Modelling the effective reproduction number of vector-borne diseases: the yellow fever outbreak in Luanda, Angola 2015–2016 as an example

Shi Zhao[1,2,3], Salihu S. Musa[4], Jay T. Hebert[4], Peihua Cao[5], Jinjun Ran[6], Jiayi Meng[7], Daihai He[4] and Jing Qin[1]

[1] School of Nursing, Hong Kong Polytechnic University, Hong Kong, China
[2] Division of Biostatistics, JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China
[3] Clinical Trials and Biostatistics Lab, Shenzhen Research Institute, Chinese University of Hong Kong, Shenzhen, China
[4] Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, China
[5] Department of Hepatobiliary Surgery II, Zhujiang Hospital, Southern Medical University, Guangzhou, China
[6] School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China
[7] School of Economics and Finance, Xi'an International Studies University, Xi'an, China

Corresponding author
Daihai He, daihai.he@polyu.edu.hk

ABSTRACT

The burden of vector-borne diseases (Dengue, Zika virus, yellow fever, etc.) gradually increased in the past decade across the globe. Mathematical modelling on infectious diseases helps to study the transmission dynamics of the pathogens. Theoretically, the diseases can be controlled and eventually eradicated by maintaining the effective reproduction number, ($\mathcal{R}_{\text{eff}}$), strictly less than 1. We established a vector-host compartmental model, and derived ($\mathcal{R}_{\text{eff}}$) for vector-borne diseases. The analytic form of the ($\mathcal{R}_{\text{eff}}$) was found to be the product of the basic reproduction number and the geometric average of the susceptibilities of the host and vector populations. The ($\mathcal{R}_{\text{eff}}$) formula was demonstrated to be consistent with the estimates of the 2015–2016 yellow fever outbreak in Luanda, and distinguished the second minor epidemic wave. For those using the compartmental model to study the vector-borne infectious disease epidemics, we further remark that it is important to be aware of whether one or two generations is considered for the transition "from host to vector to host" in reproduction number calculation.

INTRODUCTION

Vector-borne disease epidemics pose a serious threat to global health. Especially in tropical and sub-tropical regions, most vector-borne diseases are treated as a part of neglected tropical diseases (NTDs) (*Fenwick, 2012*), present endemic features, and are persistent in the interface of host and vector communities. During 2014, the historical large-scale dengue fever caused an extensive international epidemic in southern China as well as

other regions in Southeastern Asia (*Sang et al., 2015*; *Stanaway et al., 2016*). The Zika virus (ZIKV) emerged in the Pacific area in 2008 (*Duffy et al., 2009*; *Besnard et al., 2014*; *Cauchemez et al., 2016*), then in South America since 2015, and caused more than 500,000 (confirmed or probable) cases, while the true number of cases remains unclear (*He et al., 2017*; *Gao et al., 2016*; *Johansson et al., 2016*; *Ferguson et al., 2016*; *He et al., 2019*). The West Nile virus (WNV) was widespread across tropical parts globally, and was introduced into North America in 1999, which lead to an approximated 1.8 million infections from 1999 to 2010 (*Kilpatrick, 2011*). The chikungunya virus (CHIKV) hit the Americas and beyond, where tens of millions of previously unexposed persons would be at risk (*Fischer & Staples, 2014*; *Weaver & Lecuit, 2015*). In 2015–2016, the largest yellow fever (YF) outbreak (since the 1980s) occurred in Angola and the Democratic Republic of the Congo (DRC) (*Zhao et al., 2018b*; *Kraemer et al., 2017*; *Wu et al., 2016*; *Shearer et al., 2017*). After Africa, the YF continuously posed a serious threat to the unprotected population of southern Brazil, which was believed to have eradicated YF after the middle of the last century (*Barrett, 2018*). Malaria epidemics occur from time to time in many under-developed places in the tropical and sub-tropical regions (*Caminade et al., 2014*; *Murray et al., 2012*). The increasing frequency of such outbreaks over the past decades urges disease control and prevention studies (*Johansson et al., 2012*; *Kraemer et al., 2019*).

Mathematical modelling on infectious diseases are developed, and help to study the transmission dynamics of the pathogens from a theoretical point of view (*Earn et al., 2008*; *Brauer & Castillo-Chavez, 2001*; *Keeling & Rohani, 2011*; *Grenfell, Dobson & Moffatt, 1995*). Theoretically, the infectious diseases can be controlled and finally eradicated by maintaining the effective reproduction number, $\mathcal{R}_{eff}$, strictly less than 1, i.e., $\mathcal{R}_{eff} < 1$. The $\mathcal{R}_{eff}$ is the expected number of secondary cases produced by one typical infection joining in a population during its infectious period (*Van den Driessche & Watmough, 2002*). This theoretical criterion is widely adopted as a threshold to characterise the transmission dynamics and measure the disease control effectiveness (*Earn et al., 2008*; *Keeling & Rohani, 2011*). For airborne communicable diseases, e.g., respiratory diseases and most childhood infections, that transmit in the vector-free context, $\mathcal{R}_{eff}$ is given as in Eq. (1).

$$\mathcal{R}_{eff} = \mathcal{R}_0 \mathcal{S}_h, \tag{1}$$

where the $\mathcal{R}_0$ is the basic reproduction number, and the $\mathcal{S}_h$ is the susceptibility in the human population. The $\mathcal{R}_0$ is defined as the expected number of secondary cases produced by one typical infection joining in a completely susceptible population during its infectious period (*Heffernan, Smith & Wahl, 2005*). This formula in Eq. (1) has been well-studied and wildly used to quantify the transmissibility of infectious diseases (*Earn et al., 2008*; *Brauer & Castillo-Chavez, 2001*; *Keeling & Rohani, 2011*; *Grenfell, Dobson & Moffatt, 1995*).

We used the recent yellow fever outbreak in Luanda, Angola from 2015 to 2016 as an example to implement the $\mathcal{R}_{eff}$, and demonstrate the difference(s) between different forms of the $\mathcal{R}_{eff}$s of airborne diseases and vector-borne diseases. The YF outbreak included 941 reported cases with 73 deaths in Luanda, the capital city of Angola from December 2015 to June 2016 (*Zhao et al., 2018b*; *WHO, 2017*). The local authority had conducted large-scale

mass vaccination campaign since February 2016, and the vaccine program immunised approximated 55% of the local population within 6 months since started (*Zhao et al., 2018b*; *WHO, 2017*). Owing to the timely and large-scale mass vaccination campaign, it was estimated that over 5-fold of both cases and deaths were saved. The YF cases time series in Luanda and the local vaccination coverage were obtained from the situation reports released by the African Health Observatory (AHO) (*WHO, 2017*).

In this work, we establish a simple (and classic) host-vector compartmental model and derive the effective reproduction number for the transmission dynamics of the vector-borne diseases. We further explore the relationship between the disease control in terms of $\mathcal{R}_{\text{eff}}$ and the control efforts of the basic reproduction number ($\mathcal{R}_0$) and population susceptibility. The yellow fever (YF) epidemic in Luanda, Angola from December 2015 to June 2016 is studied as an illustrative example to compare the $\mathcal{R}_{\text{eff}}$ estimates in different formulations or approaches.

METHODS

Many vector-borne diseases, such as dengue fever, yellow fever, Zika fever, malaria, etc., are transmitted from vector to host as well as from host to vector. Following previous literature (*Earn et al., 2008*; *Brauer & Castillo-Chavez, 2001*; *Keeling & Rohani, 2011*; *Grenfell, Dobson & Moffatt, 1995*), the transmission mechanism can be explained by the vector-host epidemic models based on the differential equations, i.e., compartmental models.

Epidemic model for vector-borne diseases

We adopted the classic "susceptible-infected-removed" (SIR) structural framework to model both host and vector population dynamics (*Gao et al., 2016*; *Tang et al., 2016*; *Saad-Roy, Van den Driessche & Ma, 2016*; *Zhao et al., 2018b*; *Earn et al., 2008*; *Brauer et al., 2016*). We use S_h, I_h, R_h to denote the numbers of susceptible, infected and removed host population respectively. The S_v, I_v, R_v denote the numbers of susceptible, infected and removed vector population respectively. The susceptible host becomes infected by the "contact" with infectious vectors, eventually recovers, i.e., move to the recovered class, and remains protected from secondary infection. A similar path is also modelled in the vectors' population. For simplicity, we ignored the incubation period, commonly denoted by E, of the infection in the epidemic model. Based on the above descriptions, the compartmental model is formulated in Eqs. (2). Figure 1 shows the schematic diagram of model (2).

$$
\begin{cases}
S_h' = \mu_h N_h - \beta_{vh} \cdot \dfrac{S_h}{N_h} I_v - \mu_h S_h, \\[2mm]
I_h' = \beta_{vh} \cdot \dfrac{S_h}{N_h} I_v - (\gamma_h + \mu_h) I_h, \\[2mm]
R_h' = \gamma_h I_h - \mu_h R_h, \\[2mm]
S_v' = B_v(t) - \beta_{hv} S_v \cdot \dfrac{I_h}{N_h} - \mu_v S_v, \\[2mm]
I_v' = \beta_{hv} S_v \cdot \dfrac{I_h}{N_h} - (\gamma_v + \mu_v) I_v, \\[2mm]
R_v' = \gamma_v I_v - \mu_v R_v.
\end{cases}
\tag{2}
$$
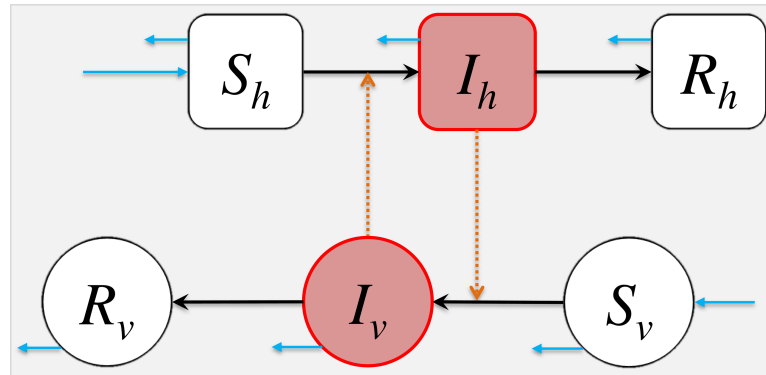
Figure 1 The schematic diagram of model (2). The S_h, I_h, R_h represent the numbers of susceptible, infected and removed host population. The S_v, I_v, R_v represent the numbers of susceptible, infected and removed vector population. The black arrows represent the infection status transition paths. The orange dashed arrows represent disease transmission paths, and the light blue arrows represent the natural birth or death of hosts or vectors. Square compartments represent the host classes (or compartments), and circular compartments represent the vector classes. The red compartments represent infected classes.

Table 1 Summary table of the parameters in model (2).

Parameter	Notation	Unit/Remark
Transmission rate from vector to host	β_{vh}	Host per vector ·time
Transmission rate from host to vector	β_{hv}	Per time
Host's disease-induced removing rate	γ_h	Per time
Vector's disease-induced removing rate	γ_v	Per time
Vector's natural recruiting rate	$B_v(t)$	Vector per time
Vector's natural death rate	μ_v	Per time
Host's natural birth/death rate	μ_h	Per time
Vector to host ratio	$m = N_v/N_h$	Vector per host
Host's susceptibility	$\mathcal{S}_h = S_h/N_h$	Unit-free
Vector's susceptibility	$\mathcal{S}_v = S_v/N_v$	Unit-free

The model parameters are summarised in Table 1, and all parameters are assumed non-negative.

The N_h and N_v are the total numbers of the hosts and vectors respectively. We have

$$N_h = S_h(t) + I_h(t) + R_h(t) = \text{constant; and}$$

$$N_v(t) = S_v(t) + I_v(t) + R_v(t) \neq \text{constant.}$$

In our model, $N_v = N_v(t)$ is time-dependent in a manner that is controlled by the mosquito birth rate $B_v(t)$, namely $N_v' = B_v(t) - \mu_v N_v$, which may not be necessarily equal to zero. To guarantee the biological reasonability, the values of all six classes should be non-negative, i.e., $\geqslant 0$. Then, $S_h, I_h, R_h \leqslant N_h$, and $S_v, I_v, R_v \leqslant N_v$. Although the vector-borne diseases could affect the vector's lifespan in some occasions, e.g., tick-borne diseases

(*Niebylski, Peacock & Schwan, 1999*), it is conventionally assumed the infected vectors can neither recover nor cause direct mortality in modelling studies, which is true for most of the mosquito-borne diseases, the most common group of vector-borne diseases. For simplicity, the disease-caused vector mortality is neglected in this modelling study. The inclusion of R_v in the model (2) seems unnatural, nevertheless letting $\gamma_v = 0$, thus $R_v = 0$, can resolve this conflict straightforwardly.

Most of the pathogens of the vector-borne diseases merely transmit via two paths, host-to-vector and vector-to-host, e.g., Dengue fever, yellow fever, Chikungunya fever and malaria, etc. For those can be transmitted via host-to-host path, commonly by sexual contact or blood transfusion, e.g., Zika fever, the host-to-host transmission is minor, and dominated by the aforementioned two paths. Therefore, we remark that although the model (2) takes a simple form, it is applicable to model the transmission dynamics of most vector-borne diseases. Model (2) also can be easily extended into more complex form, for instance the model system in *Zhao et al. (2018b)*, *Gao et al. (2016)* and *Tang et al. (2016)*. As long as the transmission paths remain from-host-to-vector and from-vector-to-host, the relationship between $\mathcal{R}_0$ and $\mathcal{R}_{\text{eff}}$ explored in this paper still holds.

Basic reproduction number

The basic reproduction number, $\mathcal{R}_0$, is the expected number of secondary cases produced by one typical infection joining in a **completely susceptible** population during its infectious period (*Heffernan, Smith & Wahl, 2005*). When $\mathcal{R}_0 < 1$, the disease would die out in the long run. While if $\mathcal{R}_0 > 1$, the disease would spread among the population and may cause a pandemic.

In epidemiology, the next generation matrix is a method used to derive the basic reproduction number, $\mathcal{R}_0$, for a compartmental model of the spread of infectious diseases. According to *Van den Driessche & Watmough (2002)*, *Diekmann, Heesterbeek & Roberts (2009)* and *Diekmann, Heesterbeek & Metz (1990)*, a systematic procedure to calculate the $\mathcal{R}_0$ by solving the dominant eigenvalue, i.e., the eigenvalue with the largest real part, of the next generation matrix, $\mathbf{G}$, at the disease-free equilibrium (DFE). Mathematically, it can be shown easily that the DFE exists and is stable for our epidemic model (2). For the next generation matrix $\mathbf{G} = \mathbf{FV}^{-1}$, the $\mathbf{F}$ is the new infection (or transmission) matrix and $\mathbf{V}$ is the infection transfer (or transition) matrix. The entry of ith row and jth column of matrix $\mathbf{F}$ is denoted by $F_{i,j}$, and $F_{i,j} = \frac{\partial \mathcal{F}_i}{\partial x_j}$ where $\mathcal{F}_i$ is the ith equation of $\mathcal{F}$ and x_j is the jth variable of the vector of infected classes. For instance, in Eqs. (2), the I_h and I_v are the infected classes. Similarly, the entry of ith row and jth column of matrix $\mathbf{V}$ is denoted by $V_{i,j}$, and $V_{i,j} = \frac{\partial \mathcal{V}_i}{\partial x_j}$ where $\mathcal{V}_i$ is the ith equation of $\mathcal{V}$ and x_j is the jth variable of the vector of infected classes. The vector $\mathcal{F}$ is the transmission rates' vector quantity, i.e., the changing rates from infected to non-infected classes, and vector $\mathcal{V}$ is the transition rates' vector quantity, i.e., the changing rates among infected classes. The $\mathbf{F}$ is the Jacobian matrix of $\mathcal{F}$, and $\mathbf{V}$ is the Jacobian matrix of $\mathcal{V}$.

For the compartmental model in Eqs. (2), the I_h and I_v are the infected classes, which should be included in the vectors ($\mathcal{F}$ and $\mathcal{V}$) of infected classes. Then, we have

$$\mathcal{F} = \begin{pmatrix} \beta_{vh} \cdot \frac{S_h}{N_h} I_v \\ \beta_{hv} S_v \cdot \frac{I_h}{N_h} \end{pmatrix}, \text{ and } \mathcal{V} = \begin{pmatrix} (\gamma_h + \mu_h) I_h \\ (\gamma_v + \mu_v) I_v \end{pmatrix}. \text{ Hence, } \mathbf{F} = \begin{pmatrix} 0 & \beta_{vh} \\ \beta_{hv} \cdot \frac{N_v}{N_h} & 0 \end{pmatrix}, \text{ and } \mathbf{V} = \begin{pmatrix} \gamma_h + \mu_h & 0 \\ 0 & \gamma_v + \mu_v \end{pmatrix}.$$

The next generation matrix, $\mathbf{G}$, is given as follows.

$$\mathbf{G} = \mathbf{F}\mathbf{V}^{-1} = \begin{pmatrix} 0 & \dfrac{\beta_{vh}}{\gamma_v + \mu_v} \\ m \cdot \dfrac{\beta_{hv}}{\gamma_h + \mu_h} & 0 \end{pmatrix},$$

where the term m is the vector to host ratio, which is defined by $m = \frac{N_v}{N_h}$. By solving the dominant eigenvalue of $\mathbf{G}$ (*Van den Driessche & Watmough, 2002*; *Diekmann, Heesterbeek & Roberts, 2009*; *Cushing & Diekmann, 2016*; *Cushing & Diekmann, 2016*), we derive the $\mathcal{R}_0$ of the model (2) in Eq. (3).

$$\mathcal{R}_0 = \sqrt{\left(\frac{m\beta_{hv}}{\gamma_h + \mu_h}\right) \cdot \left(\frac{\beta_{vh}}{\gamma_v + \mu_v}\right)} = \sqrt{m \cdot \frac{\beta_{hv}\beta_{vh}}{(\gamma_h + \mu_h)(\gamma_v + \mu_v)}}. \tag{3}$$

In Eq. (3), the first ratio under the square root, i.e., $\frac{m\beta_{hv}}{\gamma_h+\mu_h}$, represents the number of vector infections caused by one infected host, and the second, i.e., $\frac{\beta_{vh}}{\gamma_v+\mu_v}$, represents the number of host infections caused by one infected vector. The square root represents the geometric mean that takes the average number of secondary host (or vector) infections produced by a single infected host (or vector) (*Van den Driessche, 2017*).

Effective reproduction number

During an epidemic, the susceptible individuals (S_h or S_v) are gradually consumed, become infected and finally removed from the disease transmission cycle. The effective reproduction number, $\mathcal{R}_{\text{eff}}$, is the expected number of secondary cases produced by one typical infection joining in a population during its infectious period (*Van den Driessche & Watmough, 2002*). The $\mathcal{R}_{\text{eff}}$ is time-varying, which is denoted by $\mathcal{R}_{\text{eff}}(t)$ and sometimes $\mathcal{R}_t$ for the discretized situation (*Ali, Kadi & Ferguson, 2013*; *Fraser, 2007*; *Cori et al., 2013*), and $\mathcal{R}_{\text{eff}}$ quantifies the instantaneous transmissibility of the disease. By applying the approach in 'Basic Reproduction Number', the next generation matrix, $\mathbf{G}$, is given by

$$\mathbf{G} = \mathbf{F}\mathbf{V}^{-1} = \begin{pmatrix} 0 & \beta_{vh}\dfrac{S_h}{N_h} \\ \beta_{hv} \cdot \dfrac{S_v}{N_h} & 0 \end{pmatrix} \times \begin{pmatrix} (\gamma_h + \mu_h)^{-1} & 0 \\ 0 & (\gamma_v + \mu_v)^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \dfrac{S_h}{N_h} \cdot \dfrac{\beta_{vh}}{\gamma_v + \mu_v} \\ m \cdot \dfrac{S_v}{N_v} \cdot \dfrac{\beta_{hv}}{\gamma_h + \mu_h} & 0 \end{pmatrix}.$$

Then, we derive $\mathcal{R}_{\text{eff}}$ from $\mathbf{G}$.

$$\mathcal{R}_{\text{eff}} = \sqrt{m \cdot \frac{\beta_{hv}\beta_{vh}}{(\gamma_h + \mu_h)(\gamma_v + \mu_v)} \cdot \frac{S_h}{N_h} \cdot \frac{S_v}{N_v}} = \mathcal{R}_0 \sqrt{\frac{S_h}{N_h} \cdot \frac{S_v}{N_v}},$$

where $\mathcal{R}_0$ is given in Eq. (3). We further define the susceptibilities of hosts ($\mathcal{S}_h$) and vectors ($\mathcal{S}_v$) by

$$\mathcal{S}_h = \frac{S_h}{N_h}, \text{and} \quad \mathcal{S}_v = \frac{S_v}{N_v}. \tag{4}$$

In the epidemic model (2), we have $S_h, I_h, R_h \leqslant N_h$, and $S_v, I_v, R_v \leqslant N_v$. Therefore, $0 \leqslant \mathcal{S}_h, \mathcal{S}_v \leqslant 1$. Henceforth, the effective reproduction number is in Eq. (5).

$$\mathcal{R}_{\text{eff}} = \mathcal{R}_0 \sqrt{\mathcal{S}_h \mathcal{S}_v}. \tag{5}$$

Since $0 \leqslant \mathcal{S}_h, \mathcal{S}_v \leqslant 1$, we have $\mathcal{R}_{\text{eff}} \leqslant \mathcal{R}_0$. Furthermore, (i) for most of the vector-borne diseases, the vector's lifespan is much shorter than host's lifespan, e.g., mosquito's lifespan is around 7 to 60 days and tick's lifespan is around 3 months to 3 years, and (ii) the infected vectors do not recover. These two facts will lead to an outcome in model (2) that class $R_v = 0$ and class I_v is extremely small and almost zero. Therefore, for simplicity, we set $\mathcal{S}_v = 100\%$ for the remaining parts of this study.

Importantly, the relationship between $\mathcal{R}_0$ and $\mathcal{R}_{\text{eff}}$ in Eq. (5) holds whenever the transmission path of the vector-borne pathogens is from a host to a host via a vector.

An example of the yellow fever epidemic in Luanda 2015–2016
To illustrate the relationship between $\mathcal{R}_0$ and $\mathcal{R}_{\text{eff}}$ in Eq. (5), we compared different calculations or estimations of the effective reproduction number. We adopted the recent yellow fever (YF) outbreak in Luanda, Angola from 2015 to 2016 as a case study (*Zhao et al., 2018b*). The YF cases time series in Luanda and the local vaccination coverage were obtained from the situation reports released by the African Health Observatory (AHO) (*WHO, 2017*). Similar to the WHO (*WHO, 2017*) and previous literature (*Kraemer et al., 2017*; *Zhao et al., 2018b*), both probable and confirmed cases were grouped together, and were considered as the "YF cases" for further analyses (Fig. 2A).

We intend to demonstrate that the relationship between $\mathcal{R}_0$ and $\mathcal{R}_{\text{eff}}$ in Eq. (5) is more proper for pinpoint the epidemic control threshold of the vector-borne diseases than that in Eq. (1). To do so, we treat the instantaneous reproduction number, $\mathcal{R}_t$, calculated by the serial interval (SI) approach as the true effective reproduction number. The calculation of $\mathcal{R}_t$ is introduced in detail in 'Instantaneously Reproduction Number Estimation by Renewable Equation'. After we find $\mathcal{R}_t$ series, we compare the calculations of the $\mathcal{R}_{\text{eff}}$s by using Eq. (1) or Eq. (5) to the $\mathcal{R}_{\text{eff}}$. The focus of this part is to compare the two relationships between $\mathcal{R}_0$ and $\mathcal{R}_{\text{eff}}$ described in Eqs. (1) and (5), and to identify which one is more proper for pinpoint the epidemic control threshold of the vector-borne diseases.

Note that although the formulation of $\mathcal{R}_0$ in model (2) is different from that in the YF epidemic model in *Zhao et al. (2018b)*, the relationship between $\mathcal{R}_0$ and $\mathcal{R}_{\text{eff}}$ explored in Eq. (5) holds regardless of the complicity of the epidemic model.

The reproduction number, $\mathcal{R}_0(t)$ and $\mathcal{R}_{\text{eff}}(t)$, reconstruction approach proposed in *Zhao et al. (2018b)* is concisely introduced in 'Reconstruction of the Reproduction Numbers from Compartmental Model'. The instantaneous (effective) reproduction number, $\mathcal{R}(t)$ or $\mathcal{R}_t$ estimation is described in details in 'Instantaneously Reproduction Number Estimation by Renewable Equation'. We compared the two forms of effective
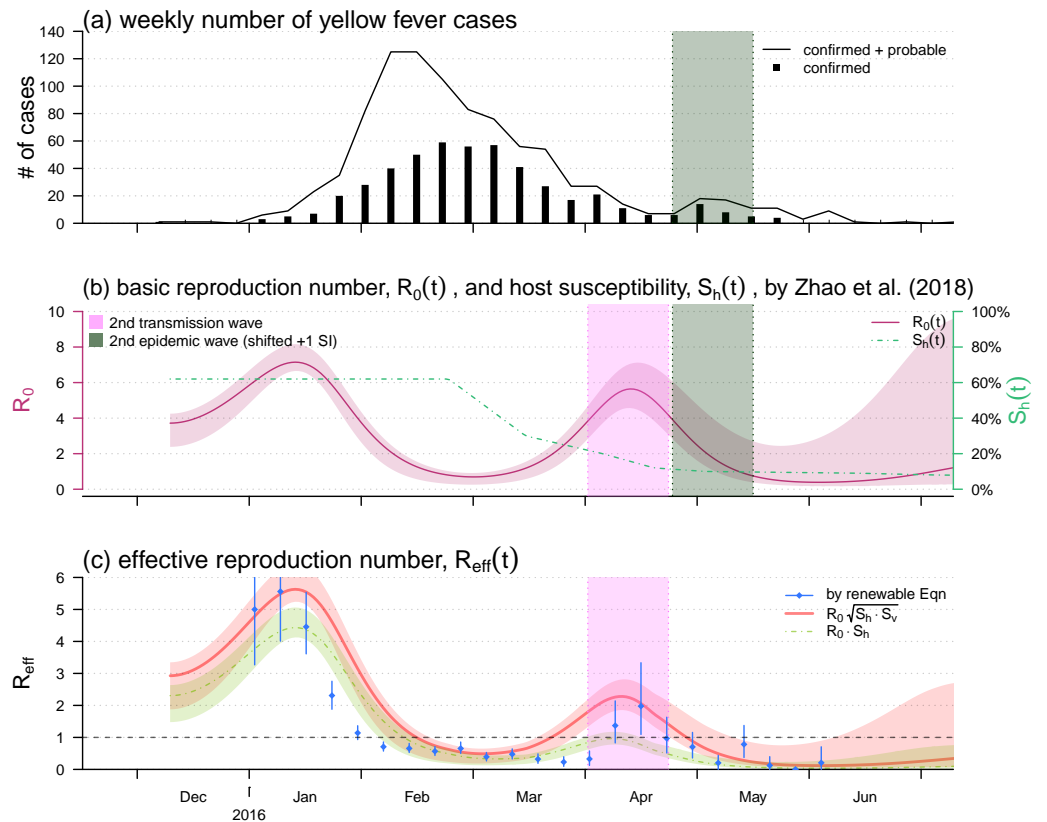
Figure 2 **The yellow fever (YF) epidemic and reproduction number estimation in Luanda, Angola from 2015 to 2016.** (A) The weekly number of YF cases time series, the second (minor) epidemic wave is shaded in grey. Panel (B) The reproduced time-varying basic reproduction number (purple curve), $\mathcal{R}_0(t)$, by using the reconstruction framework in *Zhao et al. (2018b)*, and the time-varying human hosts' susceptibility (green dashed line), $\mathcal{S}_h(t)$. The second transmission wave is highlighted in light purple by shifting one YF's serial interval (SI), averagely 23 days *Wu et al. (2016)*. The shedding area represent the 95% confidence intervals (CI). Panel (C) shows the effective reproduction numbers calculated (or estimated) by three different approaches, i.e., the $\mathcal{R}(t)$ by renewable equation (blue dots and bars) in Eq. (7), the product of $\mathcal{R}_0\sqrt{\mathcal{S}_h\mathcal{S}_v}$ (red bold curve) in Eq. (5), and the product of $\mathcal{R}_0\mathcal{S}_h$ (green dashed curve) in Eq. (1). The bars and the shedding areas represent the 95% CIs.

reproduction number, $\mathcal{R}_{\text{eff}}(t)$, as in Eqs. (1) and (5) with the estimation by using the renewable equation in Eq. (7). To summary, Table 2 listed the relevant notations in this section as well as the remains of this study.

Reconstruction of the reproduction numbers from compartmental model

We reproduce the time-varying basic reproduction number, $\mathcal{R}_0(t)$, by using the compartmental model and reconstruction approach proposed in *Zhao et al. (2018b)*. Here, we introduce the major reconstruction framework concisely, detailed procedures can be found in *Zhao et al. (2018b)*. The same framework was also implemented to study other infectious diseases (*He, Ionides & King, 2009*; *Ionides, Bretó & King, 2006*; *Shaman & Karspeck, 2012*; *Earn et al., 2012*; *Gao et al., 2016*).

Table 2 Summary table of the reproduction numbers' and susceptibilities' notations. All of the variables listed here are unit-free.

Notation	Interpretation	Formulation	Remark
$\mathcal{R}_0(t)$	Basic reproduction number	Eq. (3)	Reconstructed, 'Reconstruction of the Reproduction Numbers from Compartmental Model'
$\mathcal{S}_h(t)$	Susceptibility of hosts	Eq. (4)	Approximated, 'Reconstruction of the Reproduction Numbers from Compartmental Model'
$\mathcal{S}_v$	Susceptibility of vectors	Eq. (4)	Fixed to be 100%
$\mathcal{R}_{eff}(t)$	Effective reproduction number	Eq. (1) or Eq. (5)	Calculated, 'Reconstruction of the Reproduction Numbers from Compartmental Model'
$\mathcal{R}(t)$ or $\mathcal{R}_t$	Instantaneous (effective) reproduction number	Eq. (7)	Estimated, 'Instantaneously Reproduction Number Estimation by Renewable Equation'

The reconstructed $\mathcal{R}_0(t)$ is in the form of an exponential cubic spline function varying over the YF epidemic period, i.e., from December 2015 to June 2016. The shape of the cubic spline function are controlled by the number of nodes and value of each node. We set the nodes are evenly distributed over the YF epidemic period. The $\mathcal{R}_0(t)$ cubic spline function is estimated based on the maximal likelihood framework. We treat the compartmental model simulated number of cases time series as the theoretical case numbers, denoted by Z_i for the ith week in the epidemic period. Note that the Z_is are from the underlying time-dependent version of model (2). By contrast, the number of observed YF cases time series, denoted by C_i for the ith week, are regarded as random samples from a negative binomial (NB) process determined by the theoretical case numbers. We assume the observation noise follows an over-dispersed Poisson distribution (*Bretó et al., 2009*), and in particular, the C_is follow a Poisson process determined by Z_is. Furthermore, the rate of the Poisson process is considered to be a Gamma random variable, and thus, this leads to a NB process (*Lin et al., 2018*). The probability framework is described in Eq. (6).

$$C_i \sim \mathrm{NB}\left(\text{size} = \frac{1}{\tau}, \text{probability} = \frac{1}{1+\tau Z_i}\right),$$
$$\text{with mean} = Z_i \ \& \ \text{variance} = Z_i(1+\tau Z_i), \tag{6}$$

where the term τ is an over-dispersion parameter of the NB process that needs to be estimated. Thus, the overall log-likelihood value can be calculated by summing up to all log-probabilities of all is during the entire YF epidemic period. Therefore, the reconstructed $\mathcal{R}_0(t)$ can be estimated by finding the number of nodes and values of nodes (of the cubic spline function) with the "best fitting performance". We evaluate the fitting performance of the reconstructed $\mathcal{R}_0(t)$ by measuring the trade-off between the goodness-of-fit (in term of the log-likelihood) and the complexity of the model structure (in term of the number of parameters to be estimated). As in *Zhao et al. (2018b)*, the Bayesian information criterion (BIC) is employed to evaluate the fitting performance. By using the likelihood profile approach (*He, Ionides & King, 2009*; *He et al., 2017*; *Zhao et al., 2018a*; *Barndorff-Nielsen & Cox, 1994*; *Ionides, Bretó & King, 2006*), in which the profile of maximum log likelihood was calculated as a function of the model parameter, we estimated the 95% confidence interval (CI) of the reconstructed $\mathcal{R}_0(t)$. The simulation outcomes can be found in Fig. 3 of *Zhao et al. (2018b)*.

By using the local YF vaccination coverage, the data were publicly available via the African Health Observatory (*WHO, 2017*) as well as adopted in *Zhao et al. (2018b)*, we approximated the time-varying population susceptibility, $\mathcal{S}_h(t)$, as shown in Fig. 2B. The approximation is difference between the pre-existed population susceptibility, calculated by 1 minus the pre-existed YF-protection rate, and the time-varying local vaccination coverage.

Given $\mathcal{R}_0(t)$ and $\mathcal{S}_h(t)$, we calculated two different forms of the time-varying effective reproduction number by using the Eqs. (1) and (5) with $\mathcal{S}_v = 1$ fixed.

Instantaneously reproduction number estimation by renewable equation

The transmissibility of YF can be quantified by calculating the instantaneous (effective) reproduction number, $\mathcal{R}(t)$ and $\mathcal{R}_t$ for discrete scenarios, defined as the expected number of secondary cases generated by a single infectious individual during the infectious periods at time t. We estimated the $\mathcal{R}(t)$ from the YF cases time series by using the serial interval (SI) approach proposed by *Wallinga & Teunis (2004)*. The SI, in the epidemiology of infectious diseases, is the period of time between successive cases in a chain of transmission (*Porta, 2014*; *Fine, 2003*). If we know the distribution of the inter-arrival time of patients arrived at a clinic, we may simulate the sequence of patients arrivals. Similarly, if we know the distribution of SI, we may simulate the sequence of infections, adding that one primary infection could lead to a number of, which is determined by the reproduction number, secondary infections. Reversely, if we know the distribution of SI and the cases time series, we can reconstruct the reproductive number backwardly. This SI approach was extended by *Forsberg White & Pagano (2008)*, *Katriel et al. (2011)*, *Ali, Kadi & Ferguson (2013)*, *Fraser (2007)*, *Cori et al. (2013)* and *Wallinga & Lipsitch (2006)*, and also implemented to study several vector-borne diseases (*Ferguson et al., 2016*; *Zhao et al., 2019a*; *Wu et al., 2016*; *Zhao et al., 2019b*). Hence, the time-varying $\mathcal{R}(t)$ is estimated from the renewal equation in Eq. (7).

$$\mathcal{R}(t) = \frac{x(t)}{\int_0^\infty w(k)x(t-k)\,\mathrm{d}k}, \tag{7}$$

where $x(t)$ is the YF incidence rate at time t. The convolution term $\int_0^\infty w(k)x(t-k)\,\mathrm{d}k$ is the measurement of the total infectiousness at time t. The term $w(k)$ is the YF SI distribution that describes the distribution of the infectiousness during the period of infection. We adopted the same approach, similar methods were also implemented in *Zhao et al. (2019c)*, *Zhao et al. (2019d)*, *Cowling et al. (2009)* and *Ferguson et al. (2016)*, to find SI, $w(k)$, and epidemiology parameter setting as in *Wu et al. (2016)* as well as the references mentioned in it (*Johansson et al., 2012*; *Johansson et al., 2010*), and we had the numerical estimation of $w(k)$ with the mean SI of 23 days.

We estimated the $\mathcal{R}(t)$ of YF between January and May of 2016. After this time period, weeks of zero confirmed case appeared. The 95% confidence intervals (CI) were estimated based on the Gamma priors of each $\mathcal{R}(t)$ (*Ali, Kadi & Ferguson, 2013*; *Cori et al., 2013*).
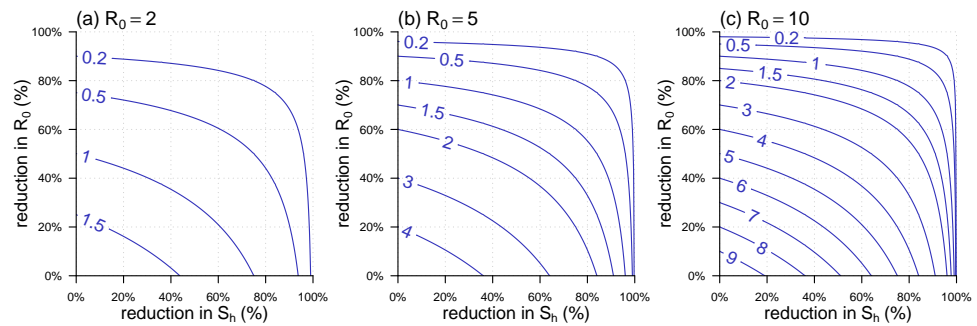
Figure 3 **The contour plot of the effective reproduction number, $\mathcal{R}_{\text{eff}}$, against the percentage reductions in $\mathcal{S}_{\text{h}}$ and $\mathcal{R}_0$.** (A –C) The scenarios of highest $\mathcal{R}_0 = 2$, $\mathcal{R}_0 = 5$, and $\mathcal{R}_0 = 10$, respectively. The labels on the blue curves show the values of $\mathcal{R}_{\text{eff}}$. The horizontal axis is the percentage reduction in $\mathcal{S}_{\text{h}}$. The vertical axis is the percentage reduction in $\mathcal{R}_0$ that is 1 minus the ratio of the reduced $\mathcal{R}_0$ over its original (or highest) value, i.e., the value in the panel label.

RESULTS

The analytic formula of $\mathcal{R}_{\text{eff}}$ was given in Eq. (5). Since in many situations, the $\mathcal{R}_0(t)$ is time-varying, we defined the percentage reduction in $\mathcal{R}_0$ as 1 minus the ratio of the reduced $\mathcal{R}_0$ over its original highest value, usually the basic reproduction number during the initial outbreak. Similarly, the percentage reduction in $\mathcal{S}_{\text{h}}$ can be defined as 1 minus the host susceptibility. We noted that both percentage reductions in $\mathcal{S}_{\text{h}}$ and $\mathcal{R}_0$ ranged from 0 to 1. The relationship of $\mathcal{R}_{\text{eff}}$ against the percentage reductions in $\mathcal{S}_{\text{h}}$ and $\mathcal{R}_0$ was shown in Fig. 3. Figs. 3(A)–3(C) were the scenarios of the highest $\mathcal{R}_0 = 2$, $\mathcal{R}_0 = 5$, and $\mathcal{R}_0 = 10$ respectively. The range of $\mathcal{R}_0$ from 2 to 10 covers most of the vector-borne diseases' basic reproduction numbers during the initial outbreak.

We used the 2015–2016 YF outbreak in Luanda to demonstrate the effective reproduction number formula in Eq. (5). The YF epidemic was shown in Fig. 2A. Subsequent to the (first) major epidemic wave that peaked in the February of 2016, we observed a second minor wave that followed the major wave, which peaked in May. We showed the reconstructed $\mathcal{R}_0(t)$ and approximated $\mathcal{S}_{\text{h}}(t)$ reproduced from *Zhao et al. (2018b)* in Fig. 2B. The local human susceptibility, $\mathcal{S}_{\text{h}}(t)$, was decreasing by the end of February and ended up less than 7% due to the timely mass vaccination campaign (*Shearer et al., 2017*; *Wu et al., 2016*; *Zhao et al., 2018b*; *Kraemer et al., 2017*). We found two peaks in the reconstructed $\mathcal{R}_0(t)$, of which the highest value was found to be 7.1 during the first wave that peaked in January. The second peak of $\mathcal{R}_0(t)$ occurred in the April of 2016 with the local maximal value of 5.6. We matched the second peak in $\mathcal{R}_0(t)$, highlighted in purple, and the minor epidemic wave in the YF incidences, highlighted in grey, by one SI shift, i.e., 23 days averagely (*Wu et al., 2016*).

Figure 2C shows the effective reproduction number calculated or estimated by Eq. (1) or Eq. (5) or the renewable equation in Eq. (7). The (first) major transmission, i.e., $\mathcal{R}_{\text{eff}}(t)$, wave peaked in January of 2016 associated with the major epidemic wave that peaked in February (Fig. 2A). We found the $\mathcal{R}_{\text{eff}}(t)$ series based on $\mathcal{R}_0(t)$ and $\mathcal{S}_{\text{h}}(t)$ (in Fig. 2B)
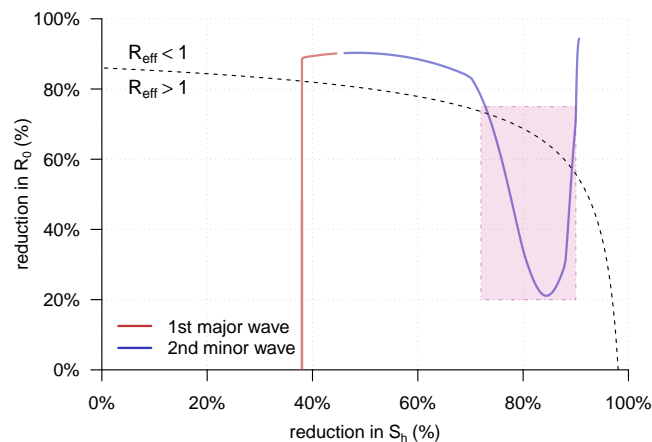
Figure 4 **The trajectory of the $\mathcal{R}_{\text{eff}}(t)$ from Eq. (5) against the percentage reductions in $\mathcal{S}_{\text{h}}$ and $\mathcal{R}_0$ of the yellow fever (YF) epidemic in Luanda, Angola from 2015 to 2016.** The horizontal and vertical axes have the same setting as in Fig. 3. Here, the original highest $\mathcal{R}_0 = 7.1$ as the same as in Fig. 2B. The black dashed curve represents the level of $\mathcal{R}_{\text{eff}} = 1$. The area under the $\mathcal{R}_{\text{eff}} = 1$ curve is for $\mathcal{R}_{\text{eff}} > 1$, and those above is for $\mathcal{R}_{\text{eff}} < 1$. The red trajectory represents the changing dynamics of $\mathcal{S}_{\text{h}}$ and $\mathcal{R}_0$ during the first major epidemic wave from December 2015 to February 2016. The purple trajectory represents the changing dynamics of $\mathcal{S}_{\text{h}}$ and $\mathcal{R}_0$ during the second minor epidemic wave from March to May 2016. The $\mathcal{R}_{\text{eff}} > 1$ part during the second transmission wave, highlighted in purple in Fig. 2C, are marked in the purple rectangle.

were (almost) synchronised, i.e., in-phase, with the estimated $\mathcal{R}_{\text{eff}}(t)$ (or $\mathcal{R}(t)$) series by the renewable equation. The highest $\mathcal{R}_{\text{eff}}(t)$ estimate by renewable equation was 5.5, and the same as the estimate by using Eq. (5) that was also 5.5, whereas the Eq. (1) version is 4.4. Similar to the trends of YF incidences and $\mathcal{R}_0(t)$, we also found a second minor wave in $\mathcal{R}_{\text{eff}}(t)$ around April, highlighted in purple. During this minor wave, the local maximal $\mathcal{R}_{\text{eff}}(t)$ estimate by renewable equation was 2.0 (95% CI [1.1–3.4]) that is larger than 1 significantly, the "$\mathcal{R}_0\sqrt{\mathcal{S}_{\text{h}}\mathcal{S}_{\text{v}}}$" version was 2.3, and the "$\mathcal{R}_0\mathcal{S}_{\text{h}}$" version is 0.9 (< 1).

Figure 4 showed the trajectory of the $\mathcal{R}_{\text{eff}}(t)$ from Eq. (5) against the percentage reductions in $\mathcal{S}_{\text{h}}$ and $\mathcal{R}_0$ of the YF epidemic in Luanda. Consistent with the observations in Fig. 2C, we found both of the two transmission waves of YF $\mathcal{R}_{\text{eff}}(t)$ moved across the disease control threshold, i.e., $\mathcal{R}_{\text{eff}} = 1$. For the second transmission waves, marked in the purple rectangle, it "broke" the $\mathcal{R}_{\text{eff}} = 1$ boundary and thus associated with the second (minor) YF epidemic wave in May of 2016, highlighted in grey in Fig. 2A.

DISCUSSION

In this work, a simple epidemic model (2) is developed to study the transmission dynamics of vector-borne diseases. We formulated the analytic form of the effective reproduction number, $\mathcal{R}_{\text{eff}}$, with respect to the basic reproduction number, $\mathcal{R}_0$, and the susceptibilities of the vector ($\mathcal{S}_{\text{v}}$) and host ($\mathcal{S}_{\text{h}}$) for vector-borne diseases in Eq. (5). The $\mathcal{R}_{\text{eff}}$ from Eq. (5) were compared with the $\mathcal{R}_{\text{eff}}$ of the classic airborne infectious disease in Eq. (1) as well as the estimation by the SI approach in Eq. (7). We re-visited the yellow fever (YF) outbreak in Luanda, and used this epidemic as an example to compare the $\mathcal{R}_{\text{eff}}$

calculation and estimation. Although there existed differences in the three $\mathcal{R}_{\text{eff}}$ series during the first transmission wave around January 2016 in Fig. 2, the $\mathcal{R}_{\text{eff}}$ values were roughly synchronised. However, for the second (minor) transmission wave around April 2016, the $\mathcal{R}_{\text{eff}}$s from Eq. (5) were consistent with the estimates from the renewable equation that were significantly larger than 1, the "$\mathcal{R}_0\mathcal{S}_{\text{h}}$" appeared inconsistent with the formers and lower than 1. According to theoretical epidemiology (*Earn et al., 2008*; *Brauer & Castillo-Chavez, 2001*; *Keeling & Rohani, 2011*; *Grenfell, Dobson & Moffatt, 1995*), the condition that $\mathcal{R}_{\text{eff}} < 1$ guarantees the disease under control. The $\mathcal{R}_0\mathcal{S}_{\text{h}}$ was calculated lower than 1 since the mid-February 2016, and this contradicted with the occurrence of the second YF epidemic wave in May. Therefore, the "$\mathcal{R}_0\mathcal{S}_{\text{h}}$" form of effective reproduction number was demonstrated unqualified for measuring the transmissibility of a vector-borne disease. On the other hand, our derived $\mathcal{R}_{\text{eff}} = \mathcal{R}_0\sqrt{\mathcal{S}_{\text{h}}\mathcal{S}_{\text{v}}}$, Eq. (5), matched the two waves of both YF incidences time series and the $\mathcal{R}(t)$ estimates by renewable equation well.

Different from the vector-free context, the $\mathcal{R}_{\text{eff}} = \mathcal{R}_0\sqrt{\mathcal{S}_{\text{h}}\mathcal{S}_{\text{v}}}$ for the vector-borne diseases indicates that the disease control effectiveness, i.e., $\mathcal{R}_{\text{eff}}$, non-linearly depends on the control of the host's susceptibility, $\mathcal{S}_{\text{h}}$. Figure 3 shows that the reduction in $\mathcal{S}_{\text{h}}$ is (relatively) less effective in reducing $\mathcal{R}_{\text{eff}}$ during the initial stage, i.e., from 0% onwards, and becomes more effective when the cumulative reduction of $\mathcal{S}_{\text{h}}$ grows. This finding suggests that directly reducing $\mathcal{R}_0$, via, e.g., vector elimination, avoiding exposure to vectors, improving treatment, etc., could be a more efficient option to control the vector-borne diseases, especially when the herd protection in the host population is difficult to build up.

Although the next generation matrix method in 'Basic Reproduction Number' is valid around a disease-free equilibrium (DFE) of model (2) (*Van den Driessche & Watmough, 2002*; *Van den Driessche, 2017*), the (asymptotic) stability of the endemic equilibrium (EE) will allow that the product of the $\mathcal{R}_0$ multiplying the susceptibility can be interpreted as $\mathcal{R}_{\text{eff}}$. The $B_v(t)$ can be treated as a constant, $\langle B_v \rangle$, when we consider a sufficiently short period of time. Hence, during this short period of time, the EE of model (2) is asymptotically stable, and the Eq. (5) holds. More precisely, Eq. (5) follows a more general version as follows,

$$\mathcal{R}_{\text{eff}}(t) = \mathcal{R}_0(t)\sqrt{\mathcal{S}_{\text{h}}(t)\mathcal{S}_{\text{v}}(t)}.$$

The global asymptotic stability (GAS) of EE can be further guaranteed as $N_v' = B_v(t) - \mu_v N_v = 0$ in model (2), and this leads to the condition that $B_v(t) = \langle B_v \rangle = \mu_v N_v$.

This work used the serial interval (SI) approach, i.e., the renewable equation, to estimate the instantaneous effective reproduction number, $\mathcal{R}(t)$, for further comparison. The estimates of $\mathcal{R}(t)$ depended on the choice of the distribution of SI, i.e., the $w(k)$ in Eq. (7). Accounting for the initial susceptibility of 63% of the YF epidemic (*WHO, 2017*), *Wu et al. (2016)* estimated that the YF basic reproduction number of $\mathcal{R}_0 = 8.3$ (95% CI [6.8–9.7]) with mean SI of 23 days, and $\mathcal{R}_0 = 11.3$ (95% CI [8.7–13.8]) with mean SI of 32 days. *Kraemer et al. (2017)* estimated that $\mathcal{R}_0 = 7.6$ (95% CI [6.3–8.9]) with mean SI 15 days. We adopted the mean SI of 23 days as in *Wu et al. (2016)* in this work to estimate the $\mathcal{R}(t)$ series. Different (but reasonable) settings on SI will distinguish the second transmission

wave (as highlighted in purple in Figs. 2C and 4) with $\mathcal{R}_{\text{eff}} > 1$ indifferently. In addition, we note that slight changes in the choice of YF SI will not affect our main results.

This work intends to estimate and compare different forms of the time-varying effective reproduction number. We have adopted two different but widely used approaches, i.e., the maximal likelihood-based reconstruction and the SI, i.e., by the renewable equation, based estimation, which includes three different formations in Eqs. (1), (5) and (7). The maximum likelihood-based reconstruction method and the SI based estimation method are different in the calculation procedures and theoretical features. The maximum likelihood-based reconstruction relies on the mechanistic disease transmission model, e.g., model (2), and thus, it follows a biologically reasonable model structure. It is able to disentangle the changing dynamics of susceptibility, $\mathcal{S}_{\text{h}}(t)$, and the basic reproduction number, $\mathcal{R}_0(t)$, based the number of cases time series and other reasonable epidemiological settings. Hence, we adopted this approach to find both $\mathcal{S}_{\text{h}}(t)$ and $\mathcal{R}_0(t)$, and further calculate the $\mathcal{R}_{\text{eff}}(t)$ in two difference forms. The SI based (renewable equation based) estimation method is to calculate descriptive statistics by nature. By directly using the number of disease cases time series and the knowledge (distribution) of SI, the $\mathcal{R}(t)$ can be estimated straightforwardly.

To derive the $\mathcal{R}_{\text{eff}}$ in Eq. (5), we used the next generation matrix approach in 'Effective Reproduction Number' and considered the transition ''from host to vector to host'' as two generations, which is consistent with (*Gao et al., 2016*; *Tang et al., 2016*; *Zhao et al., 2018b*; *Champagne et al., 2016*; *d Pinho et al., 2010*; *Musa et al., 2019*; *Wang et al., 2012*). As also remarked in *Brauer et al. (2016)* and *Van den Driessche (2017)*, some other studies treated the same transition ''from host to vector to host'' as a single combined generation (*Tennant & Recker, 2018*; *Chowell et al., 2007*; *Kucharski et al., 2016*; *Towers et al., 2016*; *Mideo & Day, 2008*). Although the two choices have the same threshold value and follow the same mathematical criteria to judge the stability of compartmental models, to be aware of their difference is crucial. We remark that if considering the aforementioned transition as a single generation, the basic reproduction number would be the square of the $\mathcal{R}_0$ in Eq. (3). In this case, the effective reproduction number is $\mathcal{R}_{\text{eff}} = \mathcal{R}_0 \mathcal{S}_{\text{h}} \mathcal{S}_{\text{v}}$. In our YF epidemic example, we demonstrated that the misuse of the Eq. (1) is likely to cause misleading or contradictory outcomes in studying the vector-borne diseases outbreak. The two forms of the reproduction numbers have different biological interpretations due to the different definitions of generations, nevertheless one can be transformed to the other.

Our modelling study, specially the derived " $\mathcal{R}_{\text{eff}} = \mathcal{R}_0 \sqrt{\mathcal{S}_{\text{h}} \mathcal{S}_{\text{v}}}$ " relationship, has limitations mainly due to the model settings and structures. As stated in the analysis parts, the relationship holds on the condition that the transmission paths remain from-host-to-vector and from-vector-to-host. Hence, when direct transmission occurs, e.g., sexual transmission between hosts in Zika virus (*Gao et al., 2016*), the $\mathcal{R}_{\text{eff}} \neq \mathcal{R}_0 \sqrt{\mathcal{S}_{\text{h}} \mathcal{S}_{\text{v}}}$. However, since the sexual transmission merely contributes very minor infections, it can be ignored in scale, and thus $\mathcal{R}_{\text{eff}} \approx \mathcal{R}_0 \sqrt{\mathcal{S}_{\text{h}} \mathcal{S}_{\text{v}}}$. The numerical results and estimates in this work are calculated with $\mathcal{S}_{\text{v}} = 1$ fixed. This $\mathcal{S}_{\text{v}} = 1$ is based on two facts that the vector's lifespan is much shorter than host's lifespan; and the infected vectors do not recover (for most of the vector-borne diseases). These two facts will lead to an outcome in model (2) that class $R_v = 0$ and class I_v is remarkably small and almost zero. As the matter of fact, I_v, i.e.,

the prevalence of disease in vectors, is likely to increase when the pathogen is extremely infectious and the vertical transmission accounts. Therefore, the surveillance on the disease prevalence in vectors would be helpful for calculating the $\mathcal{R}_{\text{eff}}$.

CONCLUSIONS

We formulate the analytic form of the $\mathcal{R}_{\text{eff}} = \mathcal{R}_0\sqrt{\mathcal{S}_h\mathcal{S}_v}$ for the vector-borne diseases. We demonstrate the $\mathcal{R}_{\text{eff}}$ formulation is consistent with the estimates of the 2015–2016 yellow fever outbreak in Luanda, and distinguishes the second minor epidemic wave significantly. We remark that it is important to be aware of whether one or two generations is considered for the transition "from host to vector to host" in the infectious diseases modelling studies.

Funding

This work was supported by a grant from the Hong Kong Polytechnic University (project no.: 1-ZE8J). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Hong Kong Polytechnic University: 1-ZE8J.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Shi Zhao conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Salihu S. Musa performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Jay T. Hebert, Peihua Cao, Jinjun Ran, Jiayi Meng and Jing Qin performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Daihai He analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

This is a methodology paper, and the method with examples is explained in the text. Data and code are available in the Supplemental Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.8601#supplemental-information.

REFERENCES

Ali ST, Kadi AS, Ferguson NM. 2013. Transmission dynamics of the 2009 influenza A (H1N1) pandemic in India: the impact of holiday-related school closure. *Epidemics* **5(4)**:157–163 DOI 10.1016/j.epidem.2013.08.001.

Barndorff-Nielsen OE, Cox DR. 1994. *Inference and asymptotics.* London: Chapman & Hall.

Barrett ADT. 2018. The reemergence of yellow fever. *Science* **361(6405)**:847–848 DOI 10.1126/science.aau8225.

Besnard M, Lastere S, Teissier A, Cao-Lormeau VM, Musso D. 2014. Evidence of perinatal transmission of Zika virus, French Polynesia, December 2013 and February 2014. *Eurosurveillance* **19(13)**:Article 20751 DOI 10.2807/1560-7917.ES2014.19.13.20751.

Brauer F, Castillo-Chavez C. 2001. *Mathematical models in population biology and epidemiology.* Vol. 40. New York: Springer.

Brauer F, Castillo-Chavez C, Mubayi A, Towers S. 2016. Some models for epidemics of vector-transmitted diseases. *Infectious Disease Modelling* **1(1)**:79–87 DOI 10.1016/j.idm.2016.08.001.

Bretó C, He D, Ionides EL, King AA. 2009. Time series analysis via mechanistic models. *The Annals of Applied Statistics* **3(1)**:319–348 DOI 10.1214/08-AOAS201.

Caminade C, Kovats S, Rocklov J, Tompkins AM, Morse AP, Colón-González FJ, Stenlund H, Martens P, Lloyd SJ. 2014. Impact of climate change on global malaria distribution. *Proceedings of the National Academy of Sciences of the United States of America* **111(9)**:3286–3291 DOI 10.1073/pnas.1302089111.

Cauchemez S, Besnard M, Bompard P, Dub T, Guillemette-Artur P, Eyrolle-Guignot D, Salje H, Van Kerkhove MD, Abadie V, Garel C. 2016. Association between Zika virus and microcephaly in French Polynesia, 2013–15: a retrospective study. *The Lancet* **387(10033)**:2125–2132 DOI 10.1016/S0140-6736(16)00651-6.

Champagne C, Salthouse DG, Paul R, Cao-Lormeau VM, Roche B, Cazelles B. 2016. Structure in the variability of the basic reproductive number (R0) for Zika epidemics in the Pacific islands. *Elife* **5**:e19874 DOI 10.7554/eLife.19874.

Chowell G, Diaz-Duenas P, Miller JC, Alcazar-Velazco A, Hyman JM, Fenimore PW, Castillo-Chavez C. 2007. Estimation of the reproduction number of dengue fever from spatial epidemic data. *Mathematical Biosciences* **208(2)**:571–589 DOI 10.1016/j.mbs.2006.11.011.

Cori A, Ferguson NM, Fraser C, Cauchemez S. 2013. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of Epidemiology* **178(9)**:1505–1512 DOI 10.1093/aje/kwt133.

Cowling BJ, Fang VJ, Riley S, Peiris JSM, Leung GM. 2009. Estimation of the serial interval of influenza. *Epidemiology* **20(3)**:344–347 DOI 10.1097/EDE.0b013e31819d1092.

Cushing JM, Diekmann O. 2016. The many guises of R0 (a didactic note). *Journal of Theoretical Biology* **404**:295–302 DOI 10.1016/j.jtbi.2016.06.017.

Diekmann O, Heesterbeek JAP, Roberts MG. 2009. The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society Interface* **7(47)**:873–885.

Diekmann O, Heesterbeek JAP, Metz JAJ. 1990. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology* **28(4)**:365–382.

d Pinho STR, Ferreira CP, Esteva L, Barreto FR, Morato e Silva VC, Teixeira MGL. 2010. Modelling the dynamics of dengue real epidemics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **368(1933)**:5679–5693 DOI 10.1098/rsta.2010.0278.

Duffy MR, Chen T-H, Hancock WT, Powers AM, Kool JL, Lanciotti RS, Pretrick M, Marfel M, Holzbauer S, Dubray C. 2009. Zika virus outbreak on Yap Island, federated states of Micronesia. *New England Journal of Medicine* **360(24)**:2536–2543 DOI 10.1056/NEJMoa0805715.

Earn DJ, Brauer F, Van den Driessche P, Wu J. 2008. *Mathematical epidemiology*. Berlin: Springer.

Earn DJD, He D, Loeb MB, Fonseca K, Lee BE, Dushoff J. 2012. Effects of school closure on incidence of pandemic influenza in Alberta, Canada. *Annals of Internal Medicine* **156(3)**:173–181 DOI 10.7326/0003-4819-156-3-201202070-00005.

Fenwick A. 2012. The global burden of neglected tropical diseases. *Public Health* **126(3)**:233–236 DOI 10.1016/j.puhe.2011.11.015.

Ferguson NM, Cucunubá ZM, Dorigatti I, Nedjati-Gilani GL, Donnelly CA, Basáñez MG, Nouvellet P, Lessler J. 2016. Countering the Zika epidemic in Latin America. *Science* **353(6297)**:353–354 DOI 10.1126/science.aag0219.

Fine PEM. 2003. The interval between successive cases of an infectious disease. *American Journal of Epidemiology* **158(11)**:1039–1047 DOI 10.1093/aje/kwg251.

Fischer M, Staples JE. 2014. Chikungunya virus spreads in the Americas—Caribbean and South America, 2013–2014. *Morbidity and Mortality Weekly Report* **63(22)**:500–501.

Forsberg White L, Pagano M. 2008. A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. *Statistics in Medicine* **27(16)**:2999–3016 DOI 10.1002/sim.3136.

Fraser C. 2007. Estimating individual and household reproduction numbers in an emerging epidemic. *PLOS ONE* **2(8)**:e758 DOI 10.1371/journal.pone.0000758.

Gao D, Lou Y, He D, Porco TC, Kuang Y, Chowell G, Ruan S. 2016. Prevention and control of Zika as a mosquito-borne and sexually transmitted disease: a mathematical modeling analysis. *Scientific Reports* **6**:28070 DOI 10.1038/srep28070.

Grenfell BT, Dobson AP, Moffatt HK. 1995. *Ecology of infectious diseases in natural populations*. Vol. 7. Cambridge, UK: Cambridge University Press.

He D, Gao D, Lou Y, Zhao S, Ruan S. 2017. A comparison study of Zika virus outbreaks in French Polynesia, Colombia and the State of Bahia in Brazil. *Scientific Reports* **7(1)**:273 DOI 10.1038/s41598-017-00253-1.

He D, Ionides EL, King AA. 2009. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *Journal of the Royal Society Interface* **7(43)**:271–283.

He D, Zhao S, Lin Q, Musa SS, Stone L. 2019. New estimates of the Zika virus epidemic attack rate in Northeastern Brazil from 2015 to 2016: a modelling analysis based on Guillain-Barre Syndrome (GBS) surveillance data. *BioRxiv* 657015.

Heffernan JM, Smith RJ, Wahl LM. 2005. Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface* **2(4)**:281–293 DOI 10.1098/rsif.2005.0042.

Ionides EL, Bretó C, King AA. 2006. Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences of the United States of America* **103(49)**:18438–18443 DOI 10.1073/pnas.0603181103.

Johansson MA, Arana-Vizcarrondo N, Biggerstaff BJ, Gallagher N, Marano N, Staples JE. 2012. Assessing the risk of international spread of yellow fever virus: a mathematical analysis of an urban outbreak in Asuncion, 2008. *The American Journal of Tropical Medicine and Hygiene* **86(2)**:349–358 DOI 10.4269/ajtmh.2012.11-0432.

Johansson MA, Arana-Vizcarrondo N, Biggerstaff BJ, Staples JE. 2010. Incubation periods of yellow fever virus. *The American Journal of Tropical Medicine and Hygiene* **83(1)**:183–188 DOI 10.4269/ajtmh.2010.09-0782.

Johansson MA, Mier-y Teran-Romero L, Reefhuis J, Gilboa SM, Hills SL. 2016. Zika and the risk of microcephaly. *New England Journal of Medicine* **375(1)**:1–4 DOI 10.1056/NEJMp1605367.

Katriel G, Yaari R, Huppert A, Roll U, Stone L. 2011. Modelling the initial phase of an epidemic using incidence and infection network data: 2009 H1N1 pandemic in Israel as a case study. *Journal of the Royal Society Interface* **8(59)**:856–867 DOI 10.1098/rsif.2010.0515.

Keeling MJ, Rohani P. 2011. *Modeling infectious diseases in humans and animals*. Princeton, USA: Princeton University Press.

Kilpatrick AM. 2011. Globalization, land use, and the invasion of West Nile virus. *Science* **334(6054)**:323–327 DOI 10.1126/science.1201010.

Kraemer MUG, Cummings DAT, Funk S, Reiner RC, Faria NR, Pybus OG, Cauchemez S. 2019. Reconstruction and prediction of viral disease epidemics. *Epidemiology & Infection* **147**:Article e34.

Kraemer MUG, Faria NR, Reiner Jr RC, Golding N, Nikolay B, Stasse S, Johansson MA, Salje H, Faye O, Wint GRW. 2017. Spread of yellow fever virus outbreak in Angola and the Democratic Republic of the Congo 2015–16: a modelling study. *The Lancet Infectious Diseases* **17(3)**:330–338 DOI 10.1016/S1473-3099(16)30513-8.

Kucharski AJ, Funk S, Eggo RM, Mallet H-P, Edmunds WJ, Nilles EJ. 2016. Transmission dynamics of Zika virus in island populations: a modelling analysis of the 2013–14 French Polynesia outbreak. *PLOS Neglected Tropical Diseases* **10(5)**:e0004726 DOI 10.1371/journal.pntd.0004726.

Lin Q, Chiu APY, Zhao S, He D. 2018. Modeling the spread of Middle East respiratory syndrome coronavirus in Saudi Arabia. *Statistical Methods in Medical Research* **27(7)**:1968–1978 DOI 10.1177/0962280217746442.

Mideo N, Day T. 2008. On the evolution of reproductive restraint in malaria. *Proceedings of the Royal Society B: Biological Sciences* **275(1639)**:1217–1224 DOI 10.1098/rspb.2007.1545.

Murray CJL, Rosenfeld LC, Lim SS, Andrews KG, Foreman KJ, Haring D, Fullman N, Naghavi M, Lozano R, Lopez AD. 2012. Global malaria mortality between 1980 and 2010: a systematic analysis. *The Lancet* **379(9814)**:413–431 DOI 10.1016/S0140-6736(12)60034-8.

Musa SS, Zhao S, Chan HS, Jin Z, He D. 2019. A mathematical model to study the 2014–2015 large-scale dengue epidemics in Kaohsiung and Tainan cities in Taiwan, China. *Mathematical Biosciences and Engineering* **16(5)**:3841–3863 DOI 10.3934/mbe.2019190.

Niebylski ML, Peacock MG, Schwan TG. 1999. Lethal effect of Rickettsia rickettsiion its tick vector (Dermacentor andersoni). *Applied and Environmental Microbiology* **65(2)**:773–778 DOI 10.1128/AEM.65.2.773-778.1999.

Porta M. 2014. *A dictionary of epidemiology.* Oxford, UK: Oxford university press.

Saad-Roy CM, Van den Driessche P, Ma J. 2016. Estimation of Zika virus prevalence by appearance of microcephaly. *BMC Infectious Diseases* **16(1)**:754 DOI 10.1186/s12879-016-2076-z.

Sang S, Gu S, Bi P, Yang W, Yang Z, Xu L, Yang J, Liu X, Jiang T, Wu H. 2015. Predicting unprecedented dengue outbreak using imported cases and climatic factors in Guangzhou, 2014. *PLOS Neglected Tropical Diseases* **9(5)**:e0003808 DOI 10.1371/journal.pntd.0003808.

Shaman J, Karspeck A. 2012. Forecasting seasonal outbreaks of influenza. *Proceedings of the National Academy of Sciences of the United States of America* **109(50)**:20425–20430 DOI 10.1073/pnas.1208772109.

Shearer FM, Moyes CL, Pigott DM, Brady OJ, Marinho F, Deshpande A, Longbottom J, Browne AJ, Kraemer MUG, O'Reilly KM. 2017. Global yellow fever vaccination coverage from 1970 to 2016: an adjusted retrospective analysis. *The Lancet Infectious Diseases* **17(11)**:1209–1217 DOI 10.1016/S1473-3099(17)30419-X.

Stanaway JD, Shepard DS, Undurraga EA, Halasa YA, Coffeng LE, Brady OJ, Hay SI, Bedi N, Bensenor IM, Castañeda-Orjuela CA. 2016. The global burden of dengue: an analysis from the Global Burden of Disease Study 2013. *The Lancet Infectious Diseases* **16(6)**:712–723 DOI 10.1016/S1473-3099(16)00026-8.

Tang B, Xiao Y, Tang S, Wu J. 2016. Modelling weekly vector control against dengue in the Guangdong Province of China. *Journal of Theoretical Biology* **410**:65–76 DOI 10.1016/j.jtbi.2016.09.012.

Tennant W, Recker M. 2018. Robustness of the reproductive number estimates in vector-borne disease systems. *PLOS Neglected Tropical Diseases* **12(12)**:e0006999 DOI 10.1371/journal.pntd.0006999.

Towers S, Brauer F, Castillo-Chavez C, Falconar AKI, Mubayi A, Romero-Vivas CME. 2016. Estimate of the reproduction number of the 2015 Zika virus outbreak in Barranquilla, Colombia, and estimation of the relative role of sexual transmission. *Epidemics* **17**:50–55 DOI 10.1016/j.epidem.2016.10.003.

Van den Driessche P. 2017. Reproduction numbers of infectious disease models. *Infectious Disease Modelling* **2(3)**:288–303 DOI 10.1016/j.idm.2017.06.002.

Van den Driessche P, Watmough J. 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences* **180(1–2)**:29–48 DOI 10.1016/S0025-5564(02)00108-6.

Wallinga J, Lipsitch M. 2006. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences* **274(1609)**:599–604.

Wallinga J, Teunis P. 2004. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *American Journal of Epidemiology* **160(6)**:509–516 DOI 10.1093/aje/kwh255.

Wang Y, Jin Z, Yang Z, Zhang Z-K, Zhou T, Sun G-Q. 2012. Global analysis of an SIS model with an infective vector on complex networks. *Nonlinear Analysis: Real World Applications* **13(2)**:543–557 DOI 10.1016/j.nonrwa.2011.07.033.

Weaver SC, Lecuit M. 2015. Chikungunya virus and the global spread of a mosquito-borne disease. *New England Journal of Medicine* **372(13)**:1231–1239 DOI 10.1056/NEJMra1406035.

WHO. 2017. African Health Observatory (AHO) of the World Health Organization (WHO). Publications website, Health Topic: Yellow Fever. Yellow fever outbreak in Angola Situation Reports. *Available at http://www.afro.who.int/publications*.

Wu JT, Peak CM, Leung GM, Lipsitch M. 2016. Fractional dosing of yellow fever vaccine to extend supply: a modelling study. *The Lancet* **388(10062)**:2904–2911 DOI 10.1016/S0140-6736(16)31838-4.

Zhao S, Lou Y, Chiu APY, He D. 2018a. Modelling the skip-and-resurgence of Japanese encephalitis epidemics in Hong Kong. *Journal of Theoretical Biology* **454**:1–10 DOI 10.1016/j.jtbi.2018.05.017.

Zhao S, Musa SS, Fu H, He D, Qin J. 2019a. Simple framework for real-time forecast in a data-limited situation: the Zika virus (ZIKV) outbreaks in Brazil from 2015 to 2016 as an example. *Parasites & Vectors* **12(1)**:344 DOI 10.1186/s13071-019-3602-9.

Zhao S, Musa SS, Meng J, Qin J, He D. 2019b. The long-term changing dynamics of dengue infectivity in Guangdong, China, from 2008–2018: a modelling analysis. *Transactions of The Royal Society of Tropical Medicine and Hygiene* **114(1)**:62–71.

Zhao S, Musa SS, Qin J, He D. 2019c. Associations between Public Awareness, Local Precipitation, and Cholera in Yemen in 2017. *The American Journal of Tropical Medicine and Hygiene* **101(3)**:521–524 DOI 10.4269/ajtmh.18-1016.

Zhao S, Musa SS, Qin J, He D. 2019d. Phase-shifting of the transmissibility of macrolide-sensitive and resistant Mycoplasma pneumoniae epidemics in Hong Kong, from 2015 to 2018. *International Journal of Infectious Diseases* **81**:251–253 DOI 10.1016/j.ijid.2019.02.030.

Zhao S, Stone L, Gao D, He D. 2018b. Modelling the large-scale yellow fever outbreak in Luanda, Angola, and the impact of vaccination. *PLOS Neglected Tropical Diseases* **12(1)**:e0006158 DOI 10.1371/journal.pntd.0006158.