PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Transmission dynamics and epidemiology of dengue: insights from age–stratified sero–prevalence surveys

Neil M. Ferguson, Christl A. Donnelly and Roy M. Anderson

Phil. Trans. R. Soc. Lond. B 1999 **354**, 757-768
doi: 10.1098/rstb.1999.0428

THE ROYAL SOCIETY

Transmission dynamics and epidemiology of dengue: insights from age-stratified sero-prevalence surveys

Neil M. Ferguson*, Christl A. Donnelly and Roy M. Anderson

Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

The relationship between infection with the four major serotypes of dengue virus and the occurrence of different forms of disease is complex and not fully understood. Interpreting longitudinal records of the incidence of serious disease to gain insight into the transmission dynamics and epidemiology of the virus is therefore complicated. Since age reflects duration of exposure, age-stratified, strain-specific serological surveys carried out at one point in time, or over a short time interval, can potentially provide a rich source of information on longitudinal patterns. This paper describes the development and application (to data collected in Thailand) of statistically rigorous methods designed to estimate time-varying, strain-specific forces of infection, and thus basic reproduction numbers, from cross-sectional serological data. The analyses provide support for the hypothesis that antibody-dependent enhancement of transmission influences observed epidemiological pattern. Age-stratified serological data also reveal evidence of a propensity for the annual incidence of infection to oscillate over time with a frequency of several years. The latter observation is consistent with the predictions of simple mathematical models of the transmission dynamics of the virus. The estimates of the basic reproduction numbers obtained are similar in magnitude for each dengue serotype, being in the range of four to six. Such values are higher than those obtained from earlier analyses, and the implications of this for dengue control are discussed.

Keywords: epidemiology; model; dengue; sero-prevalence; antibody-dependent enhancement; R_0

1. INTRODUCTION

Dengue virus infection is an important public health problem in many tropical regions of the world (Halstead 1993; Gubler & Kuno 1997; Rigau-Perez *et al*. 1998) due to the high morbidity and mortality associated with the severe forms of the disease (Halstead 1998; WHO 1986). For example, dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS) are among the leading causes of childhood hospitalizations and death in South-east Asia (Gubler 1988). The dengue virus belongs to the flaviviruses, and all four antigenically related but distinct serotypes (DEN-1 to DEN-4) are thought able to cause serious disease. In recent years the average annual incidence of dengue-related serious disease in many tropical counties has been rising dramatically, with the infection becoming endemic in cities where its occurrence was once sporadic (Gubler 1998). The reasons for this are complex and not fully understood. A variety of factors are likely to be important, including human population growth and increased aggregation in urban areas. Also of importance is an increased frequency of travel between cities and countries, and failures in public health measures such as mosquito vector control in urban slum areas in affected regions. Recent molecular epidemiological studies also reveal rapid recent evolution of the virus, where the rate is closely correlated in time with human population growth (Page & Holmes 1998). Such work also provides evidence of recombination between viral strains, which suggests that new serotypes may emerge in the future as the rate of viral evolution accelerates within large urban conglomerates.

All these factors argue for an enhanced research effort on the epidemiology and control of this important viral infection. For effective control, whether by the use of insecticides or via the use of a vaccine (none are currently available but products are in development), a better understanding of the transmission dynamics of the virus is key. For example, in the use of insecticides or vaccines, it is essential to understand the oscillatory dynamics of the virus, both seasonal and longer term, the relationship between age at infection and the incidence of serious disease, and the role of multiple infections in the occurrence of serious disease. Serological surveys provide opportunities to estimate the age- or time-specific force of infection and, concomitantly, the basic or case reproduction number R_0 (the average number of secondary cases of infection generated by one primary case in a susceptible population (Anderson & May 1991)). This in turn provides insights into how a defined degree of vector control or vaccination will influence the incidence of infection and disease, and the level of control required to block transmission. Preliminary studies have been carried out on the estimation of R_0 (Dietz 1975; Koopman *et al*. 1991; Marques *et al*. 1994; Newton & Reiter 1992), but a

*Author to whom correspondence should be addressed.

crucial step is to base such analyses on strain-stratified data.

The principal aim of this paper, therefore, is to describe the development and application of statistical methods to estimate strain-specific forces of infection over time, and strain-specific values for the basic reproduction number. These methods are applied to the analysis of age- and strain-stratified serological data collected in Thailand (Sangkawibha *et al*. 1984). A further aim is the examination of age-stratified data for further evidence of periodicity in the incidence of infection. Time-series analysis of longitudinal data on the incidence of serious disease suggests three to four year oscillations in incidence (superimposed on a background of seasonal cycles) (S. Hay and D. Rodgers, personal communication), as predicted by theoretical models of viral transmission dynamics (Ferguson *et al*. 1999).

The occurrence of the virus as four distinct serotypes raises many complications in the analysis and interpretation of serological data. Antibodies generated by exposure to any one strain are known to be cross-reactive for other strains, but they are believed only to provide strain-specific lifelong immunity to reinfection (Kuno *et al*. 1993). The immunological response on exposure to a second strain is complex and depends on factors such as patient age, strain type and the interval between exposure to one serotype and exposure to the second serotype (Halstead *et al*. 1973; Halstead & O'Rourke 1977; Halstead 1979). The high antibody titres attained after primary infection appear to generate a degree of cross-protection for a while, but if secondary exposure occurs after antibody levels begin to decline, cross-reactivity appears to act to enhance the growth rate of the new invading viral strain. This is called antibody-dependent enhancement (ADE) and its occurrence in dengue has been used to explain the aetiology of serious disease (DHF and DSS) (Halstead 1982, 1988; Kliks *et al*. 1988; Kliks 1990; Thein *et al*. 1997). In general ADE is the phenomenon by which cross-reactive antibodies generated by prior exposure to a heterologous strain enhance replication of a second invading strain. *In vitro* studies have demonstrated such replication enhancement for a variety of viruses, including dengue (Halstead *et al*. 1977; Mady *et al*. 1991), other flaviviruses (Peiris *et al*. 1981; Porterfield 1986), and HIV (Robinson *et al*. 1988; Takeda *et al*. 1988).

Cross-reactive immunological responses complicate the interpretation of age-stratified, strain-specific seroprevalence surveys. Secondary dengue infection in an individual with monotypic immunity derived from infection by a different strain elicits antibody responses that are broadly cross-reactive with all four serotypes (Okuno *et al*. 1980; Kuno *et al*. 1993). Current serological tests classify people as either having been exposed once to an identifiable serotype or exposed more than once to an unidentifiable set of dengue serotypes. This renders standard methods of estimating strain-specific forces of infection ineffective, as information on exposure to particular serotypes is lost once a second exposure to a different serotype has occurred. New immunological approaches, based on viral serotype genome studies, may resolve this complication in the future by identifying regions of the genomes that encode for serotype-specific antigens where antibodies to these strain-specific antigens remain detectable in multi-typically exposed hosts (Gritsun *et al*. 1995).

We describe new methods for estimating time-varying, strain-specific forces of infection based on cross-sectional serological survey data which record only monotypic and multitypic past infections. The basic reproductive number of each strain is then estimated under a variety of different assumptions regarding the degree of cross-protection and/or enhancement conferred by primary or later infections. These methods are applied to analyse serological survey data collected in Thailand (Sangkawibha *et al*. 1984) to obtain estimates of the force of infection and basic reproduction numbers. We also examine the resulting evidence for non-seasonal periodicity in the incidence of infection.

2. INTERPRETING SERO-PREVALENCE PROFILES

(a) *Simple model*

A sero-prevalence profile derived from a cross-sectional survey provides age-stratified estimates of the proportion of individuals that have been previously infected by some pathogen (assuming lifelong persistence of antigen-specific immunological markers (e.g. antibodies)). Usually the proportion seropositive increases with age, in which case the rate of increase with age can be interpreted as a measure of the 'intensity' of infection experienced by the population in the past. Specifically, a large increase in seroprevalence between age a and $a+1$ could result from a high risk of infection which was specific to that age window, but constant through time, or from an age-independent risk that was specific to the time period between a and $a+1$ years ago, or from any combination of the two. Hence, whilst a seroprevalence profile provides information on the overall cumulative infection risk experienced by individuals, age- and time-specific effects are invariably confounded in the absence of other data (Becker 1989).

We build on the existing methodological framework developed to estimate time- or age-dependent forces of infection from serological data (Grenfell & Anderson 1985; Becker 1989; Keiding 1991), but extend it to treat infectious agents with multiple strains that interact via the host immune response. We therefore only briefly review the basic approach (for applications, see, for example, Ades & Nokes 1993; Anderson & May 1991) before discussing the modifications required to analyse multi-strain systems.

For a single-strain directly transmitted infection, the proportion that are seropositive at age a and time t, $z(a,t)$ is related to the force (or per capita rate) of infection, $\lambda(a,t)$ thus:

$$\frac{\partial z}{\partial a}+\frac{\partial z}{\partial t}=(1-z(a,t))\lambda(a,t). \tag{1}$$

This has the solution

$$z(a,t)=\int_0^a \lambda(a-\tau,t-\tau)\mathrm{e}^{-\int_0^\tau \lambda(a-\tau',t-\tau')\mathrm{d}\tau'}\mathrm{d}\tau \tag{2}$$

$$=1-\mathrm{e}^{-\int_0^a \lambda(a-\tau,t-\tau)\mathrm{d}a'}. \tag{3}$$

Equation (2) represents an integral over all possible times of infection, $t-\tau$, where the first term, $\lambda(a-\tau,t-\tau)$, is

just the probability of infection per unit time at $t-\tau$ conditional on not already being infected (termed the infection hazard), and the second term is the probability of having escaped infection up to that time. As seen in equation (3), we can express this equivalently as one minus the probability of having escaped infection up until time t. However, we shall see that this is not always possible for complex multi-strain systems.

It should be noted that a cross-sectional sero-prevalence survey measures $z(a,t)$ for a range of values of a, but for fixed t. The slope of the sero-prevalence curve is therefore not given by equation (1), but by

$$\frac{\partial z}{\partial a} = (1-z)\left[\lambda(0,t-a) + \int_0^a \frac{\partial \lambda}{\partial a}(a-\tau,t-\tau)\mathrm{d}\tau\right]. \quad (4)$$

This illustrates clearly how age- and time-dependence in $\lambda(a,t)$ are intrinsically confounded in sero-prevalence measurements. Equation (4) is only simple in structure if $\lambda(a,t)$ is independent of either age or time:

$$\frac{\partial z}{\partial a} = (1-z)\lambda(0,t-a) \quad \text{if } \lambda(a,t)=\lambda(0,t) \quad \forall t,a \quad (5)$$

$$= (1-z)\lambda(a,0) \quad \text{if } \lambda(a,t)=\lambda(a,0) \quad \forall t,a. \quad (6)$$

For infections (such as measles) where independent case notification data gives information on temporal changes in $\lambda(a,t)$, these equations can be used to estimate the age-specific trends in the force of infection (Griffiths 1974; Grenfell & Anderson 1985). In turn these estimates can give insight into age-dependent mixing patterns in the host populations (Dietz & Schenzle 1985; Anderson & May 1985). In the case of dengue, however, no such data exist. Thus, assumptions regarding age-specific patterns of infection are required before serological data collected at a single point in time can be used to estimate temporal trends (or vice versa). For simplicity we assume that age-specific variation in mosquito biting rates, and thus in the force of infection, is not of significance. However, unlike much earlier work, we do not assume that the force of infection is also constant through time.

(b) *Multiple strain infections*

For a multi-strain system, one would ideally want to be able to identify independently, through serological testing, all strains with which an individual had been infected. As mentioned above, this is not possible for dengue, due to the cross-reactivity of the antibody response. In practice, only monotypic prevalences for each strain and the prevalence of all multitypic strain combinations can be measured. In the absence of any cross-protective or ADE response following primary infection, the force of infection $\lambda_i(a,t)$ for each strain may be calculated from the proportions, $z_i(a,t)$, who have been exposed solely to strain i, and the proportion, $x(a,t)$, who remain completely susceptible (never infected with any strain):

$$\frac{\partial x}{\partial a} + \frac{\partial x}{\partial t} = -x\lambda(a,t-a), \quad (7)$$

$$\frac{\partial z_i}{\partial a} + \frac{\partial z_i}{\partial t} = x\lambda(a,t-a) - \sum_{k\neq i}\lambda_k(a,t). \quad (8)$$

Again, these equations can be written equivalently in terms of integrals over possible infection times and/or ages:

$$x(a,t) = \mathrm{e}^{-\int_0^a \sum_k \lambda_k(a-\tau,t-\tau)\mathrm{d}\tau}, \quad (9)$$

$$z_i(a,t) = \left[\mathrm{e}^{-\int_0^a \sum_{k\neq i}\lambda_k(a-\tau,t-\tau)\mathrm{d}\tau}\right]\left[1-\mathrm{e}^{-\int_0^\infty \lambda_i(a-\tau,t-\tau)\mathrm{d}\tau}\right], \quad (10)$$

$$= x(a,t)\left[\mathrm{e}^{\int_0^a \lambda_i(a-\tau,t-\tau)\mathrm{d}\tau} - 1\right]. \quad (11)$$

The expression for $x(a,t)$ is just the probability of escaping infection with any strain up to time t, whilst the expression for $z_i(a,t)$ is the probability of being infected with strain i but escaping infection with any other strain. The proportion multitypically infected is just $1-x(a,t)-\sum_i z_i(a,t)$.

So far we have assumed strain independence: that the probability of being infected with one strain, and infectiousness following infection with that strain, are independent of past exposure to other strains. ADE or cross-immunity may invalidate this assumption, and this may have an effect on the observed sero-prevalence profiles. Specifically, if exposure to one strain either increases or decreases susceptibility to a second strain, then the above model is invalid. However, before describing a more suitable model for this case, it is worth emphasizing exactly what is meant in this context by enhancement (or reduction) of susceptibility. It corresponds to a phenomenon in which the probability that strain-specific antibodies (and hence permanent immunity) are generated following exposure to a strain depends on the past history of exposure to heterologous strains. This should be carefully distinguished from enhancement of infectiousness, in which the probability of generating a strain-specific immune response is identical for primary or secondary infections, but the degree of viraemia may differ, potentially leading to different levels of infectiousness. The latter phenomenon might raise the overall force of infection for a strain compared with that expected otherwise, but on its own does not change the relative infection hazards experienced by completely susceptible and monotypically exposed individuals. Hence serology is unable to detect such enhancement–inhibition, and the model above remains applicable for estimating forces of infection.

We can, however, examine serological data for the signature of susceptibility enhancement–inhibition. The basic partial differential equation model remains little changed:

$$\frac{\partial x}{\partial a} + \frac{\partial x}{\partial t} = -x\lambda(a,t-a), \quad (12)$$

$$\frac{\partial z_i}{\partial a} + \frac{\partial z_i}{\partial t} = x\lambda(a,t-a) - \sum_{k\neq i}\rho_{ik}\lambda_k(a,t), \quad (13)$$

where ρ_{ij} represents the enhancement ($\rho_{ij}>1$) or inhibition ($\rho_{ij}<1$) of susceptibility to infection with strain j seen following infection with strain i. (Note that $\rho_{ii}=0$, representing permanent strain-specific immunity.) However, the solution of these equations is considerably more complex. Whilst the expression for $x(a,t)$ remains unchanged, the probability of only being infected with

strain i by age a and time t, $z_i(a,t)$, now must take account of the different risk of infection by heterologous strains experienced before and after infection with strain i. It can be shown that

$$z_i(a,t) = x(a,t)\int_0^a \lambda_i(a-\tau,t-\tau)\mathrm{e}^{\int_0^\tau \sum_k (\rho_{ik}-1)\lambda_k(a-\tau',t-\tau')\mathrm{d}\tau'}\mathrm{d}\tau. \tag{14}$$

3. THE BASIC REPRODUCTION NUMBER

R_0, the basic reproduction number of an infectious agent, is defined to be the number of secondary infections caused by one primary infection in an entirely susceptible population (MacDonald 1957, 1965; Anderson & May 1991). It is key in determining whether a pathogen can invade a host population (which requires $R_0 > 1$), and the extent and rate of spread of a pathogen into that population. As such it plays a critical role in determining the intensity of control measures required to eradicate an endemic disease. For a simple directly transmitted pathogen in a homogeneously mixing population, the minimum proportion of the population required to be effectively vaccinated for disease elimination is given by $p_c = 1 - 1/R_0$ (Anderson & May 1991).

Here we address the question of how R_0 can be estimated from serological data. Since R_0 is a measure of transmission potential, we need to extend the formalism developed above to model transmission. In other words, instead of estimating the force of infection $\lambda(a,t)$ as an independent parameter, we represent the time evolution of $\lambda(a,t)$ as a function of the transmission dynamics of the pathogen in the host population.

Beginning with the simple model of a single-strain infection conferring permanent immunity, we add an infectious class of individuals, defining $y(a,t)$ to be the proportion of individuals of age a at time t which are infectious. We consider this group to overlap with the exposed class $z(a,t)$, so that individuals move into both on infection, but leave the infectious class at some higher rate determined by the duration of infectiousness. Thus,

$$\frac{\partial y}{\partial a} + \frac{\partial y}{\partial t} = (1 - z(a,t))\lambda(a,t) - \sigma y, \tag{15}$$

where σ is the rate of loss of infectiousness (1/duration of infectiousness). The transmission model is then completed by relating the force of infection to the proportion of the population infectious. The simplest approach is to assume homogenous (mass-action) mixing, where the force of infection is just proportional to the total proportion of the population infectious, namely

$$\lambda(t) = \int_0^\infty \beta f(a')\, y(a',t)\mathrm{d}a', \tag{16}$$

where $f(a)$ is the proportion of the population of age a (we neglect disease-induced mortality and time-varying demography here) and β is the transmission coefficient (assumed constant through time and across age classes), representing the probability per unit time of an infectious individual infecting a susceptible. In the case of vector-borne diseases, we can generalize the model to include the effect of age-dependent biting rates and/or susceptibility by allowing β to vary with age.

R_0 is just the probability per capita per unit time of infection, multiplied by the average duration of infectiousness, namely $R_0 = \beta/\sigma$, if the natural mortality rate, $\mu(a)$, is always much less than σ. Assuming age-independent $\lambda(t)$, to estimate R_0 we differentiate equation (16) with respect to time, and substitute equation (15) into the resulting expression, to obtain

$$\begin{aligned}\frac{\mathrm{d}\lambda}{\mathrm{d}t} =& \beta\lambda\left(1 - \int_0^\infty f(a')z(a',t)\mathrm{d}a'\right) \\ &- \sigma\lambda - \beta\int_0^\infty \mu(a')f(a')\,y(a',t)\mathrm{d}a',\end{aligned} \tag{17}$$

where the last term represents the effect of natural mortality (note $\partial f/\partial a = -\mu(a)f(a)$). This term can be neglected if $\mu(a) \ll \sigma$, giving

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t} = \beta\lambda\left(1 - \int_0^\infty f(a')z(a',t)\mathrm{d}a'\right) - \sigma\lambda. \tag{18}$$

Since $R_0 = \beta/\sigma$, this gives

$$R_0 = \frac{(\mathrm{d}\lambda(t)/\mathrm{d}t)/(\sigma\lambda(t)) + 1}{1 - \int_0^\infty f(a')z(a',t)\mathrm{d}a'}. \tag{19}$$

For an infection at endemic equilibrium ($\mathrm{d}\lambda/\mathrm{d}t = 0$), this reduces to the result that R_0 is equal to the inverse of the proportion of the population remaining susceptible to the pathogen at endemic equilibrium (Anderson & May 1991). However, equation (19) has more general application to situations when β is thought to be unvarying with time (at least between years, i.e. neglecting seasonal fluctuations), but the force of infection is changing with time, either because the disease has not yet reached equilibrium or because the transmission dynamics are intrinsically oscillatory.

For multi-strain pathogens where the strains do not interact via the host immune response, the above result translates into an expression for the strain-specific basic reproduction number, R_{0i}:

$$R_{0i} = \frac{(\mathrm{d}\lambda_i(t)/\mathrm{d}t)/(\sigma\lambda_i(t)) + 1}{1 - \int_0^\infty f(a')w_i(a',t)\mathrm{d}a'}, \tag{20}$$

where $w_i(a,t)$ is the proportion of individuals of age a at time t who have ever been exposed to strain i. In the case of dengue, w_i is unmeasurable, due to the difficulty in distinguishing individual strains in multitypically infected individuals by current serological methods, so it is necessary to estimate w_i from the estimated $\lambda_i(t)$, thus

$$w_i(a,t) = 1 - \mathrm{e}^{-\int_0^a \lambda_i(a-\tau,t-\tau)\mathrm{d}\tau}. \tag{21}$$

As has been noted in relation to malaria (Gupta *et al.* 1994), using non-strain-specific data and equation (19) to estimate R_0 for dengue would result in an estimate that was the product of the strain-specific R_{0i}s calculated above (in the absence of strain interactions). If such estimates were used in intervention strategy design, the difficulty of controlling infection incidence might severely be overestimated.

For interacting strains, we need to distinguish between infectious individuals with primary and secondary infections. Denoting the proportion infectious following primary infection with strain i by $y_i(a,t)$, and the proportion

infectious following secondary infection with strain i after prior exposure to strain j by $y_{ji}(a,t)$, we can write down equations for their dynamics analogous to equation (15) using equations (12) and (13):

$$\frac{\partial y_i}{\partial a}+\frac{\partial y_i}{\partial t}=x(a,t)\lambda(a,t)-\sigma y_i, \tag{22}$$

$$\frac{\partial y_{ji}}{\partial a}+\frac{\partial y_{ji}}{\partial t}=\rho_{ji}z_j\lambda_i(a,t)-\sigma y_{ji}, \tag{23}$$

where the second equation allows for different susceptibilities of hosts to secondary infection depending on the strain causing primary infection (ρ_{ji}). Note that we implicitly assume here that exposure to two strains generates permanent immunity to all others.

Assuming homogenous mass-action mixing, λ_i is given by

$$\lambda_i(t)=\beta\int_0^\infty f(a')\left[y_i(a',t)+\sum_j \phi_{ji}y_{ji}(a')\right]\mathrm{d}a', \tag{24}$$

where ϕ_{ji} represents the degree to which infectiousness is enhanced or inhibited by prior exposure to strain j (with $\phi_{ii}=0$).

As in the single-strain case, the equations for the infectious proportions can be eliminated, giving one equation for the time-evolution of λ_i:

$$\frac{\mathrm{d}\lambda_i}{\mathrm{d}t}=\beta\lambda\int_0^\infty f(a')\left[x(a',t)+\sum_j \phi_{ji}\rho_{ji}z_j(a',t)\right]\mathrm{d}a'-\sigma\lambda, \tag{25}$$

where we have again assumed $\mu(a)\ll\sigma$. Hence R_{0i} is given by:

$$R_{0i}=\frac{(\mathrm{d}\lambda_i(t)/\mathrm{d}t)/(\sigma\lambda_i(t))+1}{\int_0^\infty f(a')\left[x(a',t)+\sum_j \phi_{ji}\rho_{ji}z_j(a',t)\right]\mathrm{d}a'}. \tag{26}$$

An interesting point to note here is that whilst we are able to estimate the degree of susceptibility enhancement–inhibition (ρ_{ij}) from serological data, it is impossible to estimate the degree of infectiousness enhancement–inhibition, ϕ_{ij}. Generally, therefore, if we assume $\phi_{ij}=1$, then we will overestimate R_0 if enhancement is a significant effect, and underestimate it if inhibition dominates. Furthermore, our assumption that exposure to two strains generates immunity means that equation (26) may overestimate R_{0i}. The only way in which R_{0i} can therefore be precisely estimated using serological data is from a population in which only strain i is circulating.

We can use equation (26) to get an upper bound on R_{0i} by assuming complete cross-protection between all strains ($\rho_{ij}=0$ or $\phi_{ij}=0$). A lower bound is, however, impossible to estimate, since if ADE were important enough epidemiologically, it might be possible for one or more of the R_{0i} to be below one, with strains having evolved to exist 'co-operatively' with each other. In this context, the concept of R_0 as defined in relation to an entirely susceptible population breaks down, and it is necessary to consider the criteria for co-existence of two or more strains (Ferguson *et al*. 1999).

4. PARAMETER ESTIMATION

Given a sero-prevalence survey of N individuals performed at time t_0, the N_k individuals in each age-class k (where $k=1,\ldots,m$) can be classified into unexposed, n_{xk}, monotypically exposed to strain i, n_{ik}, and multitypically exposed, $N_k-n_{xk}-\sum_i n_{ik}$. In order to use the framework developed above to estimate strain-specific and time-varying forces of infection, we approximate the continuous $\lambda_i(t)$ profiles in a stepped piecewise manner with knot locations coincident with age-class boundaries, such that

$$\lambda_i(t)=\lambda_{ik}\quad\text{for } t_0-a_k\leqslant t<t_0-a_{k-1}, \tag{27}$$

where $(a_0,a_1,\ldots,a_{m+1})$ are the boundaries of the age classes used.

The multinomial log-likelihood of the data is then given by

$$\begin{aligned}l(\lambda_{ik},\rho_{ij})=\sum_{k=1}^m\Bigg[&n_{xk}\ln[x(a_k,t_0)]+\sum_i n_{ik}\ln[z_i(a_k,t_0)]\\&+\left(N_k-n_{xk}-\sum_i n_{ik}\right)\ln\Bigg[1-x(a_k,t_0)\\&-\sum_i z_i(a_k,t_0)\Bigg]\Bigg],\end{aligned} \tag{28}$$

(ignoring an additive constant). Here $x(a_k,t_0)$ is calculated from the λ_{ik} using equations (27) and (9). Similarly the $z_i(a_k,t_0)$ are calculated using equation (14). The advantage of using a stepped piecewise representation of $\lambda_i(a,t)$ is that it enables easy analytical evaluation of the integrals in equations (7) and (14).

Maximum likelihood estimates of the parameters are obtained through maximization of equation (28) using nonlinear optimization methods. The goodness-of-fit of the model can be judged by the comparison of the maximized log-likelihood, denoted l_{MAX}, and the saturated log-likelihood, denoted l_{SAT}. The saturated log-likelihood is defined as

$$\begin{aligned}l_{\mathrm{SAT}}=\sum_{k=1}^m\Bigg(&n_{xk}\ln\left[\frac{n_{xk}}{N_k}\right]+\sum_i n_{ik}\ln\left[\frac{n_{ik}}{N_k}\right]\\&+\left(N_k-n_{xk}-\sum_i n_{ik}\right)\ln\left[1-\frac{n_{xk}+\sum_i n_{ik}}{N_k}\right]\Bigg).\end{aligned} \tag{29}$$

The statistic $X^2=2(l_{\mathrm{MAX}}-l_{\mathrm{SAT}})$ is less for better fitting models, and is asymptotically χ^2-distributed with $d=D-P$ degrees of freedom, where $D=5m$ (given four dengue strains) being the number of degrees of freedom in the data, and P being the number of parameters fitted. A model with a goodness-of-fit p-value greater than greater than 0.05 is regarded as fitting the data well.

Calculating the boundaries of the 95% confidence region around the best-fit parameter set is impractical for high-dimensional, nonlinear models, so we approximate univariate confidence intervals by determining the range along the parameter coordinate axes through the best fit point in which the log-likelihood is within $\frac{1}{2}\chi^2_{1,0.95}$ of the maximum log-likelihood, l_{MAX}.

5. APPLICATION

Sangkawibha *et al*. (1984) reported on neutralizing antibodies specific to dengue virus antigens in serum samples from 1009 children of less than one to ten years of age in Rayong. The sera were collected in early 1980 prior to the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

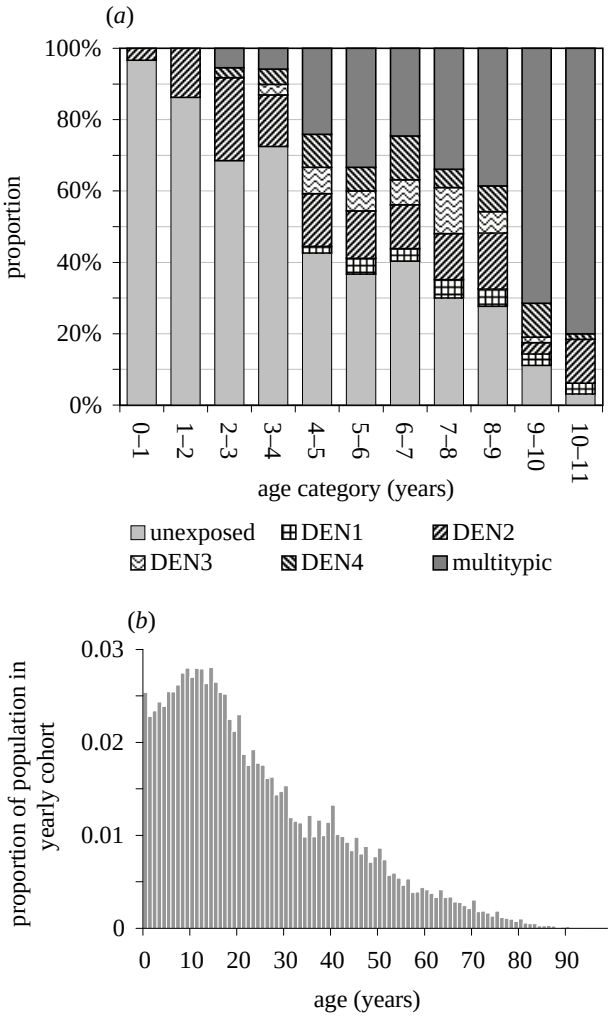

Figure 1. (*a*) Sero-prevalence estimates for the proportion unexposed, seropositive for a single strain or multitypically sero-positive. Data were collected from 1009 children between zero and ten years of age in Rayong, Thailand in 1980 (Sangkawibha *et al*. 1984). (*b*) Age profile of population in Rayong, Thailand calculated from 1980 census data.

dengue epidemic of that year. The data are shown in figure 1*a*, categorized by age-group as those with no antibody (susceptible), those with monotype antibodies to strains 1, 2, 3 and 4, and those with multitype antibodies.

Annual age data for the population of Rayong province in Thailand (National Statistical Office 1983) were obtained from the 1980 population and housing census. The reported population size for the province was 339 196 and the age profile is shown in figure 1*b*.

We applied the methodology developed above to obtain estimates of the strain-specific force of infection between 1969 and 1979, and of R_0 in 1980. We fitted a nested set of models in order to test whether adding complexities such as ADE of susceptibility, or allowing for the forces of infection to vary with time, significantly improved goodness-of-fit. The models (labelled 1 to 6) are defined as follows.

(i) Model 1. Non-interacting strains ($\rho_{ij} = 1$ for $i \neq j$), constant forces of infection ($\lambda_{ik} = \lambda_i$)—four parameters fitted.

(ii) Model 2. Interacting strains, with identical enhancement–inhibition for all strain combinations ($\rho_{ij} = \rho_{..}$ for $i \neq j$), constant forces of infection—five parameters fitted.

(iii) Model 3A. Interacting strains, with enhancement–inhibition dependent on strain of primary infection ($\rho_{ij} = \rho_{i.}$ for $i \neq j$), constant forces of infection—eight parameters fitted.

(iv) Model 3B. Interacting strains, with enhancement–inhibition dependent on strain of secondary infection ($\rho_{ij} = \rho_{.j}$ for $i \neq j$), constant forces of infection—eight parameters fitted.

(v) Model 4. Non-interacting strains ($\rho_{ij} = 1$ for $i \neq j$), time-varying forces of infection (λ_{ik} fitted for $k = 1, \ldots, 11$)—44 parameters fitted.

(vi) Model 5. Interacting strains, with identical enhancement–inhibition for all strain combinations ($\rho_{ij} = \rho_{..}$ for $i \neq j$), time-varying forces of infection—45 parameters fitted.

(vii) Model 6A. Interacting strains, with enhancement–inhibition dependent on strain of primary infection ($\rho_{ij} = \rho_{i.}$ for $i \neq j$), time-varying forces of infection—48 parameters fitted.

(viii) Model 6B. Interacting strains, with enhancement–inhibition dependent on strain of secondary infection ($\rho_{ij} = \rho_{.j}$ for $i \neq j$), time-varying forces of infection—48 parameters fitted.

The difference between models 3A and 3B (or 6A and 6B) is the manner in which the full matrix of ρ_{ij} value is simplified. Models 3A and 6A assume enhancement–inhibition of susceptibility is entirely determine by the serotype of the primary infection, and does not depend which strain the individual is later exposed to. Models 3B and 6B assume the converse—that the degree of enhancement–inhibition depends on the serotype of the secondary strain.

Tables 1 and 2 list the parameter estimates (except for the time-varying λ_{ik}) and goodness-of-fit statistics for the above models, together with corresponding estimates of R_{0i}. Note that all models allowing for time-varying forces of infection fit significantly better than those which do not ($p < 0.001$), and that models allowing for strain interaction also fit significantly better than those which assume strain independence ($p < 0.001$). Furthermore models 3A and 3B fit significantly better than model 2 ($p = 0.001$ and $p = 0.017$, respectively), and similarly models 6A and 6B fit significantly better than model 5 ($p = 0.028$ and $p = 0.030$, respectively).

Where we assume the λ_i values do not vary with time, there are sufficient degrees of freedom to fit all elements of the ρ_{ij} matrix, but the resulting model (estimates not presented) does not fit any better than model 3A ($p > 0.999$).

Figures 2–5 show the estimated past force of infection time-series for models 4, 5, 6A and 6B. Whilst the detailed structure and magnitude of the pattern of these time-series differ, it is striking to note that all suggest significant three- or four-year oscillations in dengue infection incidence. Clearly, before such a conclusion can be definitively confirmed, analysis of other stratified serological surveys will be required, but it is supported by data from hospitalization reports (Briseno-Garcia *et al*. 1996) and clinical observation (Dove 1998). Oscillations of this type of period are also predicted by theoretical studies of the

Table 1. *Parameter estimates (with approximate 95% confidence intervals for fitted parameters) for models assuming constant force-of-infection over time (models 1–3)*

model	parameter estimates		R_0 estimates		goodness-of-fit		
			method 1	method 2	X^2	d.f.	p-value
1	λ_i				175.97	51	<0.001
DEN 1	0.027 (0.021,0.034)		3.49	4.82			
DEN 2	0.075 (0.066,0.085)		4.39	5.63			
DEN 3	0.046 (0.039,0.054)		3.74	5.08			
DEN 4	0.043 (0.036,0.051)		3.69	5.03			
2	λ_i	$\rho_{..}$ 1.68 (1.47,1.91)			138.91	50	<0.001
DEN 1	0.022 (0.017, 0.028)		2.92	4.00			
DEN 2	0.067 (0.059,0.075)		3.79	4.86			
DEN 3	0.039 (0.033,0.046)		3.15	4.26			
DEN 4	0.036 (0.030,0.043)		3.10	4.21			
3A	λ_i	$\rho_{i.}$			123.35	47	<0.001
DEN 1	0.009 (0.006,0.012)	0.00 (0.00,0.46)	3.50	4.40			
DEN 2	0.106 (0.096,0.116)	5.41 (4.58,5.41)	4.97	6.48			
DEN 3	0.024 (0.019,0.029)	0.50 (0.15,0.91)	4.00	4.73			
DEN 4	0.026 (0.021,0.031)	0.87 (0.47,1.36)	4.01	4.84			
3B	λ_i	$\rho_{.j}$			128.73	47	<0.001
DEN 1	0.013 (0.011,0.016)	10.18 (8.19,12.41)	1.38	1.69			
DEN 2	0.084 (0.073,0.095)	0.00 (0.00,0.33)	7.73	8.47			
DEN 3	0.031 (0.026,0.037)	2.83 (1.90,3.88)	3.29	4.07			
DEN 4	0.036 (0.029,0.045)	0.24 (0.00,1.09)	6.87	7.70			

transmission dynamics of the virus (Ferguson *et al*. 1999; Gupta *et al*. 1998), where they are generated by immune-mediated competition or cooperation between strains.

Figures 2–5 also suggest very high dengue incidence between 1969 and 1971. This arises from the significant jump in multitypic sero-prevalence between ages nine and ten (figure 1*a*). In the absence of additional sero-prevalence data from surveys performed a few years before or after that analysed here, it cannot be ruled out that this increase is in fact due to age-dependent effects. However, it is difficult to conceive of age-dependence explaining the oscillatory nature of the incidence patterns seen in our results.

The R_0 estimates in table 1 suffer slightly from our inability to accurately estimate the $(1/\lambda_i(t))(\mathrm{d}\lambda_i(t)/\mathrm{d}(t)) = \mathrm{d}\ln\lambda_i(t)/\mathrm{d}t$ correction term in equation (26). However, given that the duration of infectiousness for dengue is thought to be around three to four days ($\sigma \simeq 100$), and the observed oscillations are occurring on an approximately three-year time-scale, it is likely that this term will always have a magnitude significantly less than unity. Also, since the sero-prevalence data used did not cover people over 11 years of age, approximation methods are needed to estimate R_0 for the entire population. We used two methods, both of which produced very similar results. The first (method 1) assumed $\lambda_i(t)$ prior to 1969 was constant and equal to the average of the estimates over the period 1969–79, and then used equations (9) and (14) to calculate estimates of $x(a,t_0)$ and $z_i(a,t_0)$ (for $t_0 = 1980$) which were used to estimate R_{0i} using equation (26). The second method (method 2) just uses the $\lambda_i(t)$ estimates for the 1969–79 period to calculate estimates of $x(a,t_0)$ and $z_i(a,t_0)$ for $0 \leqslant a \leqslant 11$, and then estimates R_{0i} from equation (26) under the assumption that those under 11 years of age make up the whole population (i.e. renormalizing $f(a)$ on [0,11]). We then

Table 2. *Parameter estimates (with approximate 95% confidence intervals for fitted parameters) for models fitting time-varying forces-of-infection (models 4–6)*

			R_0 estimates		goodness-of-fit		
model		parameter estimates	method 1	method 2	X^2	d.f.	p-value
4					42.67	11	<0.001
	DEN 1		4.72	4.96			
	DEN 2		5.93	5.91			
	DEN 3		4.85	5.13			
	DEN 4		4.98	5.23			
5		$\rho_{..}$			22.63	10	0.012
		1.48					
		(1.31,1.68)					
	DEN 1		4.12	4.33			
	DEN 2		5.38	5.39			
	DEN 3		4.28	4.53			
	DEN 4		4.36	4.59			
6A		$\rho_{i.}$			13.50	7	0.061
	DEN 1	0.00	4.29	4.47			
		(0.00,0.34)					
	DEN 2	0.72	5.12	5.06			
		(0.46,1.06)					
	DEN 3	1.23	4.61	4.81			
		(0.89,1.61)					
	DEN 4	3.35	5.50	5.75			
		(2.80,3.99)					
6B		$\rho_{.j}$			13.70	7	0.057
	DEN 1	0.58	5.64	5.89			
		(0.03,1.32)					
	DEN 2	2.96	4.12	4.20			
		(2.53,3.44)					
	DEN 3	1.33	4.21	4.73			
		(0.62,2.15)					
	DEN 4	0.00	7.86	7.60			
		(0.00,0.40)					

make use of the general result that R_0 is proportional to the life expectancy of the population for an endemic disease (Anderson & May 1991) to rescale these R_0 estimates by the ratio of the real life expectancy of the Rayong population (36.6 years, under the assumption that the population structure is stable) to that assumed when making the original 'reduced population' estimates (10.5 years).

In estimating $\lambda_i(t)$ and R_{0i}, we have made the simplifying assumption that the population structure of the Rayong population is stable. In fact, as is clear from figure 1*b*, this assumption is incorrect, as the population size has increased dramatically over the last 50 years (with the birth rate only starting to stabilize in the mid-1970s). Life expectancy has also increased in the last 50 years. However, extending the analysis to take account of changing population structure considerably complicates the mathematics and requires extensive longitudinal population data, and is therefore not considered further here (especially since another assumption, that the transmission coefficients of dengue strains have been constant over time, is also likely to be invalid over time-scales of the order of 50 years).

It should be noted that the values of R_0 obtained in this paper are considerably higher than estimates from past work (Koopman *et al*. 1991; Marques *et al*. 1994), which have ranged between 1.3 and 2.4, depending on the population being studied. Whilst some of this discrepancy may be due to different intensities of transmission in different regions, much is due to the methodological assumptions adopted in earlier work. Marques *et al*. (1994) estimated R_0 for dengue in Sao Paulo state, Brazil, from the rate of growth of clinical cases seen early in the 1990–91 epidemic, and obtained estimates between 1.6 and 2.4 for different cities. However, this method of estimation requires very precise knowledge of the generation time of the disease (the mean time between one individual being infected and another individual being infected by the first, via a vector), and assumes that the transmission coefficient of the disease is unvarying early in the epidemic. In fact, both the transmission coefficient and generation time will depend on vector density, and it is known that vector densities vary significantly throughout the year on a seasonal basis. In this case, dengue epidemics would be expected to be synchronized with the seasonal cycle in vector density, so estimates of R_0 made from data collected at the start of an epidemic are likely to be lower than those seen from estimates made at the peak of the epidemic. Furthermore, the complex relationship between dengue infection and disease (Halstead 1988; Kliks 1990; Thein *et al*. 1997), and

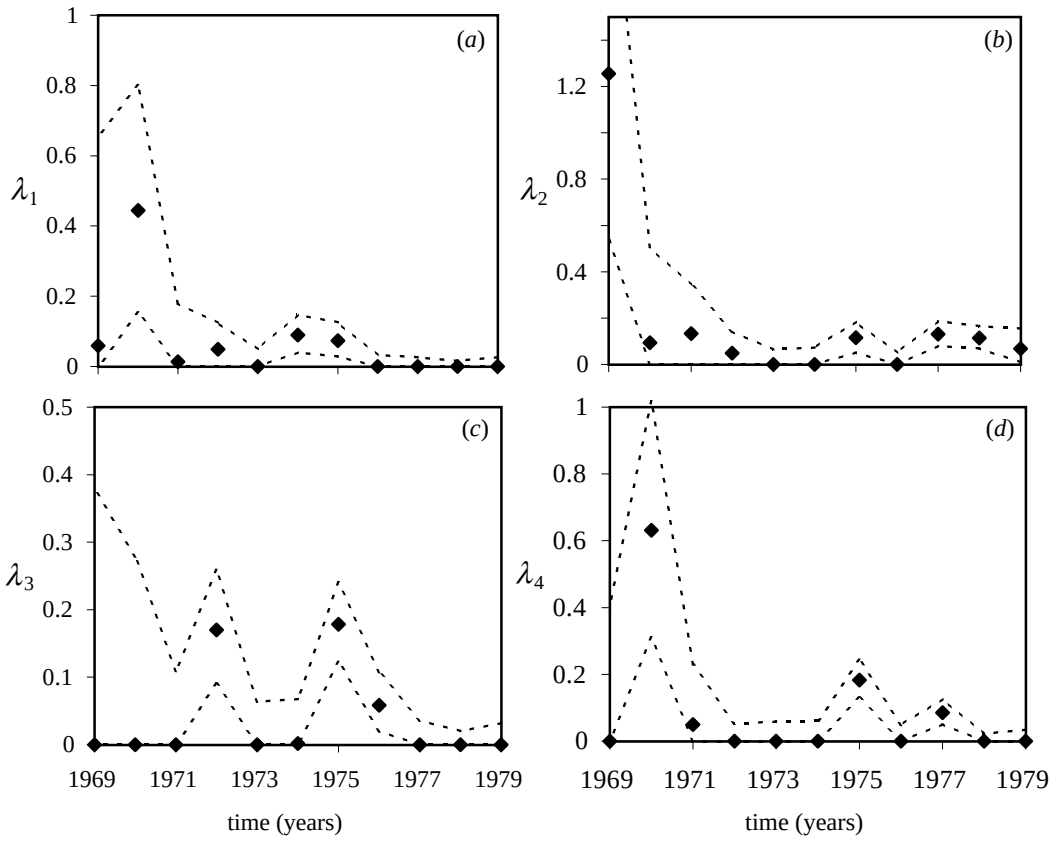

Figure 2. Estimated strain-specific force of infection estimates ($\lambda_i(t)$) for the period 1969–79 (points), with approximated univariate 95% confidence intervals (dashed lines), calculated using model 4 (see text). (*a*) DEN 1; (*b*) DEN 2; (*c*)DEN 3; (*d*) DEN 4.

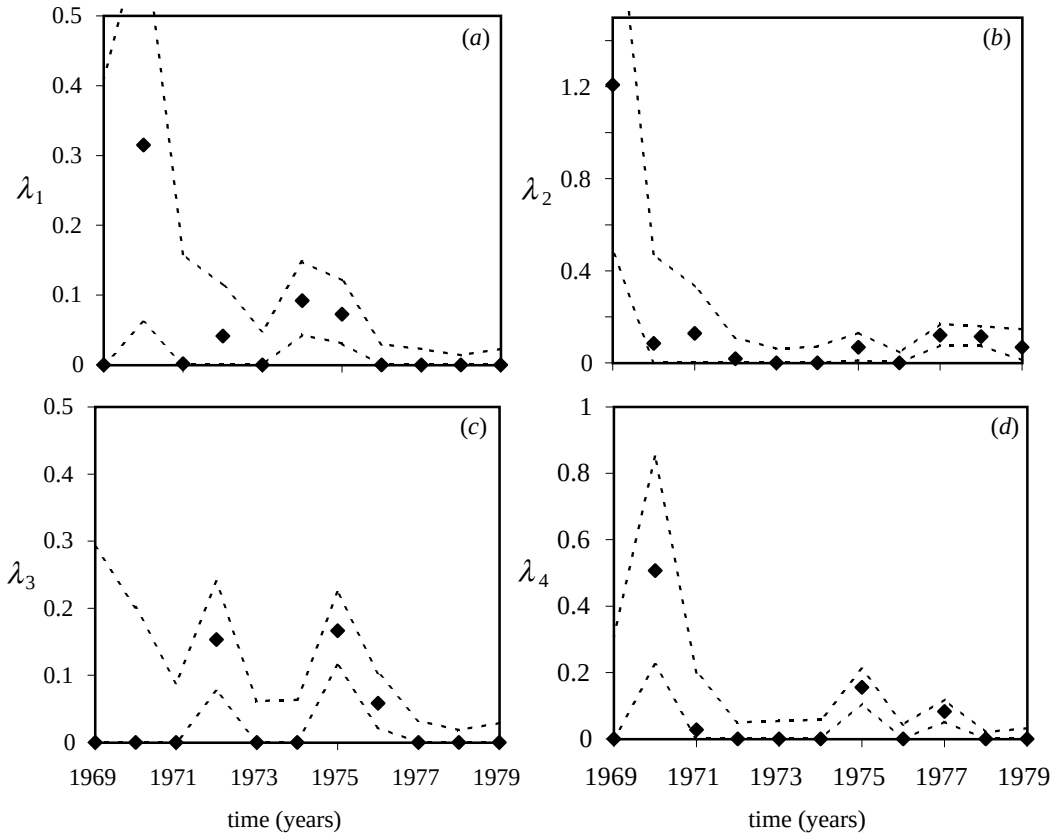

Figure 3. Estimated strain-specific force of infection estimates ($\lambda_i(t)$) for the period 1969–79 (points), with approximated univariate 95% confidence intervals (dashed lines), calculated using model 5 (see text). (*a*) DEN 1; (*b*) DEN 2; (*c*)DEN 3; (*d*) DEN 4.

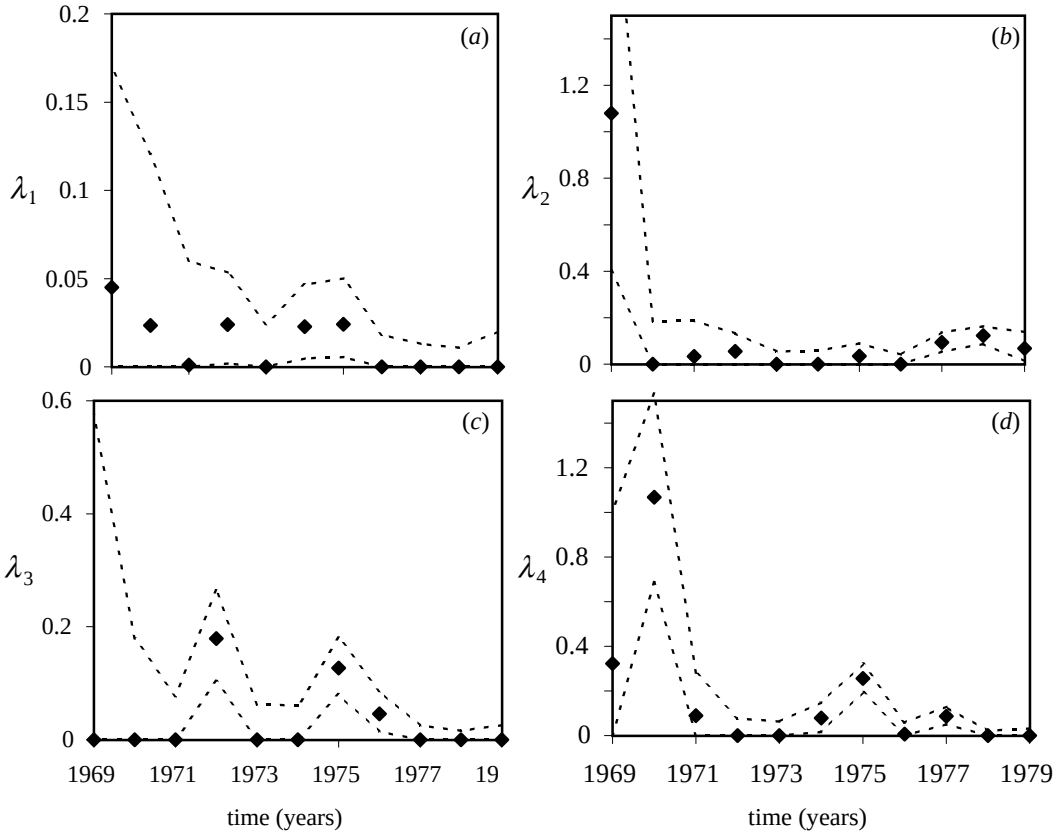

Figure 4. Estimated strain-specific force of infection estimates ($\lambda_i(t)$) for the period 1969–79 (points), with approximated univariate 95% confidence intervals (dashed lines), calculated using model 6A (see text). (*a*) DEN 1; (*b*) DEN 2; (*c*)DEN 3; (*d*) DEN 4.

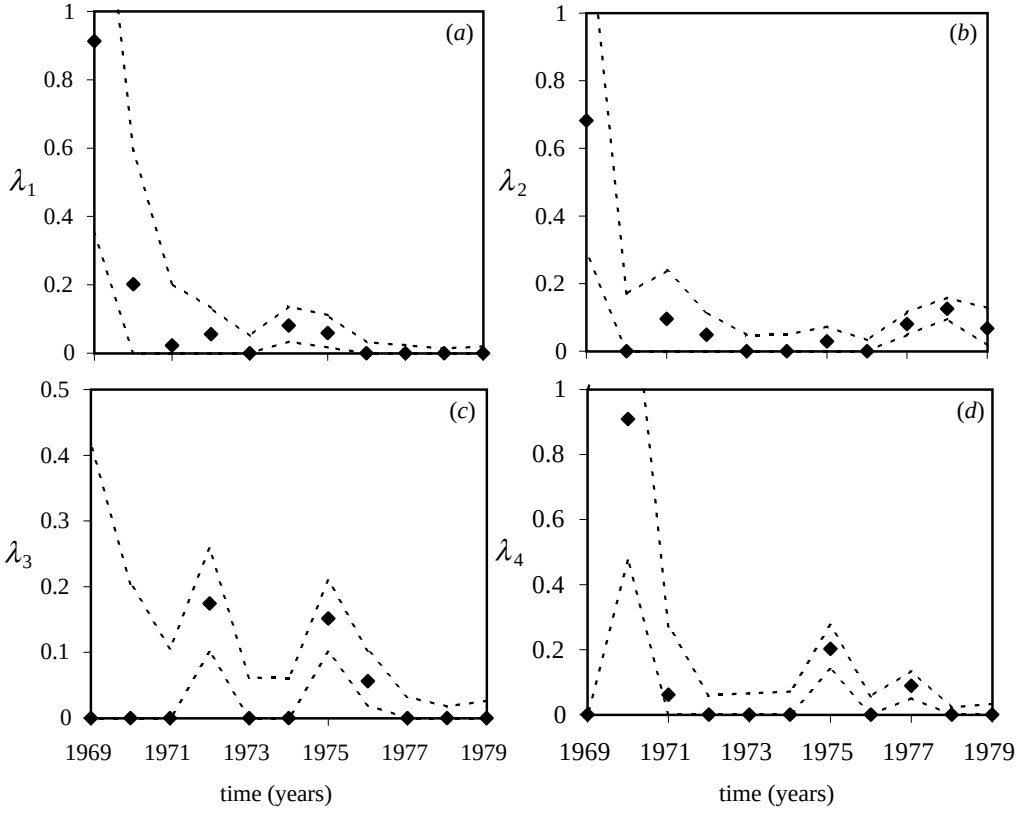

Figure 5. Estimated strain-specific force of infection estimates ($\lambda_i(t)$) for the period 1969–79 (points), with approximated univariate 95% confidence intervals (dashed lines), calculated using model 6B (see text). (*a*) DEN 1; (*b*) DEN 2; (*c*)DEN 3; (*d*) DEN 4.

potentially variable diagnosis rates, make estimation of R_0 from longitudinal disease data problematic.

Koopman *et al*. (1991) estimated R_0 for 70 locations in Mexico from serological data using a result derived from closed population (no births or deaths) stochastic epidemic theory. Because at the time of the survey dengue had been endemic (with seasonally induced oscillations in incidence) in Mexico for at least eight years, this assumption may have given rise to significant underestimation of R_0. The authors interpreted a prevalence of 0.9 as indicating $R_0 \simeq 2.5$. Under the assumption that dengue was endemic, the same prevalence would have given an estimate of $R_0 = 10$, since the proportion exposed to an endemic infection is given by $z = 1 - 1/R_0$. Reanalysis of the Mexico data using age-stratification to take account of the fact that the disease was not yet in endemic equilibrium would therefore be likely to give estimates for R_0 in the areas of Mexico with most intense transmission that would be consistent with those presented here for Rayong, Thailand.

6. CONCLUSIONS

The methods described in this paper can be employed to gain estimates of per capita rates or forces of infection for any virus that may persist as a set of closely related serotypes. In turn these estimates permit assessment of the degree of vaccination required to block transmission via the calculation of the values for the strain-specific R_0 values. Approximate values for the level of cohort vaccine uptake, p, required to eradicate transmission may then be derived given the relationship $p > 1 - 1/R_0$. Where many strains are involved and the vaccine consists of a cocktail of antigens from all strains, the magnitude of p_c will be set by the strain with the highest R_0 value.

In the case of dengue infection, the key objective was to obtain estimates of the basic reproduction number for each serotype in a defined location from the strain-specific forces of infection. For the best-fitting model investigated (model 6A), the estimated R_0 values show relatively little variation by serotype. Furthermore, the R_0 estimates are broadly similar in magnitude for all models explored, despite wide variation in goodness-of-fit to the serological data.

Conservatively picking the model which gives the largest average R_0 value across all strains (5.6, from model 6B, method 2), the magnitude of p_c is 0.82. In other words, if a vaccine becomes available to protect against infection by all four serotypes, roughly 85% of each birth cohort would have to be immunized before their first birthday to eradicate transmission. As noted above, this value is considerably higher than that which would have been calculated on the basis of earlier estimates of R_0 for dengue (Koopman *et al*. 1991; Marques *et al*. 1994), and implies that—at least in endemic areas—dengue will require considerable effort to eliminate even once a vaccine becomes available. Of course, the precise value of the average R_0—and thus p_c—will depend on geographical location, since both vector and viral reproductive rates have been shown to be temperature-dependent (Koopman *et al*. 1991; Holmes *et al*. 1998).

For simplicity, the methods developed in this paper to estimate R_0 ignore many sources of heterogeneity that influence the force of infection, the incubation and infectious periods of the infection, and other factors. Of particular importance is age-dependence in the force of infection, the severity of which can strongly influence the best age at which to immunize, and the manner in which cohort immunization influences the incidence of infection and disease. The methods described in this paper for multiple strains can easily be extended to cover age-dependence in exposure to infection, but either data on age-dependence in exposure or additional longitudinal data on infection incidence is required before such effects can be resolved.

The main focus of attention in the current paper is the treatment of strain structure in the estimation of rates of infection and the handling of various assumptions concerning how the different strains influence each other when all are co-circulating in the same human community. Of the various models fitted to the serological data from Thailand, the model which assumed both a time-varying force of infection and strain interaction (enhancement) gave the best performance with the maximum likelihood fitting procedure adopted for parameter estimation. This result therefore adds some additional support to the hypothesis that ADE is significant in the epidemiology of dengue infection and associated disease. However, to date, the analysis was not able to address the question of whether any particular sequence of exposure to the different serotypes is more likely to lead to serious disease.

An interesting implication of our results is that the force of infection varies on a three- to four-year time period. Inspection of the limited longitudinal data on the incidence of dengue-related disease in children in different locations also provides a strong indication of both seasonal and longer-term periodicity. This can be detected despite the complication of a rise in the annual incidence of dengue-related disease over the past few decades. Our analysis of the cross-sectional serological data provides further support for the presence of the periodicity in the transmission dynamics of the virus. Recent theoretical studies of the transmission dynamics of dengue taking into account the multi-strain composition, cross-protective immune responses and possible antibody-dependent enhancement of transmission suggest epidemiological mechanisms which might generate such oscillatory behaviour (Ferguson *et al*. 1999).

In future, more refined analyses may become possible, given the development of new serological techniques that can define the precise mix of strains experienced by individuals with more than one exposure. A more pessimistic view would be that the current rapid rate of evolution of the virus may lead to the emergence of new serotypes that will further complicate serological studies. Given the severity of the diseases induced by dengue infection and the rapid rise in the global incidence of the infection in tropical regions of the world over the past two decades, it is to be hoped that investment is increased, hastening progress in vaccine development. If, and when, a vaccine becomes available, the methods described in this paper can help to define the targets for cohort-based vaccination programmes.

N.M.F. thanks the Royal Society for research grant support. C.A.D. and R.M.A. thank the Wellcome Trust.

REFERENCES

Ades, A. & Nokes, D. 1993 Modeling age- and time-specific incidence from seroprevalence: toxoplasmosis. *Am. J. Epidemiol*. **137**, 1022–1034.

Anderson, R. & May, R. 1985 Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes. *J. Hyg. Camb*. **94**, 365–436.

Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans: dynamics and control*, 2nd edn. Oxford University Press.

Becker, N. G. 1989 *Analysis of infectious disease data*. London: Chapman & Hall.

Briseno-Garcia, B. (and 9 others) 1996 Potential risk for dengue hemorrhagic fever: the isolation of serotype dengue 3 in Mexico. *Emerg. Infect. Dis*. **2**, 133–135.

Dietz, K. 1975 Transmission and control of arbovirus diseases. In *Epidemiology* (ed. D. Ludwig & K. L. Cooke), pp. 104–121. Philadephia, PA: Soceity for Industrial and Applied Mathematics.

Dietz, K. & Schenzle, D. 1985 Proportionate mixing models for age-dependent infection transmission. *J. Math. Biol*. **22**, 117–120.

Dove, A. 1998 Dengue fever on the increase. *Nature Med*. **4**, 543.

Ferguson, N., Anderson, R. & Gupta, S. 1999 The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple strain pathogens. *Proc. Natl Acad. Sci. USA* **96**, 790–794.

Grenfell, B. & Anderson, R. 1985 The estimation of age-related rates of infection from case notifications and serological data. *J. Hyg. Camb*. **95**, 419–436.

Griffiths, D. 1974 A catalytic model of infection for measles. *Appl. Statist*. **23**, 330–339.

Gritsun, T., Holmes, E. C. & Gould, E. 1995 Analysis of flavivirus envelope proteins reveals variable domains that reflect their antigenicity and may determine their pathogenesis. *Virus Res*. **35**, 307–321.

Gubler, D. 1988 Dengue. In *The arboviruses: epidemiology and ecology*, vol. II (ed. T. Monath), pp. 223–260. Boca Raton, FL: CRC Press, Inc.

Gubler, D. 1998 Dengue and dengue hemorrhagic fever. *Clin. Microbiol. Rev*. **11**, 480–498.

Gubler, D. & Kuno, G. (ed.) 1997 *Dengue and dengue hemorrhagic fever*. New York: CAB International.

Gupta, S., Trenholme, K., Anderson, R. & Day, K. 1994 Antigenic diversity and the transmission dynamics of *Plasmodium falciparum*. *Science* **263**, 961–963.

Gupta, S., Ferguson, N. & Anderson, R. 1998 Chaos, persistence and evolution of strain structure in antigenically diverse infectious agents. *Science* **280**, 912–915.

Halstead, S. 1979 In vivo enhancement of dengue virus infection in rhesus monkeys by passively transferred antibody. *J. Infect. Dis*. **140**, 527–533.

Halstead, S. 1982 Immune enhancement of viral infection. In *Progress in allergy* (ed. P. Kallos), p. 301. Basel, Switzerland: S. Karger.

Halstead, S. 1988 Pathogenesis of dengue: challenges to molecular biology. *Science* **239**, 476–481.

Halstead, S. 1993 Global epidemiology of dengue: health systems in disarray. *Trop. Med*. **35**, 137–146.

Halstead, S. B. & O'Rourke, E. J. 1977 Dengue viruses and mononuclear phagocytes. I. Infection enhancement by non-neutralising antibody. *J. Exp. Med*. **146**, 201–217.

Halstead, S., Chow, J. & Marchette, N. 1973 Immunological enhancement of dengue virus replication. *Nature New Biol*. **243**, 24–26.

Halstead, S. B., O'Rourke, E. J. & Allison, A. C. 1977 Dengue viruses and mononuclear phagocytes. II. Identity of blood and tissue leukocytes supporting in vitro infection. *J. Exp. Med*. **146**, 218–229.

Holmes, E., Bartley, L. & Garnett, G. 1998 The emergence of dengue: past, present, and future. In *Emerging infections* (ed. R. Krause), pp. 301–325. London: Academic Press.

Keiding, N. 1991 Age-specific incidence and prevalence: a statistical perspective. *J. R. Statist. Soc.* A **154**, 371–412.

Kliks, S. 1990 Antibody-enhanced infection of monocytes as the pathogenetic mechanism for severe dengue illness. *AIDS Res. Hum. Retroviruses* **6**, 993–998.

Kliks, S., Nimmanitya, S., Nisalak, A. & Burke, D. 1988 Evidence that maternal antibodies are important in the development of dengue hemorrhagic fever in infants. *Am. J. Trop. Med. Hyg*. **38**, 411–419.

Koopman, J., Prevots, R., Marin, M., Dantes, H., Aquino, M., Longini, I. & Amor, J. 1991 Determinants and predictors of dengue infections in Mexico. *Am. J. Epidemiol*. **133**, 1168–1178.

Kuno, G., Gubler, D. & Oliver, A. 1993 Use of the 'original antigenic sin' theory to determine the serotypes of previous dengue infections. *Trans. R. Soc. Trop. Med. Hyg*. **87**, 103–105.

MacDonald, G. 1957 *The epidemiology and control of malaria*. Oxford University Press.

MacDonald, G. 1965 The dynamics of helminth infections, with special reference to schistosomes. *Trans. R. Soc. Trop. Med*. **59**, 489–506.

Mady, B., Erbe, D., Kurane, I., Fanger, M. & Ennis, F. 1991 Antibody-dependent enhancement of dengue virus infection mediated by bispecific antibodies against cell surface molecules other than FC receptors. *J. Immunol*. **147**, 3139–3144.

Marques, C., Forattini, O. & Massad, E. 1994 The basic reproduction number for dengue fever in Sao Paulo state, Brazil: 1990–1991 epidemic. *Trans. R. Soc. Trop. Med. Hyg*. **88**, 88–89.

National Statistical Office 1983 1980 population & housing census: Changwat Rayong. Technical Report. Bangkok, Thailand.

Newton, E. & Reiter, P. 1992 A model of the transmission of dengue fever with an evaluation of the impact of ultra-low volume (ULV) insecticide applications on dengue epidemics. *Am. J. Trop. Med. Hyg*. **47**, 709–720.

Okuno, Y., Fukunaga, T., Srisupaluck, S., Kasemsarn, P., Dharakul, C. & Sangkawibha, N. 1980 Serological and virological studies on outpatients with dengue hemorrhagic fever (DHF) in Chanthaburi Province, Thailand. I. Serological evidence on paired sera from DHF patients by neutralization (N), hemagglutination inhibition (HI) and staining tests. *Biken J*. **23**, 113–121.

Page, R. D. M. & Holmes, E. C. 1998 *Molecular evolution: a phylogenetic approach*. Oxford: Blackwell Science.

Peiris, J., Gordon, S., Unkeless, J. & Porterfield, J. 1981 Monocolonal, anti-FC receptor IGG blocks antibody enhancement of viral replication in macrophages. *Nature* **289**, 189–191.

Porterfield, J. 1986 Antibody-dependent enhancement of viral infectivity. *Adv. Vir. Res*. **31**, 335–355.

Rigau-Perez, J., Clark, G., Gubler, D., Reiter, P., Sanders, R. & Vorndam, A. 1998 Dengue and dengue haemorrhagic fever. *Lancet* **352**, 971–977.

Robinson, W., Montefiori, D. & Mitchell, W. 1988 Antibody-dependent enhancement of HIV type 1 infection. *Lancet* **1**, 790–794.

Sangkawibha, N., Rojanasuphot, S., Ahandrik, S., Viriyapongse, S., Jatanasen, S., Salitul, V., Phanthumachinda, B. & Halstead, S. 1984 Risk factors in dengue shock syndrome: a prospective epidemiologic study in Rayong Thailand. I. The 1980 outbreak. *Am. J. Epidemiol*. **120**, 653–669.

Takeda, A., Tuazon, C. & Ennis, F. 1988 Antibody-enhanced infection by HIV via FC receptor-mediated entry. *Science* **242**, 580–583.

Thein, S., Aung, M. M., Shwe, T. N., Aye, M., Zaw, A., Aye, K., Aye, K. M. & Aaskov, J. 1997 Risk factors in dengue shock syndrome. *Am. J. Trop. Med. Hyg*. **56**, 566–572.

WHO 1986 Dengue hemorrhagic fever: diagnosis, treatment and control. Technical Report. Geneva: World Health Organization.