + \nu^i_\mathcal{J}) I^i_\mathcal{J}, \qquad \text{for all } \mathcal{J} \text{ and } i \notin \mathcal{J}, \tag{3.1c}$$

$$\frac{\mathrm{d}T_\mathcal{J}}{\mathrm{d}t} = \sum_{i\in\mathcal{J}} \gamma^i_\mathcal{J} I^i_{\mathcal{J}\setminus i} + \chi \left(\sum_{i\in\mathcal{J}} \lambda^i(t) T_{\mathcal{J}\setminus i} - \sum_{i\notin\mathcal{J}} \lambda^i(t) T_\mathcal{J} \right) - (\mu + \delta) T_\mathcal{J}, \qquad \text{for all } \mathcal{J} \neq \emptyset, \tag{3.1d}$$

where the force of infection for strain i with a small level of seasonal forcing ϵ is

$$\lambda^i(t) = [1 + \epsilon \cos(2\pi t)] \left(\sum_{\mathcal{J}\subset\mathcal{K}\setminus i} \frac{\beta^i_\mathcal{J} I^i_\mathcal{J}}{N} \right), \qquad \text{for all } i \in \mathcal{K}. \tag{3.1e}$$

The set of dengue serotypes, DEN-1 to DEN-4, are represented by $\mathcal{K} = \{1, 2, 3, 4\}$ so that the current serotype index is $i \in \mathcal{K}$ and the previous serotype index is $\mathcal{J} \subset \mathcal{K}$. The parameter $1/\mu$ is the average life span; $1/\delta$ is the average duration of heterologous cross immunity; $\sigma^i_\mathcal{J}$ is the susceptibility-reduction factor, relative to people who have had no previous infections, so $\sigma^i_\emptyset = 1$; $\nu^i_\mathcal{J}$ is the mortality rate for the individuals currently infected with strain i having previously experienced strains in set $\mathcal{J}$; $1/\gamma^i_\mathcal{J}$ is the average duration of infectiousness; $\beta^i_\mathcal{J}$ is the transmission rate; and the parameter χ is the odds of a temporarily immune person gaining clinical cross-protection from an infectious mosquito bite, relative to the chance of infection of a person with no previous infections.

With this general model defined, the parameters can be specified to explore the different possible hypothesized effects of ADE. This model is quite general, allowing differences in parameters for different current serotypes i and different previous serotypes $\mathcal{J}$. Due to the huge numbers of parameters this allows in general and the lack of strong evidence for serotype-specific differences in dengue epidemiology, we will take the parameters to vary only with the number of previous serotypes, $|\mathcal{J}|$, and, more specifically, whether a person has had none vs. 1 or more previous serotypes, as is generally hypothesized to be the determining factor for ADE.

3.1.1 Enhancement of Transmission

One hypothesis for the consequence of ADE on the dynamics of dengue is that ADE causes increased transmission of dengue virus, which is controlled by the parameter $\beta^i_{\mathcal{J}}$ in our model. Without enhancement of transmission, we define

$$\beta^i_{\mathcal{J}} = \beta_1, \qquad \text{for all } \mathcal{J} \text{ and } i, \tag{3.2}$$

so that all infections have equal transmission, regardless of the current or previous serotypes. In contrast, to incorporate the hypothesis of increased transmission in our model, we choose $\beta^i_{\mathcal{J}}$ to be higher for the individuals who have previously been exposed to one or more dengue strains than the individuals who have not been exposed to any of the dengue strains. So we define the $\beta^i_{\mathcal{J}}$ as

$$\beta^i_{\mathcal{J}} = \begin{cases} \beta_1 & \text{if } |\mathcal{J}| = 0, \\ \beta_2 & \text{if } |\mathcal{J}| \geq 1, \end{cases} \qquad \text{for all } i, \tag{3.3}$$

where $\beta_2 > \beta_1$.

3.1.2 Enhancement of Susceptibility

The susceptibility-reduction factor $\sigma^i_{\mathcal{J}}$ is used to explore the increased susceptibility as a consequence of ADE. Without enhancement of susceptibility, we define

$$\sigma^i_{\mathcal{J}} = 1, \qquad \text{for all } \mathcal{J} \text{ and } i, \tag{3.4}$$

so that all individuals have equal susceptibility, regardless of the current serotype and previous serotypes they might have had. In contrast, to incorporate increased susceptibility as an effect of ADE, an individual who has experienced infection with at least 1 dengue serotype previously is more susceptible than an individual who has not experienced infection from any serotypes. So increased susceptibility is incorporated in the model as

$$\sigma^i_{\mathcal{J}} = \begin{cases} 1 & \text{if } |\mathcal{J}| = 0, \\ \sigma_1 & \text{if } |\mathcal{J}| \geq 1, \end{cases} \qquad \text{for all } i, \tag{3.5}$$

where $\sigma_1 > 1$.

Moreover, it is not clear whether infection with third and fourth serotypes is possible. We will also use the hypotheses that they are possible, as is incorporated above, or that they are not possible by setting

$$\sigma^i_{\mathcal{J}} = 0 \qquad \text{for } |\mathcal{J}| \geq 2 \text{ and all } i, \tag{3.6}$$

in addition to the possible enhancement of susceptibility to second infections.

3.1.3 Mortality

Higher dengue-induced mortality is generally associated with secondary infections, which is controlled by the parameter $\nu^i_{\mathcal{J}}$ in our model. Without enhancement of mortality, we define

$$\nu^i_{\mathcal{J}} = \nu_1, \qquad \text{for all } \mathcal{J} \text{ and } i, \tag{3.7}$$

so that all individuals have equal mortality, regardless of the current serotype and previous serotypes they might have had. On the other hand, we include increased mortality in the model by

$$\nu^i_{\mathcal{J}} = \begin{cases} \nu_1 & \text{if } |\mathcal{J}| = 0, \\ \nu_2 & \text{if } |\mathcal{J}| \geq 1, \end{cases} \qquad \text{for all } i, \tag{3.8}$$

where $\nu_2 > \nu_1$.

3.1.4 Duration of Infection

To our knowledge, no hypothesis has been formulated about duration of infection ($1/\gamma^i_{\mathcal{J}}$ in our model) being different depending on a person's previous number of serotypes. With no difference in duration of infection, we define

$$\gamma^i_{\mathcal{J}} = \gamma_1, \qquad \text{for all } \mathcal{J} \text{ and } i, \tag{3.9}$$

so that all individuals have equal duration, regardless of the current serotype and previous serotypes they might have had. On the other hand, we include a longer duration of infection in the model by

$$\gamma^i_{\mathcal{J}} = \begin{cases} \gamma_1 & \text{if } |\mathcal{J}| = 0, \\ \gamma_2 & \text{if } |\mathcal{J}| \geq 1, \end{cases} \qquad \text{for all } i, \tag{3.10}$$

where $\gamma_2 < \gamma_1$.

3.1.5 Cross-Protection

There are two different hypotheses about temporary cross-protection that have been proposed. One hypothesis is that there is a period of cross-protection after an infection during which people cannot be infected with another serotype. The other hypothesis is that during this period of cross-protection, people can be infected with another serotype, but the cross-protection prevents any symptoms whatsoever, so they have an asymptomatic infection resulting in immunity to this new serotype. This is Nagao and Koelle's clinical cross-protection and we refer to cross-protection without the possibility of asymptomatic infection as traditional cross-protection.

The parameter δ is the rate of leaving the cross-protection state, $T_{\mathcal{J}}$, so that the mean duration of cross-protection is $1/\delta$. The assumption of no cross protection is then formally taking $1/\delta \to 0$ or $\delta \to \infty$, given by the simplified model created by removing the $T_{\mathcal{J}}$ classes from the model and replacing the differential equations for $S_{\mathcal{J}}$ for $\mathcal{J} \neq \emptyset$ with

$$\frac{\mathrm{d}S_{\mathcal{J}}}{\mathrm{d}t} = \sum_{i \in \mathcal{J}} \gamma^i_{\mathcal{J}} I^i_{\mathcal{J}\setminus\{i\}} - \sum_{i \notin \mathcal{J}} \sigma^i_{\mathcal{J}} \lambda^i(t) S_{\mathcal{J}} - \mu S_{\mathcal{J}}, \qquad \text{for all } \mathcal{J} \neq \emptyset, \tag{3.11}$$

so that upon recovery from an infection with one serotype, people immediately become susceptible to the serotypes they have not yet experienced.

On the other hand, with cross-protection, the parameter χ in the model (3.1) allows us to switch between a model with traditional cross-protection and model with clinical cross-protection. Setting $\chi = 0$ excludes clinical cross-protection, giving traditional cross-protection, while allowing $\chi > 0$ includes clinical cross-protection.

3.2 Parameter Estimation

In order to apply proposed mathematical models to study the dynamics of the disease, we need to resolve few important problems, one of which is to estimate the unknown parameters in the models. It is challenging to resolve this problem for a system of differential equations. Generally, there is no closed-form solution and there are many unknown parameters. In addition, direct data on population-level parameters such as transmission or susceptibility are extremely limited, due to an inability to feasibly conduct experiments of infection in people, and instead must be estimated from indirect population-level data. Researchers usually substitute some of the unknown parameters with estimates from previous studies. Here we suggest a Bayesian approach to find good estimates for the unknown parameters. In Bayesian terminology, the information from previous studies is regarded as prior knowledge, which is combined with the clinical data to perform the statistical inference on the unknown parameters. We outline this Bayesian approach which uses Monte Carlo Markov Chain (MCMC) simulation to estimate the unknown parameters of the model.

Given a set of differential equations with unknown (free) parameters θ and the clinical data $D(t_i)$ at the discrete time points $\{t_1, t_2, ...\}$, the aim is to find a set of free parameters so that the model fits the data at those time points. Let $y(t_i \mid \theta)$ be the time series produced by the mathematical model at the same discrete time point t_i's for which data is available. An error function is assigned to this data and aim is to minimize the error. The most common error function is the mean square error, which can be written as:

$$E = \sum_i \Big(D_i - y(t_i \mid \theta) \Big)^2 \tag{3.12}$$

The classic way to minimize the error is to take the derivative of the error function with respect to parameters and find the minimum points. However, for most of the mathematical models this is done numerically using some iterative schemes. Moreover, the error functions may have multiple local minima and we need to find the global optimum which can be a very difficult task. We use the Bayesian MCMC approach to tackle this difficulty. We start with assuming some probability distribution for our error function. Generally a normal distribution is a good assumption for error distribution. Hence, we assume that the error function in (3.12), obeys a normal distribution with

zero mean and standard deviation σ. We can then write

$$P(E) \propto \exp\left(-\frac{\sum_i (D_i - y(t_i \mid \theta))^2}{2\sigma^2}\right) \tag{3.13}$$

Now our model points $y(t_i \mid \theta)$ depend on the free parameters, so it can be rewritten as a likelihood function of the parameters given the data:

$$L(\theta) \propto \exp\left(-\frac{\sum_i (D_i - y(t_i \mid \theta))^2}{2\sigma^2}\right) \tag{3.14}$$

In order to use Bayesian inference to estimate the parameters, it is assumed that parameters have probability distribution and we need to have a prior distribution for all the free parameters. If we have some prior information about the parameters, then it can be used to assign some prior distribution to the free parameters. However, if we do not know any specific, definite information which can be used to assign the prior distribution to the parameters, then we can choose a prior which is flat. In this case the prior is called a noninformative prior. A noninformative prior expresses a vague or general information about the parameters such as "the parameter is positive" or "the parameter is less than some limit". Once prior distribution is defined for the unknown parameters, we can write the probability of parameters given the data points D using the Bayes theorem:

$$P(\theta \mid D) = \frac{P(D \mid \theta)P(\theta)}{P(D)}, \tag{3.15}$$

where $P(D \mid \theta)$ is the likelihood function $L(\theta)$ from (3.14), $P(\theta \mid D)$ is called the posterior distribution, $P(\theta)$ is the prior distribution and $P(D)$ is called the evidence, which is an integral of the likelihood over the prior distribution of the parameters:

$$P(D) = \int P(D \mid \theta)P(\theta)\ \mathrm{d}\theta. \tag{3.16}$$

We notice from the expression for the likelihood (3.14) that the likelihood is maximal when the error is minimal. The method of maximum likelihood is to simply find the set of the parameter values that maximize the likelihood function. We use a simple Metropolis algorithm to do a maximum likelihood estimation. The Metropolis algorithm uses an acceptance–rejection rule to converge to the required posterior distribution. The algorithm is as follows:

1. We start with some initial guess for the parameter values. It is done by drawing a starting point θ^0 for which $P(\theta^0 \mid D) > 0$ from a starting distribution $P_0(\theta)$ or we may choose some previous estimates for the parameter values if available.

2. For each step $t = 1, 2, 3, ...$

   (a) We propose a new set of parameter values by sampling a proposal θ^* from a Jumping distribution,$J(\theta^* \mid \theta^{t-1})$ at step t. The jumping distribution $J(\theta^* \mid \theta^{t-1})$ must be symmetric for this algorithm.

   (b) We generate a random uniform number between 0 and 1 and call it α. The likelihood ratio is calculated as
   $$r = \min\left\{\frac{P(\theta^* \mid D)}{P(\theta^{t-1} \mid D)}, 1\right\} \tag{3.17}$$

   (c) Set
   $$\theta^t = \begin{cases} \theta^* & \text{if } \alpha < r, \\ \theta^{t-1} & \text{otherwise.} \end{cases} \tag{3.18}$$

We run this algorithm for a long enough time and pick the parameter set that had the maximum likelihood. Generally, we just keep the parameter values after some burn in period. Since, there are no good theorem about the convergence rate of the MCMC, so it is difficult to decide how long we need to run. Generally, we run the MCMC, increase the time and run again and see how much answer converges. If parameters start to converge, we can be confident that we have run the simulation long enough. In principle, if we wait long enough MCMC will sample all of the guess space and find the global optimum. However, in practice, it can get stuck in local minima for a long time before getting out.

We will use Bayesian MCMC and clinical data from Thailand (Nisalak et al., 2003) to estimate the parameter values for our model (3.1) with the different hypotheses about dengue epidemiology. Indeed, we have begun implementing MCMC, so far only similar but simpler models. We are carefully considering whether the error function (3.14) is appropriate for our models.

3.3 Model Selection

It is possible to have several possible mathematical models that could explain the data available and we need to know that which of the model is best. The first thing which can be done is to see which model fits the data best. However, a more complex model with more parameters will likely fit the data better than a less complex model with fewer parameters, independent of whether it is actually a better model. For example we could always use a polynomial model with as many parameters as data points and it would fit the data perfectly but explain nothing. It is important to balance how good a model fits the data with the complexity of the model. Some of the popular methods to address this issue are *Akaike Information Criterion (AIC)* and *Bayes Information Criterion (BIC).*

The AIC selection procedure is based on the idea of finding the model that best approximates the true, but unknown, mechanisms generating the data. Models are ranked according to the AIC defined as

$$AIC(M) = -2\log(L) + 2K \tag{3.19}$$

where L is the maximized value of the likelihood of the model and K is the number of parameters. The best model, from a set of models that have been selected, is the one that minimizes AIC. The interpretation of the AIC is simple: the likelihood L increases as more terms are added to the model, which makes AIC smaller, but as the terms are added, K increases, which makes AIC larger. Thus AIC incorporates a trade-off between goodness-of-fit and parsimony of explanation.

The BIC selection procedure is to choose the candidate model with the highest probability given data. The BIC is closely related to the AIC and is defined as

$$BIC(M) = -2\log(L) + K\log(n) \tag{3.20}$$

where L is the maximized value of the likelihood of the model, K is the number of parameters and n is the number of data points, the number of observations, or equivalently, the sample size. The model with minimum BIC is selected. BIC is very similar to AIC, as both have a penalty term for the number of parameters in the model. However, this penalty is larger in the BIC than in the related AIC. In other words, complex models are more penalized by BIC than the simple ones as compared to AIC.

We will use AIC and BIC on the results from Bayesian MCMC on our models with the clinical data from Thailand (Nisalak et al., 2003) to select the most parsimonious model for dengue transmission. The results of this model selection could be a significant contribution to understanding dengue transmission, as shown by the competing models in the literature (e.g. Cummings et al., 2005; Adams et al., 2006; Wearing and Rohani, 2006; Nagao and Koelle, 2008).

Chapter 4

Optimization of Dengue Vaccine

One of the principal strategies for reducing the disease burden of infectious diseases is vaccination. As significant progress has been made in the development of a dengue vaccine recently (Guy et al., 2011), a licensed vaccine is expected to be available in near future. It is important to evaluate the vaccination policies before distribution of vaccine to allocate resources and to minimize disease burden. When vaccine availability is limited, it becomes imperative to allocate the vaccines optimally. Since second infections are likely more severe than primary ones, an efficient vaccine program would aim at vaccinating children old enough to have experienced a primary infection, although still young to not yet have experienced a secondary one. Once this group is protected, the vaccine program would then shift its focus to younger children, and possibly to older people who may have a small fraction who may not yet have had a secondary infection or who may be at risk of third or fourth infections. Thus in order to allocate vaccine optimally, it is essential to choose target age groups for the vaccine that most efficiently reduce the disease burden.

We will select a basic model that most parsimoniously fits the clinical data from Thailand (Nisalak et al., 2003) using the AIC and BIC on the results from Bayesian MCMC on our models from chapter 3. The model is then extended to an age-structured model and then we will use numerical optimization routine to identify the target age groups for the allocation of vaccine.

4.1 Age-Structured Model

To determine the optimal age-specific distribution of different numbers of available vaccine, we must incorporate age into the model. This is done by adding discete age groups in the model, where the first age group is made up of newborns. As time progresses individuals move up to subsequent age groups or die. We formulate an age-structured dengue model with $n+1$ age groups with vaccination by extending our dengue model (3.1). We define $S_{a,\mathcal{J}}$ to be the class of uninfected individuals in age class a who have previously been exposed to strains in set $\mathcal{J}$. The class of infectious individuals $I_{a,i,\mathcal{J}}$, and class of temporarily immune individuals $T_{a,\mathcal{J}}$, are defined in a similar fashion as $S_{a,\mathcal{J}}$. An additional class R_a of recovered individuals is added to keep track of vaccinated people who are immune to all the serotypes. The extended dengue model is given by

$$\frac{\mathrm{d}S_{0,\emptyset}}{\mathrm{d}t} = \sum_{\alpha=0}^{n} B_\alpha N_\alpha - A_0 S_{0,\emptyset} - \sum_{i\in\mathcal{K}} \lambda_0^i(t) S_{0,\emptyset} - \mu_0 S_{0,\emptyset} - V_0 S_{0,\emptyset}, \tag{4.1a}$$

$$\frac{\mathrm{d}S_{a,\emptyset}}{\mathrm{d}t} = A_{a-1} S_{a-1,\emptyset} - A_a S_{a,\emptyset} - \sum_{i\in\mathcal{K}} \lambda_a^i(t) S_{a,\emptyset} - \mu_a S_{a,\emptyset} - V_a S_{a,\emptyset}, \tag{4.1b}$$

$$\frac{\mathrm{d}S_{a,\mathcal{J}}}{\mathrm{d}t} = A_{a-1} S_{a-1,\mathcal{J}} - A_a S_{a,\mathcal{J}} + \delta T_{a,\mathcal{J}} - \sum_{i\notin\mathcal{J}} \sigma_{\mathcal{J}}^i \lambda_a^i(t) S_{a,\mathcal{J}} - \mu_a S_{a,\mathcal{J}} - V_a S_{a,\mathcal{J}}, \tag{4.1c}$$

$$\frac{\mathrm{d}I_{a,i,\mathcal{J}}}{\mathrm{d}t} = A_{a-1} I_{a-1,i,\mathcal{J}} - A_a I_{a,\mathcal{J}} + \sigma_{\mathcal{J}}^i \lambda_a^i(t) S_{a,\mathcal{J}} - (\mu_a + \gamma_{\mathcal{J}}^i + \nu_{\mathcal{J}}^i) I_{a,i,\mathcal{J}}, \tag{4.1d}$$

$$\begin{aligned}\frac{\mathrm{d}T_{a,\mathcal{J}}}{\mathrm{d}t} &= A_{a-1} T_{a,\mathcal{J}} - A_a T_{a,\mathcal{J}} + \sum_{i\in\mathcal{J}} \gamma_{\mathcal{J}}^i I_{a,i,\mathcal{J}\setminus i} - (\mu+\delta) T_{a,\mathcal{J}} \\ &\quad + \chi \left(\sum_{i\in\mathcal{J}} \lambda_a^i(t) T_{a,\mathcal{J}\setminus i} - \sum_{i\notin\mathcal{J}} \lambda_a^i(t) T_{a,\mathcal{J}} \right),\end{aligned} \tag{4.1e}$$

$$\frac{\mathrm{d}R_a}{\mathrm{d}t} = \sum_{a=0}^{n} \sum_{\mathcal{J}} V_a S_{a,\mathcal{J}} + A_{a-1} R_{a-1} - A_a R_a - \mu_a R_a, \tag{4.1f}$$

where the force of infection for strain i with a small level of seasonal forcing ϵ is

$$\lambda_a^i(t) = [1 + \epsilon \cos(2\pi t)] \left(\sum_{\mathcal{J}\subset\mathcal{K}\setminus i} \beta_{\mathcal{J}}^i \sum_{\alpha=0}^{n} \frac{I_{a,i,\mathcal{J}}}{N_\alpha} \right), \tag{4.1g}$$

B_α is the birth rate in age group α, N_α is the number of individuals in age group α, V_a is the vaccination rate for age group a and A_a is the aging constant for age group a and is calculated as reciprocal of the width of age group, but with $A_n = 0$, so that the last age group accumlates all people above a certain age.

4.2 Objective Functions

We will determine the optimal age-specific distribution of a limited number of vaccine doses according to few different outcome measures (objective functions). The different objective functions we will use are total number of deaths averted, total number of infections averted, years of life saved and disability-adjusted life years saved. Years of life lost (YLL) takes into account the age at which death occurs by giving greater weight to deaths at younger age and lower weight to deaths at older age. Disability-adjusted life years (DALYs) is a health gap measure to determine years of 'healthy' life lost by virtue of being in state of poor health or disability. The objective functions can be calculated from the age-structured model (4.1) as follows:

1. Total number of deaths from infection:

$$\int_0^T \sum_i \sum_{\mathcal{J}} \sum_a \nu^i_{\mathcal{J}} I_{a,i,\mathcal{J}} \mathrm{d}t. \tag{4.2}$$

2. Total number of infections:

$$\int_0^T \sum_i \sum_{\mathcal{J}} \sum_a \sigma^i_{\mathcal{J}} \lambda^i_a S_{a,\mathcal{J}} \mathrm{d}t. \tag{4.3}$$

3. Years of Life Lost (YLL):

$$\int_0^T \sum_a Y_a \sum_{\mathcal{J}} \sum_i \nu^i_{\mathcal{J}} I_{a,\mathcal{J}} \mathrm{d}t, \tag{4.4}$$

   where Y_a is the measure of years of life lost for an individual in age group a,

4. Disability-adjusted lost years (DALYs):

$$\int_0^T \sum_i \sum_{\mathcal{J}} \sum_a D^i_{\mathcal{J}} \sigma^i_{\mathcal{J}} \lambda^i_a S_{a,\mathcal{J}} \mathrm{d}t, \tag{4.5}$$

   where $D^i_{\mathcal{J}}$ is the measure of DALYs for an individual who is currently infected with serotype i and has previously been exposed to strains in set $\mathcal{J}$. In particular, $D^i_{\mathcal{J}}$ for $|J| > 0$ will be large relative to $D^i_{\emptyset}$.

4.3 Vaccine Optimization

The maximum total number of vaccinations per unit time imposes a constraint on the optimal vaccine allocation problem. The sum of vaccination rates per unit time for each age group is bounded by the total number of vaccine available per unit time and the constraint along with the non-negativity of each group is given by

$$\sum_a V_a S_a \leq V_T, \tag{4.6}$$

where $V_a \geq 0$ and V_T is the total number of available vaccine per unit time for the population.

Now, given the age-structured model and an objective function, our aim is to find a vaccine allocation for all the age groups in the model such that it satisfies the constraint (4.6) and minimizes (optimizes) the objective function. To do this we will use numerical optimization algorithms, particularly we will consider algorithms based on quasi-Newton methods and the Nelder–Mead simplex method. The results will then be compared for different objective functions to better understand the basic tradeoffs involved in different optimal allocations as Medlock and Galvani (2009) did for influenza. Incorporating the demographics of the population due to the larger timescale of dengue epidemics and the interaction of multiple serotypes might lead to substantial complications over this previous work.

Bibliography

B. Adams, E.C. Holmes, C. Zhang, M.P. Mammen Jr., S. Nimmannitya, S. Kalayanarooj, and M. Boots. Cross-protective immunity can account for the alternating epidemic pattern of dengue virus serotypes circulating in Bangkok. *Proc Natl Acad Sci USA*, 103(38):14234–14239, 2006.

R.M. Anderson and R.M. May. Infectious diseases of humans: dynamics and control. *New York*, 757, 1991.

Derek A. T. Cummings, Ira B. Schwartz, Lora Billings, Leah B. Shaw, and Donald S. Burke. Dynamic effects of antibody-dependent enhancement on the fitness of viruses. *Proc Natl Acad Sci USA*, 102(42):15259–15264, 2005.

L. Esteva and H. Mo Yang. Mathematical model to assess the control of Aedes aegypti mosquitoes by the sterile insect technique. *Mathematical biosciences*, 198(2):132–147, 2005. ISSN 0025-5564.

L. Esteva and C. Vargas. Influence of vertical and mechanical transmission on the dynamics of dengue disease. *Mathematical Biosciences*, 167(1):51–64, 2000. ISSN 0025-5564.

N. Ferguson, R. Anderson, and S. Gupta. The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, 96(2):790, 1999.

DA Focks, DG Haile, E. Daniels, and GA Mount. Dynamic life table model for Aedes aegypti (Diptera: Culicidae): analysis of the literature and model development. *Journal of Medical Entomology*, 30(6):1003–1017, 1993a. ISSN 0022-2585.

DA Focks, DG Haile, E. Daniels, and GA Mount. Dynamic life table model for Aedes aegypti (Diptera: Culicidae): simulation results and validation. *Journal of Medical Entomology*, 30(6): 1018–1028, 1993b. ISSN 0022-2585.

D.A. Focks, E. Daniels, D.G. Haile, and J.E. Keesling. A simulation model of the epidemiology of urban dengue fever: literature analysis, model development, preliminary validation, and samples of simulation results. *The American journal of tropical medicine and hygiene*, 53(5):489, 1995.

D.A. Focks, R.J. Brenner, J. Hayes, and E. Daniels. Transmission thresholds for dengue in terms of Aedes aegypti pupae per person with discussion of their utility in source reduction efforts. *The American journal of tropical medicine and hygiene*, 62(1):11, 2000.

Duane J. Gubler. Dengue and dengue hemorrhagic fever. *Clin Microbiol Rev*, 11(3):480–496, 1998a. ISSN 0893-8512.

Duane J. Gubler. Resurgent vector-borne diseases as a global health problem. *Emerg Infect Dis*, 4 (3):442–450, 1998b.

B. Guy, J. Almond, and J. Lang. Dengue vaccine prospects: a step forward. *The Lancet*, 377(9763): 381–382, 2011. ISSN 0140-6736.

Maria G. Guzmán and Gustavo Kouri. Dengue: an update. *Lancet Infect Dis*, 2(1):33–42, 2002. ISSN 1473-3099.

María G. Guzmán, Gustavo Kouri, Jose Bravo, Luis Valdes, Susana Vasquez, and Scott B. Halstead. Effect of age on outcome of secondary dengue 2 infections. *International Journal of Infectious Diseases*, 6(2):118–124, 2002.

M.G. Guzman, S.B. Halstead, H. Artsob, P. Buchy, J. Farrar, D.J. Gubler, E. Hunsperger, A. Kroeger, H.S. Margolis, E. Martínez, et al. Dengue: a continuing global threat. *Nature Reviews Microbiology*, 8:S7–S16, 2010. ISSN 1740-1526.

Scott B. Halstead. Dengue. *Lancet*, 370:1644–1652, 2007.

Isao Kawaguchi, Akira Sasaki, and Michael Boots. Why are dengue virus serotypes so distantly related? enhancement and limiting serotype similarity between dengue virus strains. *Proc Biol Sci*, 270(1530):2241–2247, 2003.

Jennifer L. Kyle and Eva Harris. Global spread and persistence of dengue. *Annu Rev Microbiol*, 62: 71–92, 2008. doi: 10.1146/annurev.micro.62.081307.163005.

Paula Mendes Luz, Cláudia Toress Codeço, Eduardo Massad, and Claudio José Struchiner. Uncertainties regarding dengue modeling in Rio de Janeiro, Brazil. *Mem Inst Oswaldo Cruz*, 98(7): 871–878, 2003.

J. Medlock and A.P. Galvani. Optimizing influenza vaccine distribution. *Science*, 325(5948):1705, 2009. ISSN 0036-8075.

Yoshiro Nagao and Katia Koelle. Decreases in dengue transmission may act to increase the incidence of dengue hemorrhagic fever. *Proc Natl Acad Sci USA*, 105(6):2238–2243, 2008.

Ananda Nisalak, Timothy P. Endy, Suchitra Nimmannitya, Siripen Kalayanarooj, Usa Thisayakorn, Robert M. Scott, Donald S. Burke, Charles H. Hoke, Bruce L. Innis, and David W. Vaughn. Serotype-specific dengue virus circulation and dengue disease in Bangkok, Thailand from 1973 to 1999. *Am J Trop Med Hyg*, 68(2):191–202, 2003. ISSN 0002-9637.

H. Nishiura. Mathematical and statistical analyses of the spread of dengue. *Dengue Bulletin*, 30:51, 2006. ISSN 1020-895X.

World Health Organization, Special Programme for Research, Training in Tropical Diseases, World Health Organization. Dept. of Control of Neglected Tropical Diseases, World Health Organization. Epidemic, Pandemic Alert, and Response. *Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control*. World Health Organization, 2009. ISBN 9241547871.

R. Rico-Hesse. Microevolution and virulence of dengue viruses. *Adv Virus Res*, 59:315–341, 2003.

Albert B. Sabin. Research on dengue during World War II. *Am J Trop Med Hyg*, 1(1):30–50, 1952.

D.S. Shepard, J.A. Suaya, S.B. Halstead, M.B. Nathan, D.J. Gubler, R.T. Mahoney, D.N.C. Wang, and M.I. Meltzer. Cost-effectiveness of a pediatric dengue vaccine. *Vaccine*, 22(9-10):1275–1280, 2004. ISSN 0264-410X.

J. F. Siler, Milton W. Hall, and A. Parker Hitchens. Dengue: its history, epidemiology, mechanism of transmission, etiology, clinical manifestations, immunity, and prevention. *Philipp J Sci*, 29: 1–304, 1926.

Helen J. Wearing and Pejman Rohani. Ecological and immunological determinants of dengue epidemics. *Proc Natl Acad Sci USA*, 103(31):11802–11807, 2006.