Title: Estimating dengue transmission intensity from case-notification data from multiple countries.

Authors: Natsuko Imai[1], Ilaria Dorigatti[1], Simon Cauchemez[2], Neil M. Ferguson[1]

Affiliations: 1. MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, United Kingdom. **2.** Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Paris, France.

Abstract:

Background: Despite being the most widely distributed mosquito-borne viral infection estimates of dengue transmission intensity and associated burden remain ambiguous. With advances in the development of novel control measures, obtaining robust estimates of average dengue transmission intensity is key for assessing the burden of disease and the likely impact of interventions.

Methodology/Principle Findings: We estimated the force of infection (λ) and corresponding basic reproduction numbers (R_0) by fitting catalytic models to age-stratified incidence data identified from the literature. We compared estimates derived from incidence and seroprevalence data and assessed the level of under-reporting of dengue disease. In addition, we estimated the relative contribution of primary to quaternary infections to the observed burden of dengue disease incidence. The majority of R_0 estimates ranged from one to five and the force of infection estimates from incidence data were consistent with those previously estimated from seroprevalence data. The baseline reporting rate (or the probability of detecting a secondary infection) was generally low (<25%) and varied within and between countries.

Conclusions/Significance: As expected, estimates varied widely across and within countries, highlighting the spatiotemporally heterogeneous nature of dengue transmission. Although seroprevalence data provide the maximum information, the incidence models presented in this paper provide a method for estimating dengue transmission intensity from age-stratified incidence data, which will be an important consideration in areas where seroprevalence data are not available.

Author Summary

With 40% of the world's population at risk of infection, dengue imposes a significant public health burden. Yet estimates of baseline transmission intensity are still sparse, making it difficult to implement efficient control programs. The authors used incidence data, which are abundant compared to seroprevalence data, to estimate dengue transmission intensity in 13 countries. Estimates derived from incidence data were comparable to those from seroprevalence data, an important conclusion for areas where seroprevalence data are not available. Additionally the estimated baseline reporting rates and the contribution of primary to tertiary/quaternary infections to observed disease in each country will help to highlight potential weaknesses in the country or region's surveillance system.

Introduction:

Dengue is the most widely distributed mosquito-borne viral infection, but assessment of its geographic variation in transmission remains challenging. Analysis based on mapping the probability of occurrence of dengue estimated that dengue causes 390 million annual infections worldwide [1]. However, these estimates relied on assuming a direct linear correlation between probability of occurrence and incidence, rather than estimating transmission intensity as quantified by the force of infection or reproduction number. Here we develop methods to estimate transmission intensity from routine, age-stratified surveillance data on suspected dengue case incidence.

All four serotypes of dengue virus (DENV-1, 2, 3, and 4) can cause dengue fever with the risk of severe dengue increasing with subsequent heterologous infections. Once infected, individuals develop long-lived protective homotypic immunity and short-lived heterotypic immunity [2,3]. Once antibody levels wane below the threshold required to provide protection, antibody-dependent enhancement (ADE) becomes a risk, leading to secondary heterologous infection having an increased risk of causing clinically apparent disease [4,5]. Hence, while the majority of primary dengue infections are asymptomatic [6,7], secondary heterologous infection has been identified as a major risk factor for symptomatic and severe dengue [8–10]. Therefore the majority of cases seen in hospitals [11] or reported via surveillance systems [12] tend to be secondary infections [7].

In previous work, we estimated dengue transmission intensity from age-stratified seroprevalence data but highlighted the relative paucity of seroprevalence data compared with routine surveillance data on the incidence of suspected dengue [13]. This reflects dengue fever, dengue haemorrhagic fever (DHF), and dengue shock syndrome (DSS) being notifiable diseases in most countries [14–18]. Indeed, in many countries, incidence reports are the only type of data available. However the clinical diagnostic criteria vary and different countries have their own reporting standards [19]. The World Health Organisation (WHO) collates surveillance data from dengue affected countries via its DengueNet system, but the data are not always updated regularly and there can be inconsistencies with other sources (e.g. WHO regional offices or countries) of national and subnational data [19].

The lack of systematic data on dengue incidence, the lack of standardized reporting procedures or diagnostic criteria, and the lack of integration between private and public sectors makes accurate estimation of the true dengue burden difficult [20]. Previous studies have attempted to estimate the burden of dengue and associated economic costs in South East Asia and South America by calculating expansion factors from systematic literature reviews, collation of existing data, and population-based cohorts [20–24]. However, the lack of standardisation also affects the validity of expansion factors (calculated by dividing the cumulative incidence of dengue cohort studies by that from passive data at national and local levels) as estimates of underreporting. Due to the wide spectrum of clinical manifestations and the lack of routine laboratory testing, dengue is globally underreported and analyses of officially reported dengue numbers need to take this into account [25].

While reported incidence levels cannot be relied upon to directly quantify disease burden, the age distribution of dengue cases provides more reliable information on dengue transmission intensity. Here we propose an approach for estimating average transmission intensity - as quantified by the force of infection or basic reproduction number (R_0) – from age-stratified incidence data. We compare estimates derived from seroprevalence and incidence data and assess the level of under-reporting of dengue disease. In addition, we estimate the relative contribution of primary to quaternary infections to the observed burden of dengue disease incidence.

Methods:

Literature Search

Web of Knowledge and PubMed were searched for age-stratified incidence data since 1980 as we were interested in contemporary dengue transmission and wanted to be consistent with our previous study where we collated age-stratified seroprevalence data [13]. Search terms used were ‘dengue’ and ‘age’ and (‘incidence’ or ‘cases’ or ‘notifications’ or ‘notified cases’) with inclusion criteria mapped to subject headings. Additional web-based searches were performed to augment the primary literature search. Data were extracted from published datasets

where authors reported age-stratified incidence data with corresponding population age-structure data.

Estimating the force of infection and reporting rates

We consider a population stratified into M age groups and denote a_j and a_{j+1} the lower and upper age bounds respectively of age group j ($j=0,\ldots, M-1$). Our model assumes perfect homotypic protection following infection with any serotype. Thus, an individual can experience a maximum of four dengue infections in their life (corresponding to the four dengue serotypes). Ideally, we would allow forces of infection to vary by serotype (DENV-1 to DENV-4). However as serotype-specific data were not available, we assumed circulating serotypes were equally transmissible, i.e. had the same force of infection, λ, which does not vary over time. The incidence of primary infections (I_1) for any one serotype for people in an age group j is calculated as the integral of the probability of being seronegative to all four strains at age a multiplied by four times the constant serotype-specific infection hazard, λ, (since primary infection can occur with any of the four serotypes). Age a spans the range $[a_j, a_{j+1}]$, as described by the bounds of integrations (equation 1).

$$\begin{aligned} I_1(j) &= \int_{a_j}^{a_{j+1}} 4\lambda\left(e^{-\lambda a}\right)^4 da \\ &= \int_{a_j}^{a_{j+1}} 4\lambda e^{-4\lambda a} da = \left(e^{-4\lambda a_j} - e^{-4\lambda a_{j+1}}\right) \end{aligned} \quad (1)$$

The incidence of secondary, tertiary, and quaternary infections in age group j ($I_2(j)$, $I_3(j)$, and $I_4(j)$) are calculated in a similar fashion. If fewer than four serotypes have circulated in an area, then the number of infections an individual can have changes accordingly. Full details are given in the Supporting Information (SI Text).

The average observed annual disease incidence rate per person in age group j is then given by the weighted sum of the primary to quaternary infection rates (equation 2):

$$D(j) = \frac{\rho}{w(j)}\left\{I_2(j) + \gamma_1\left(I_1(j) + \gamma_3\left(I_3(j) + I_4(j)\right)\right) + B\right\} \quad (2)$$

where $w(j) = a_{j+1} - a_j$ is the width of age group j, ρ is the probability that a secondary infection results in a detected dengue case (reporting rate), γ_1 is the probability that a primary infection is detected relative to a secondary infection, and γ_3 is the probability that a tertiary or quaternary infection is detected relative to a primary infection. Here B is a baseline risk of disease used to represent any non-dengue related illnesses that are misdiagnosed as dengue, and was only estimated when fitting suspected dengue incidence data where laboratory confirmation was lacking.

We assumed that secondary infections were more likely to be symptomatic than primary infections [7,26] and that post-secondary infections were even less likely to be symptomatic than primary infections, i.e. $\rho > \gamma_1 > \gamma_3$. Single values of γ_1 and γ_3 were estimated per country. For datasets that reported DHF only, we assumed that DHF cases only arose from secondary infections and set γ_1 and γ_3 to zero [27,28]. Where less than four serotypes were in circulation, we adjusted our calculation of expected incidence accordingly – full details are given in the SI Text. Where data on the age distribution of the population was not provided in the source publications, the population age-structure closest to the survey population was used (taken from census data or from United Nations estimates) [29]. For the first model variant examined (model 1), we assumed a single baseline reporting rate (ρ) across all age groups. We also explored whether baseline reporting rates might differ with age (model 2) by estimating different reporting rates in children (ρ_{young}) and adults (ρ_{old}), also fitting the age threshold ($a_{threshold}$) defining the boundary between these groups (ρ_{young} for age $a < a_{threshold}$, otherwise ρ_{old}).

Where incidence data were available for multiple years, we fitted models 1 and 2 to individual years (model variants 1A and 2A). We also examined fitting to the cumulative incidence across the observation period, as this gives a better estimate of the long-term average distribution of incidence across age groups (models 1B and 2B). When fitting to the cumulative incidence we calculated the expected disease incidence by multiplying the annual expected disease incidence by the number of years in the study. Overall, for models 1A and 1B, we estimated up to 5 parameters

(λ, ρ, γ_1, γ_3 and B), while for models 2A and 2B we estimated up 7 parameters (λ, ρ_{young}, ρ_{old}, $a_{threshold}$, γ_1, γ_3 and B). All models were fitted to the data using a Metropolis-Hasting Markov Chain Monte Carlo (MH MCMC) algorithm using a Dirichlet-multinomial log-likelihood with uniform priors in version 3.1.0 of the R statistical language [30]. Full details are given in the SI Text.

Calculating the Basic Reproduction Number (R_0)

We assumed dengue transmission was at endemic equilibrium and that the force of infection (λ) was constant in time. Since we did not have serotype-specific data, we additionally assumed that all serotypes in circulation were equally transmissible, i.e. had the same force of infection and basic reproduction number, and that there were no interactions between serotypes. We estimated a strain-specific basic reproduction number (R_0) from the single force of infection (λ) estimated under two different assumptions about the number of infections required to acquire complete immunity. Under assumption one, complete protection is acquired upon quaternary infection. Under assumption two, complete protection is reached after secondary infection (or if tertiary and quaternary infections occur, they are not infectious). These assumptions match that of our previous work estimating the force of infection from serological data and allowed us to compare the R_0 estimates obtained from both types of data [13]. Full details are given in the SI Text.

Comparing force of infection estimates by data type

We used weighted regression to assess how comparable force of infection estimates obtained from cumulative incidence data were with those derived from seroprevalence data described previously [13] and from four additional seroprevalence datasets (see SI Text, Table S1). Location- and time-matched incidence and serology data were not available, so we matched datasets by country, region, and survey year. Since seroprevalence data represent all past infections, we compared force of infection estimates with those obtained from cumulative incidence data rather than yearly incidence data where possible (see SI Text, Table S2 for full details on pairings). We used the weighted regression method described by Ripley and Thompson [31] which explicitly accounts for measurement errors in both force of infection estimates from seroprevalence data (y-axis) and incidence data (x-axis) to

estimate the MLE line. This was implemented using the *deming* package in R [32]. Full details are given in the SI Text.

Results:

We identified 23 papers reporting incidence data. Fig. 1 describes the search process and Table 1 summarises the studies identified. Seven papers reported age-stratified incidence data from multiple years, one paper reported data where the number of serotypes in circulation had changed over the survey years, six papers reported cumulative age-stratified incidence data, eight papers reported age-stratified incidence data from a single year, and two papers reported age-stratified incidence data from multiple countries.

Fig. 1: Flowchart describing the literature search process for age-stratified incidence data.

The identified studies provided a total of 34 datasets from 13 countries. The years included ranged from 1978 to 2011. The dataset reporting incidence data from 1978 was included since data were presented for the ten-year time period of 1978 – 1988 [33]. Of the 23 papers reporting incidence data, ten reported dengue incidence at the national level and only two studies reported cases detected via active as well as passive surveillance. Three additional surveys were obtained from the Ministry of Health in Thailand that reported age-specific incidence from Bangkok (2000), Ratchaburi (2000), and Rayong (2010) [34].

As expected, force of infection estimates varied widely between countries, with less variation seen within countries. Fig. 2 shows the distribution of the total force of infection (λ_{total}) grouped by country (calculated by multiplying the serotype-specific force of infection by the number of serotypes in circulation). Individual estimates are given in the SI Text.

Fig. 2: Total force of infection and corresponding R_0 estimates from the model fitted to the incidence data grouped by country. Each dot represents the posterior median estimate and the error bars show the 95% CrI for each dataset. The box represents the country-specific central estimate calculated by taking the mean values of the MCMC output for each country (the line and limits of the box represents the posterior median and the 95% CrI respectively). R_0 assumption one: complete protection acquired upon quaternary infection, assumption two: complete protection reached after secondary infection.

Table 1: Summary of cross-sectional incidence datasets identified and associated demographics. *w/active surveillance.

Country	Survey Year	Region	Diagnosis	# serotypes in circulation	DF/DHF/DSS	Age range sampled	Estimated denominator population of study (3sf)^	Population size of study region	Urban/Rural	Ref
Brazil	1995-2001	Pernambuco State	Lab confirmed	2	All cases	0 - 80+	8360000	8.5M	Urban/Rural	[35]
	2002-2006	Pernambuco State	Lab confirmed	3	All cases	0 - 80+	8360000	8.5M	Urban/Rural	[35]
	2000-2009	Vitoria	Lab confirmed	3	All cases	0 - 80+	292000	0.28M–0.32M	Urban/Rural	[36]
	2001-2006	Amazon	Clinical	4	All cases	0 - 70	3480000	23.6M	Rural/Urban	[37]
Cambodia	2006-2008	Kampong Chan Province	Lab confirmed*	4	All cases	0 - 20	805000	90000	Urban/Rural	[38]
	2006-2007	Kampong Chan	Lab confirmed	4	All cases	0-14	14500	90000	Urban/Rural	[39]
China	1978-1988	Guangzhou	Clinical	4	All cases	0 - 71+	69700000	11.64M	Urban	[33]
	1989-1999	Guangzhou	Clinical	4	All cases	0 - 71+	69000000	11.64M	Urban	[33]
	2000-2009	Guangzhou	Clinical	4	All cases	0 - 71+	39500000	11.64M	Urban	[33]
	2005-2011	Guangdong	Clinical	4	All cases	0 - 80+	88900000	104.3M	Urban	[14]
Laos	2000-2006	National	Clinical/Lab	4	All cases	0 - 15+	4980000	5.4M	Urban/Rural	[40]
	2010	Savannakhet Province	Clinical	4	All cases	0 - 40+	4880000	0.83M	Urban	[41]
	2010	National	Clinical	4	All cases	0 - 40+	6390000	6.5M	Urban/Rural	[42]
Nicaragua	1999-2001	Leon	Lab confirmed	3	All cases	0 - 55	360000	0.39M	Urban	[43]
Philippines	1998-2005	National	Clinical/Lab	4	All cases	0 - 15+	71700000	77.7M	Urban/Rural	[40]
Puerto Rico	2006	Patillas	Lab confirmed	4	All cases	0 - 40+	16700	20200	Urban	[44]
	2007	National	Lab confirmed	4	All cases	0 - 70+	3820000	3.8M	Urban/Rural	[45]
	2010	National	Lab confirmed	4	All cases	0 - 70+	3720000	3.7M	Urban/Rural	[12]
	1994	National	Lab confirmed	3	All cases	0 - 75+	3530000	3.5M	Urban/Rural	[46]
	1995-1997	National	Lab confirmed	3	All cases	0 - 75+	3530000	3.5M	Urban/Rural	[46]
Singapore	1999-2005	National	Clinical/Lab	4	All cases	0 - 15+	2620000	4M	Urban	[40]
	2005	National	Lab confirmed	4	DF/DHF	0 - 80	3450000	4.3M	Urban	[47]
	2005	National	Lab confirmed	4	All cases	0 - 55+	4270000	4.3M	Urban	[48]
	2007	National	Lab confirmed	4	All cases	1 - 55+	4590000	4.6M	Urban	[48]
Sri Lanka	1997	National	Clinical	4	DF/DHF	0 - 65	17300000	17.3M	Urban/Rural	[49]

Table 1 continued (2/2)

Country	Survey Year	Region	Diagnosis	# serotypes in circulation	DF/DHF/DSS	Age range sampled	Estimated denominator population of study^	Population size of study region	Urban/Rural	Ref
Sri Lanka	1996-2005	National	Clinical	4	All cases	0 - 15+	17700000	17.3M	Urban/Rural	[40]
Taiwan	2003-2009	Kaohsiung City	Lab confirmed	4	All cases	0 - 74+	10600000	1.5M	Urban	[50]
Thailand	2000-2010	National	Clinical	4	All cases	0-65+	797000	66.4M	Urban/Rural	[51]
	2006-2007	Ratchaburi	Lab confirmed	4	All cases	0-14	6380	38208	Urban	[39]
	2000	Bangkok	Lab confirmed	4	DHF	0-65	5050	6355144	Urban	[34]
	2000	Ratchaburi	Lab confirmed	4	DHF	0-65	1370	791217	Urban	[34]
	2010	Rayong	Lab confirmed	4	All cases	0-72	1060	616916	Urban	[34]
Vietnam	1998-2009	Hanoi	Lab confirmed	4	All cases	0 - 80	6350000	6.5M	Urban	[52]
Yemen	2010	Hadramout	Lab confirmed	3	All cases	0 - 55+	797000	0.7M	Urban/Rural	[53]

^Calculated from the population size and reported incidence if survey numbers were not given in the source ref.

Estimates of R_0 varied according to the assumptions made regarding host immunity. Assuming only primary and secondary infections are infectious (assumption two) gave up to two-fold higher estimates of R_0 than when assuming tertiary and quaternary infections are also infectious (Fig. 2). This is consistent with our previous results analysing seroprevalence data [13]. Some force of infection estimates in Cambodia were very high, perhaps a result of the active surveillance undertaken as part of the study by Vong *et al* [38] (for all parameter estimates see SI Text). The baseline-reporting rate (ρ), defined as the probability of detecting a secondary infection, was below 15% when averaged across all studies (Fig. 3). The median probability of detecting a primary infection relative to that of detecting a secondary infection (γ_1) was less than 25% for the majority of datasets. However, the credible intervals for some γ_1 estimates were wide. The data proved uninformative about the contribution of post-secondary infections to disease incidence, as our estimates of γ_3 (Fig. 3) reflected the prior distribution assumed for that parameter (uniform from 0 to 1).

Fig. 3: Summary of estimated reporting rates showing the baseline reporting rate or probability of detecting a secondary infection (ρ), the probability of detecting a primary infection (γ_1) relative to a secondary infection, and the probability of detecting a tertiary/quaternary infection (γ_3) relative to a primary infection. Each point represents the posterior median estimate and the error bars show the 95% CrI for each dataset. The box represents the country-specific central estimate calculated by taking the mean values of the MCMC output for each country (the line and limits of the box represents the posterior median and the 95% CrI respectively). A single overall value of γ_1 and γ_3 were estimated per country.

The baseline reporting rates (ρ) varied substantially by country (Fig. 3), likely reflecting differences in healthcare seeking behaviour and surveillance. Generally, estimated reporting rates in the Americas were higher than in South East Asia, with Singapore having the highest rate within SE Asia. Reporting rates also varied within each country depending on survey year or survey region, which may reflect

differences in local healthcare systems or changes in public awareness after epidemics.

We used weighted regression to compare the force of infection estimates obtained from age-stratified seroprevalence data to cumulative incidence data. Estimates obtained from the model fitted to the cumulative incidence data were largely comparable to force of infection estimates from seroprevalence data (Fig. 4). The majority of the total force of infection (λ_{total}) estimates from incidence data (calculated by multiplying the serotype-specific force of infection by the number of serotypes in circulation) were comparable to those obtained from seroprevalence data when λ_{total} was smaller than ~0.1 with greater uncertainty as the force of infection increased. In two of the three locations in Thailand where region and time matching seroprevalence and incidence data were available [34], the force of infection estimates obtained from the models fitted to incidence data and serology data had overlapping 95% credible intervals. In Ratchaburi the estimate obtained from seroprevalence data was smaller than that from incidence data (Fig. 5).

Fig. 4: Comparison of weighted deming regression of force of infection estimates by country from cumulative incidence data and seroprevalence data. Each point is weighted depending on the error in both serology and incidence estimates, represented by the size of circles (larger circles indicating greater weight, i.e. smaller error).

Fig. 5: Posterior median estimates of the total force of infection from the model fitted to incidence data (model 1) and model A (as described in [13]**) to age-stratified seroprevalence data (serology) from Thailand where incidence and serology data were available from the same year and location .**

Discussion

From a literature search we selected 23 papers reporting age-stratified case notification data in 13 countries from 1978-2010. For each dataset we estimated dengue transmission intensity as quantified by the force of infection (λ) and the basic reproduction number (R_0). Where possible we fitted to the cumulative incidence data as fitting to yearly incidence data gave less stable estimates (model fits to yearly incidence data are given in the SI Text) The total force of infection (λ_{total}) estimated from cumulative incidence data were then compared with previous λ estimates from seroprevalence data.

The incidence model presented in this paper provides a method for estimating dengue transmission intensity in areas where seroprevalence data are not available. Force of infection estimates and corresponding basic reproduction numbers varied widely across and within countries as expected, highlighting the heterogeneous nature of dengue transmission spatially and temporally. The majority of our R_0 estimates ranged from 1 to 5, similar to our estimates obtained from seroprevalence data [13]. Similarly to our serology-based estimates, force of infection estimates were generally higher in South East Asia than for Latin America. Since we had no serotype-specific notification data, we assumed that all serotypes were equally transmissible. If serotype-specific notification data were available, serotype-specific forces of infection could be estimated. Although we assumed that dengue transmission intensity does not vary with age, is constant in time and equal for all serotypes in circulation, previous studies have shown that transmissibility can differ substantially not only between serotypes [13,54] but also seasonally, yearly [54], and spatially [55]. However, given the available data it was not possible to estimate serotype-specific or time-varying forces of infection. Multiple cross-sectional surveys or cohort studies are required to estimate how forces of infection have changed by age over time, and serotype-specific data are needed to resolve differences between serotypes.

Due to the lack of incidence and serology data collected in the same year and region, we matched cumulative incidence and serology datasets according to the year or region (see SI Text). While overall estimates from incidence data were comparable

with those derived from seroprevalence data, it would nonetheless be beneficial to validate this model with more incidence and serology datasets collected simultaneously in the same geographical location.

Generally, estimated reporting rates (ρ) in the Americas were higher than those in South East Asia with Singapore having the highest rate within South East Asia, consistent with their well-established dengue surveillance program [56]. Reporting rate estimates also varied within each country depending on survey year or survey region reflecting variation in healthcare and surveillance systems [19]. Reporting rates are also likely to change in response to recent or current epidemics which affect public awareness of dengue and thus healthcare seeking behaviour [57]. Additionally, in an epidemic year clinicians may preferentially diagnose a febrile illness as dengue without laboratory testing [58]. We hypothesised that severity or disease reporting differed by age group and estimated age-dependent reporting rates (ρ_{young} and ρ_{old}) and the age at which reporting rates changed ($A_{threshold}$). However due to the wide age bands of the available data, we were not able to explore this fully. Full details are given in the SI Text.

Since the majority of notified dengue cases are diagnosed as secondary dengue infections [4,5,7,11,12,59], we assumed that the probability of detecting a primary case would be smaller than the probability of detecting a secondary case, and that the probability of detecting a tertiary or quaternary case would be smaller than the probability of detecting a primary case ($\gamma_3 < \gamma_1 < \rho$). The probability of detecting a primary case was consistently low relative to a secondary case (Fig. 3) at less than 50%, the majority being under 25%. However, we were not able to estimate the probability of detecting a tertiary/quaternary case (relative to a primary case) from the available data. A prospective cohort study in Nicaragua found that the proportion of inapparent to symptomatic infection did not differ according to whether an individual had a primary, secondary, or tertiary infection [60].

Overall, the impact of cross-immunity and the contribution of tertiary and quaternary infections to onward transmission are controversial. While there is evidence that tertiary and quaternary infections occur [54,61], there is little quantitative data on the infectiousness or severity of such infections relative to primary and secondary

infections. Additionally, clinically apparent tertiary or quaternary infections are not routinely reported, nor can they be tested for retrospectively [61]. Wikramaratna *et al.* showed that tertiary and quaternary infections allows for the high seroprevalence at very young ages observed in Haiti [62] and Nicaragua [63] better than when assuming complete protection after two heterologous infections [61].

Since the majority of dengue infections are mild or asymptomatic, even sensitive healthcare systems can substantially underestimate true rates of infection even for the supposedly more severe secondary infections, as shown by the low baseline reporting rates [3,11]. Furthermore, dengue has a wide spectrum of clinical manifestations making it difficult to accurately diagnose in the first instance [20]. Our estimates from Thailand (Fig. 5) shows that even with data from the same location and year, it is difficult to make reliable comparisons between estimates obtained from seroprevalence and incidence data. We were also comparing force of infection estimates from seroprevalence data to those from incidence data from a single year, which may have contributed to the observed discrepancy. Although incidence data are the most abundant form of data available on dengue transmission, surveillance systems and reporting procedures are not standardised within or across countries making it very difficult to reliably compare estimates [20]. Laboratory capacity and general public health infrastructure and surveillance systems vary widely and there is often no integration between private and public health sectors. With such variable data, it is very difficult to estimate dengue burden (or transmission intensity) consistently. Since non-serotype specific serological (IgG) surveys are relatively inexpensive to collect, it would be beneficial for such seroprevalence data to be collected routinely. Such data would provide better baseline estimates of overall transmission intensity against which incidence based-estimates could be calibrated to assess changes in transmission and identify weaknesses in surveillance systems.

Acknowledgments

We thank Isabel Rodriguez-Barraquer from Johns Hopkins Bloomberg School of Public Health and Derek Cummings from the University of Florida for generously sharing the incidence data from Rayong, Ratchaburi, and Bangkok.

References

1. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, et al. (2013) The global distribution and burden of dengue. Nature 496: 504–507. doi:10.1038/nature12060.
2. Reich NG, Shrestha S, King AA, Rohani P, Lessler J, et al. (2013) Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity. J R Soc Interface 10: 20130414. doi:10.1098/rsif.2013.0414.
3. Endy TP, Chunsuttiwat S, Nisalak A, Libraty DH, Green S, et al. (2002) Epidemiology of inapparent and symptomatic acute dengue virus infection: A prospective study of primary school children in Kamphaeng Phet, Thailand. Am J Epidemiol 156: 40–51. doi:10.1093/aje/kwf005.
4. Halstead SB (2007) Dengue. Lancet 370: 1644–1652.
5. Simmons CP, Farrar JJ, Nguyen v V, Wills B (2012) Dengue. N Engl J Med 366: 1423–1432. doi:10.1056/NEJMra1110265.
6. Burke DS, Nisalak A, Johnson DE, Scott RM (1988) A prospective study of dengue infections in Bangkok. Am J Trop Med Hyg 38: 172–180.
7. Guzman MG, Kouri G, Valdes L, Bravo J, Alvarez M, et al. (2000) Epidemiologic studies on dengue in Santiago de Cuba, 1997. Am J Epidemiol 152: 793–799. doi:10.1093/aje/152.9.793.
8. Halstead SB, Scanlon JE, Umpaivit P, Udomsakdi S (1969) Dengue and chikungunya virus infection in man in Thailand, 1962-1964. IV. Epidemiologic studies in the Bangkok metropolitan area. Am J Trop Med Hyg 18: 997–1021.
9. Halstead SB (1988) Pathogenesis of dengue: challenges to molecular biology. Science (80-) 239: 476–481.
10. Halstead SB, Orourke EJ (1977) DENGUE VIRUSES AND MONONUCLEAR PHAGOCYTES .1. INFECTION ENHANCEMENT BY NON-NEUTRALIZING ANTIBODY. J Exp Med 146: 201–217. doi:10.1084/jem.146.1.201.
11. Gordon A, Kuan G, Mercado JC, Gresh L, Avilés W, et al. (2013) The Nicaraguan Pediatric Dengue Cohort Study: Incidence of Inapparent and Symptomatic Dengue Virus Infections, 2004–2010. PLoS Negl Trop Dis 7: e2462. doi:10.1371/journal.pntd.0002462.
12. Sharp TM, Hunsperger E, Santiago GA, Munoz-Jordan JL, Santiago LM, et al. (2013) Virus-specific differences in rates of disease during the 2010 Dengue epidemic in Puerto Rico. PLoS Negl Trop Dis 7: e2159. doi:10.1371/journal.pntd.0002159.
13. Imai N, Dorigatti I, Cauchemez S, Ferguson NM (2015) Estimating Dengue Transmission Intensity from Sero-Prevalence Surveys in Multiple Countries. PLoS Negl Trop Dis 9: e0003719. doi:10.1371/journal.pntd.0003719.
14. Guo R, Lin J, Li L-H, Ke C-W, He J, et al. (2014) The Prevalence and Endemic Nature of Dengue Infections in Guangdong, South China: An Epidemiological, Serological, and Etiological Study from 2005-2011. PLoS One 9: e85596. doi:10.1371/journal.pone.0085596.
15. George R, Lam SK (1997) Dengue virus infection: The Malaysian experience. Ann Acad Med Singapore 26: 815–819.

16. de Lima VLC, Figueiredo LTM, Correa HR, Leite OF, Rangel O, et al. (1999) Dengue fever: a post-epidemic sero-epidemiological survey in an urban setting at a northwestern county of S. Paulo State - Brazil. Rev Saude Publica 33: 566–574.

17. Li ZJ, Yin WW, Clements A, Williams G, Lai SJ, et al. (2012) Spatiotemporal analysis of indigenous and imported dengue fever cases in Guangdong province, China. BMC Infect Dis 12. doi:10.1186/1471-2334-12-132.

18. Nandi J, Sharma RS, Dasgupta RK, Katyal R, Dutta PK, et al. (2009) Epidemiological analysis of hospitalized cases of dengue fever/dengue haemorrhagic fever and extent of breeding of Aedes aegypti in major hospitals in the National Capital Territory of Delhi (NCT Delhi), 2005-2009. Dengue Bull 33: 130–139.

19. Ruberto I, Marques E, Burke DS, Van Panhuis WG (2015) The Availability and Consistency of Dengue Surveillance Data Provided Online by the World Health Organization. PLoS Negl Trop Dis 9: e0003511. doi:10.1371/journal.pntd.0003511.

20. Shepard DS, Undurraga E a., Betancourt-Cravioto M, Guzmán MG, Halstead SB, et al. (2014) Approaches to Refining Estimates of Global Burden and Economics of Dengue. PLoS Negl Trop Dis 8: e3306. doi:10.1371/journal.pntd.0003306.

21. Undurraga EA, Halasa YA, Shepard DS (2013) Use of Expansion Factors to Estimate the Burden of Dengue in Southeast Asia: A Systematic Analysis. PLoS Negl Trop Dis 7: 15. doi:10.1371/journal.pntd.0002056.

22. Beaute J, Vong S, Beauté J (2010) Cost and disease burden of dengue in Cambodia. BMC Public Health 10: (31 August 2010) – (31 August 2010). doi:10.1186/1471-2458-10-521.

23. Shepard DS, Undurraga EA, Halasa YA (2013) Economic and Disease Burden of Dengue in Southeast Asia. PLoS Negl Trop Dis 7. doi:10.1371/journal.pntd.0002055.

24. Suaya JA, Shepard DS, Siqueira JB, Martelli CT, Lum LCS, et al. (2009) Cost of Dengue Cases in Eight Countries in the Americas and Asia: A Prospective Study. Am J Trop Med Hyg 80: 846–855.

25. Toan NT, Rossi S, Prisco G, Nante N, Viviani S (2015) Dengue epidemiology in selected endemic countries: factors influencing expansion factors as estimates of underreporting. Trop Med Int Heal 00: n/a – n/a. doi:10.1111/tmi.12498.

26. Guzman MG, Kouri G, Valdes L, Bravo JJ, Vazquez S, et al. (2002) Enhanced severity of secondary dengue-2 infections: death rates in 1981 and 1997 Cuban outbreaks. Rev Panam Salud Publica 11: 223–227. doi:10.1590/s1020-49892002000400003.

27. Sangkawibha N, Rojanasuphot S, Ahandrik S, Viriyapongse S, Jatanasen S, et al. (1984) RISK-FACTORS IN DENGUE SHOCK SYNDROME - A PROSPECTIVE EPIDEMIOLOGIC-STUDY IN RAYONG, THAILAND .1. THE 1980 OUTBREAK. Am J Epidemiol 120: 653–669.

28. Anantapreecha S, Chanama S, A-nuegoonpipat a, Naemkhunthot S, Sa-Ngasang a, et al. (2005) Serological and virological features of dengue fever

and dengue haemorrhagic fever in Thailand from 1999 to 2002. Epidemiol Infect 133: 503–507. doi:10.1017/s0950268804003541.

29. United Nations Population Division D of E and SA (2013) World Population Prospects: The 2012 Revision.

30. R Core Team, R Development Core Team R (2012) R: A language and environment for statistical computing. R Found Stat Comput 1: 409. doi:10.1007/978-3-540-74686-7.

31. Ripley BD, Thompson M (1987) Regression techniques for the detection of analytical bias. Analyst 112: 377. doi:10.1039/an9871200377.

32. Therneau T (2014) deming: Deming, Thiel-Sen and Passing-Bablock Regression.

33. Luo L, Liang H, Hu Y, Liu W, Wang Y, et al. (2012) Epidemiological, virological, and entomological characteristics of dengue from 1978 to 2009 in Guangzhou, China. J Vector Ecol 37: 230–240. doi:10.1111/j.1948-7134.2012.00221.x.

34. Bureau of Epidemiology, Ministry of Public Health of Thailand (n.d.) Bureau of Epidemiology, Ministry of Public Health of Thailand (YEAR) Annual epidemiological surveillance report. Nonthaburi, Thailand.

35. Cordeiro MTT, Schatzmayr HGG, Nogueira RMR, De Oliveira VF, De Melo WT, et al. (2007) Dengue and dengue hemorrhagic fever in the State of Pernambuco, 1995-2006. Rev Soc Bras Med Trop 40: 605–611. doi:10.1590/s0037-86822007000600001.

36. Cardoso IM, Areias Cabidelle A de S, Borges P de C e L, Lang CF, Calenti FG, et al. (2011) Dengue: clinical forms and risk groups in a high incidence city in the Southeastern region of Brazil. Rev Soc Bras Med Trop 44: 430–435.

37. Penna G, Pinto LF, Soranz D, Glatt R (2009) High incidence of diseases endemic to the Amazon region of Brazil, 2001-2006. Emerg Infect Dis 15: 626–632. doi:10.3201/eid1504.081329.

38. Vong S, Khieu V, Glass O, Ly S, Duong V, et al. (2010) Dengue incidence in urban and rural Cambodia: results from population-based active fever surveillance, 2006-2008. PLoS Negl Trop Dis 4: e903. doi:10.1371/journal.pntd.0000903.

39. Wichmann O, Yoon I-K, Vong S, Limkittikul K, Gibbons R V, et al. (2011) Dengue in Thailand and Cambodia: an assessment of the degree of underrecognized disease burden based on reported cases. PLoS Negl Trop Dis 5: e996. doi:10.1371/journal.pntd.0000996.

40. Anker M, Arima Y (2011) Male–female differences in the number of reported incident dengue fever cases in six Asian countries. West Pacific Surveill Response … 2: 1–7. doi:10.5365/wpsar.2011.2.1.002.

41. Prasith N, Keosavanh O, Phengxay M (2011) Assessment of gender distribution in dengue surveillance data, the Lao People's Democratic Republic. West Pacific Surveill Response J 4: 1–9. doi:10.5365/wpsar.2012.3.4.020.

42. Khampapongpane B, Lewis HC, Ketmayoon P, Phonekeo D, Somoulay V, et al. (2014) National dengue surveillance in the Lao People's Democratic Republic, 2006–2012: epidemiological and laboratory findings. West Pacific Surveill Response J 5: 2006–2012. doi:10.5365/wpsar.2014.5.1.001.

43. Hammond SN, Balmaseda A, Perez L, Tellez Y, Saborio SI, et al. (2005) Differences in dengue severity in infants, children, and adults in a 3-year hospital-based study in Nicaragua. Am J Trop Med Hyg 73: 1063–1070.

44. Ramos MM, Argueello DF, Luxemburger C, Quinones L, Munoz JL, et al. (2008) Epidemiological and clinical observations on patients with dengue in Puerto Rico: Results from the first year of enhanced surveillance - June 2005-May 2006. Am J Trop Med Hyg 79: 123–127.

45. Tomashek KM, Rivera A, Munoz-Jordan JL, Hunsperger E, Santiago L, et al. (2009) Description of a Large Island-Wide Outbreak of Dengue in Puerto Rico, 2007. Am J Trop Med Hyg 81: 467–474.

46. Rigau-Pérez JG, Ayala-López A, Vorndam AV, Clark GG, Rigau-Perez JG, et al. (2001) Dengue activity in Puerto Rico during an interepidemic period (1995-1997). Am J Trop Med Hyg 64: 75–83.

47. Koh BKW, Lee CN, Kita Y, Choon ST, Li WA, et al. (2008) The 2005 dengue epidemic in Singapore: Epidemiology, prevention and control. Ann Acad Med Singapore 37: 538–545.

48. Ler TS, Ang LW, Yap GSL, Ng LC, Tai JC, et al. (2011) WPRO | Epidemiological characteristics of the 2005 and 2007 dengue epidemics in Singapore – similarities and distinctions. West Pacific Surveill Response J 2: 24–29.

49. Kulatilaka TA, Jayakuru WS (1998) Control of dengue/dengue haemorrhagic fever in Sri Lanka. Dengue Bull 22: 53–59.

50. Lin C-H, Schiøler KL, Jepsen MR, Ho C-K, Li S-H, et al. (2012) Dengue outbreaks in high-income area, Kaohsiung City, Taiwan, 2003-2009. Emerg Infect Dis 18: 1603–1611. doi:10.3201/eid1810.111929.

51. Limkittikul K, Brett J, L'Azou M (2014) Epidemiological Trends of Dengue Disease in Thailand (2000–2011): A Systematic Literature Review. PLoS Negl Trop Dis 8: e3241. doi:10.1371/journal.pntd.0003241.

52. Cuong HQ, Hien NT, Duong TN, Phong T V, Cam NN, et al. (2011) Quantifying the Emergence of Dengue in Hanoi, Vietnam: 1998-2009. PLoS Negl Trop Dis 5: e1322–e1322. doi:10.1371/journal.pntd.0001322.

53. Ghouth AS, Amarasinghe A, Letson GW, Bin Ghouth AS (2012) Dengue outbreak in Hadramout, Yemen, 2010: an epidemiological perspective. Am J Trop Med Hyg 86: 1072–1076. doi:10.4269/ajtmh.2012.11-0723.

54. Reiner RC, Stoddard ST, Forshey BM, King A a, Ellis AM, et al. (2014) Time-varying, serotype-specific force of infection of dengue virus. Proc Natl Acad Sci U S A. doi:10.1073/pnas.1314933111.

55. Mammen MP, Pimgate C, Koenraadt CJM, Rothman AL, Aldstadt J, et al. (2008) Spatial and Temporal Clustering of Dengue Virus Transmission in Thai Villages. PLoS Med 5: 1605–1616. doi:10.1371/journal.pmed.0050205.

56. Beatty ME, Stone A, Fitzsimons DW, Hanna JN, Lam SK, et al. (2010) Best Practices in Dengue Surveillance: A Report from the Asia-Pacific and Americas Dengue Prevention Boards. PLoS Negl Trop Dis 4. doi:10.1371/journal.pntd.0000890.

57. Wong LP, AbuBakar S (2013) Health Beliefs and Practices Related to Dengue

Fever: A Focus Group Study. PLoS Negl Trop Dis 7. doi:10.1371/journal.pntd.0002310.

58. World Health Organization (2007) Scientific working group report on dengue. Geneva, Switz World Heal Organ.

59. Guzman MG, Kouri G (2002) Dengue: an update. Lancet Infect Dis 2: 33–42. doi:10.1016/s1473-3099(01)00171-2.

60. Montoya M, Gresh L, Mercado JC, Williams KL, Vargas MJ, et al. (2013) Symptomatic Versus Inapparent Outcome in Repeat Dengue Virus Infections Is Influenced by the Time Interval between Infections and Study Year. PLoS Negl Trop Dis 7. doi:10.1371/journal.pntd.0002357.

61. Wikramaratna PS, Simmons CP, Gupta S, Recker M (2010) The Effects of Tertiary and Quaternary Infections on the Epidemiology of Dengue. PLoS One 5. doi:10.1371/journal.pone.0012347.

62. Halstead SB, Streit TG, Lafontant JG, Putvatana R, Russell K, et al. (2001) Haiti: Absence of dengue hemorrhagic fever despite hyperendemic dengue virus transmission. Am J Trop Med Hyg 65: 180–183.

63. Balmaseda A, Standish K, Mercado JC, Matute JC, Tellez Y, et al. (2010) Trends in patterns of dengue transmission over 4 years in a pediatric cohort study in Nicaragua. J Infect Dis 201: 5–14. doi:10.1086/648592.

Striking Image Caption

A female *Aedes aegypti* taking a blood meal. Photo Credit: James Gathany. Content Provider: CDC/ Prof. Frank Hadley Collins, Dir., Cntr. for Global Health and Infectious Diseases, Univ. of Notre Dame

Supporting Information Captions

S1 Text: Supporting information file containing methods, results, and extra figures.