A generalized differential equation compartmental model of infectious disease transmission

Scott Greenhalgh [a, *], Carly Rozins [b]

[a] *Siena College, Department of Mathematics, 515 Loudon Road, Loudonville, NY, 12211, USA*
[b] *York University, Department of Science and Technology Studies, 170 Campus Walk, York University, ON, Canada*

ARTICLE INFO

ABSTRACT

For decades, mathematical models of disease transmission have provided researchers and public health officials with critical insights into the progression, control, and prevention of disease spread. Of these models, one of the most fundamental is the SIR differential equation model. However, this ubiquitous model has one significant and rarely acknowledged shortcoming: it is unable to account for a disease's true infectious period distribution. As the misspecification of such a biological characteristic is known to significantly affect model behavior, there is a need to develop new modeling approaches that capture such information. Therefore, we illustrate an innovative take on compartmental models, derived from their general formulation as systems of nonlinear Volterra integral equations, to capture a broader range of infectious period distributions, yet maintain the desirable formulation as systems of differential equations. Our work illustrates a compartmental model that captures any Erlang distributed duration of infection with only 3 differential equations, instead of the typical inflated model sizes required by traditional differential equation compartmental models, and a compartmental model that captures any mean, standard deviation, skewness, and kurtosis of an infectious period distribution with 4 differential equations. The significance of our work is that it opens up a new class of easy-to-use compartmental models to predict disease outbreaks that do not require a complete overhaul of existing theory, and thus provides a starting point for multiple research avenues of investigation under the contexts of mathematics, public health, and evolutionary biology.

1. Introduction

The compartmental model of Kermack and McKendrick (Kermack and McKendrick, 1932, 1933, 1991) is arguably one of the greatest development in disease modeling. The formulation of this model, in its original form as a system of nonlinear Volterra integral equations (Diekmann et al., 1995), provides a general characterization of the transmission cycle between susceptible individuals and a disease that propagates throughout an environment (Brauer, 2010). Despite this generality, the vast majority of disease modelers prefer differential equation compartmental models (Blackwood & Childs, 2018). While this

* Corresponding author.
  E-mail addresses: sgreenhalgh@siena.edu (S. Greenhalgh), crozins@yorku.ca (C. Rozins).
  Peer review under responsibility of KeAi Communications Co., Ltd.

particular formulation of compartmental models has distinct advantages, such as the non-requirement of specialist knowledge for analysis and well-developed numerical methods for their computation (Roberts et al., 2015), they are a special case of the aforementioned system of nonlinear Volterra integral equations (Hethcote & Tudor, 1980). Specifically, one obtains traditional differential equation compartmental models (TDECM) from the nonlinear Volterra integral equations by imposing only two assumptions: 1) the "total infectivity at time t" (Brauer, 2008, 2010) corresponds to disease incidence, and 2) the duration of infection follows either an exponential or an Erlang distribution. Unfortunately, the vast majority of diseases do not have a duration of infection that follows these distributions (Bailey, 1956; Bailey & Others, 1975; Feng et al., 2007; Gough, 1977; Kelly-Hope & McKenzie, 2009; Lloyd, 2001), and force-fitting such a distributional structure is known to have a massive effect on the behavior and quality of model predictions (Earn, 2000). Regardless of these issues, the compartmental models obtained by the two traditional assumptions have undergone many extensions. A few noteworthy examples include modifying the force of infection to account for the saturation of infection in a population (Hethcote & van den Driessche, 1991; Ponciano & Capistrán, 2011), behavioral characteristics (Alexander & Moghadas, 2005; Liu et al., 1986, 1987; Ruan & Wang, 2003), modification of the recovery rate to capture disease burnout (Greenhalgh et al., 2015a), and the inclusion of additional disease stages (Hethcote & Tudor, 1980). Furthermore, the applications of these models have also grown considerably from just predicting a disease's trajectory. Today, these traditional compartmental models are used to evaluate the health benefits and cost-effectiveness of public health policies and disease interventions (Greenhalgh et al., 2015b, 2018), gauge the potential for disease virulence evolution (Day & Proulx, 2004), predict dominant influenza strains (Casagrandi et al., 2006), investigate the complexities of disease co-infection (Tchuenche et al., 2009), among many others. However, despite this growth in theory and application, the generalization of the very foundational assumptions that simplify systems of nonlinear Volterra integral equations to differential equation compartmental models remains largely undeveloped.

In what follows, we propose new assumptions to reduce nonlinear Volterra integral equations to differential equation compartmental models. The biological motivation for these new assumptions stems from the idea that the quantity of disease in a population may be described by person-days of infection, rather than disease incidence alone. Consequently, we extend current models from solely tracking disease incidence to tracking the number of person-days of infected individuals. To account for tracking person-days of infected individuals, we make the assumptions in our analysis that 1) the total infectivity at time t corresponds to disease incidence and 2) the duration of infection is distributed according to a non-homogeneous analog of the exponential distribution. Under these assumptions, we derive a generalized differential equation compartmental models (GDECM) to track the flow of person-days susceptible to disease, person-days infected to disease, and person-days recovered from disease, in addition to providing model equilibria, and disease reproductive numbers (Heffernan et al., 2005).

Essential in the development of GDECM is the use of survival analysis (supplemental materials). Specifically, the development of GDECM requires the hazard function and the mean residual waiting-time of a distribution (Finkelstein, 2002a, 2002b; Gupta & Bradley, 2003) to describe a time-varying average duration of infection. With this in mind, we demonstrate the consistency of our GDECM to traditional compartmental models when the duration of infection follows an exponential distribution. We also demonstrate GDECM with an infectious period that follows an Erlang distribution and illustrate how this choice has a representation as either a system of 3 ODEs, regardless of the Erlang distribution parameters, or as a larger system of ODEs, which feature a chain of infected individual equations (Hurtado & Kirosingh, 2019; Krylova & Earn, 2013; Lloyd, 2001). Finally, we consider an infectious period that is Pearson distributed. In choosing the Pearson distribution, we develop a model that is capable of accounting for any possible mean, standard deviation, skewness, and kurtosis of the infectious period. Thereby, we provide a simple approach for measuring how altering the infectivity profile of a disease, as described by the first four statistical moments, influences the trajectory of the disease.

2. Methods

In what follows, we develop a GDECM to describe the progression of a disease throughout a population. To obtain such models, we impose new assumptions on the notion of the total infectivity at time t (Brauer, 2008, 2010) used in the integral equations of Kermack and McKendrick. Before deriving GDECM, we briefly overview the quantity of person-days of infection, review the formulation of TDECM from the integral equations of Kermack and McKendrick, and then present the GDECM. In addition, we illustrate the behavior of GDCEM in the absence of transmission, provide equilibria, and their reproductive numbers.

2.1. The infectious period and duration of infection

To derive the duration of infection distribution from the infectious period distribution we consider data on newly infectious individuals, I_{new}, and characterize the infectious period by the survival function $Q(t,z)$. Thus, we obtain the number of infected individuals remaining infectious at time x who where infected at time z as

$$\widehat{I}(x,z) = I_{new}(z)Q(x,z), \ x \geq z,$$

and the total number of infected individuals at time x by

$$I(x) = I_{new}(0)Q(x,0) + \int_0^x I_{new}(w)Q(x,w)dw.$$

On the basis that $\widehat{I}(x,z)/I(x)$ is a normalizing weight, we have that the conditional survival function of the duration of infection distribution for all currently infectious individuals with infection time prior to x is

$$P(t,x) = \frac{\widehat{I}(x,0)}{I(x)} Q\left(t, 0|x, 0\right) + \int_0^x \frac{\widehat{I}(x,z)}{I(x)} Q(t, z|x, z)\, dz, \tag{1}$$

where $Q(t,z|x,z) = \frac{Q(t,z)}{Q(x,z)}$, and $t \geq x \geq z$.

Given the formulation of $P(t,x)$, we assume that the associated hazard rate (Clark et al., 2003a, 2003b; Finkelstein, 2002a, 2002b) does not depend on the time of infection x, namely

$$\eta(t) = -\frac{1}{P(t,x)} \frac{dP}{dt}(t,x),\ t \geq x.$$

Note, this is a slight generalization of the classical assumption for an exponentially distributed infectious period, as when $Q(t,0) = e^{-\gamma(t-0)}$, it follows that $Q(t,z|x,z) = e^{-\gamma(t-x)}$, and $P(t,x) = e^{-\gamma(t-x)}$. Thus, the hazard rate for an exponentially distributed infectious period is $\eta(t) = \gamma$. The generalization we consider is a non-homogeneous analog of the exponential distribution, namely $Q(t,0) = e^{-\int_0^t \eta(w)dw}$, which implies $Q(t,z|x,z) = e^{-\int_x^t \eta(w)dw}$. Under such an assumed form of $Q(t,z)$, we have that $\frac{dP(t,x)}{dt} = -\eta(t)P(t,x)$, and the hazard rate naturally follows.

2.2. Person-days of infection

To provide insight on the quantity person-days of infection we consider a hypothetical epidemic. Solely for illustration purposes, we consider new incidence to be reported weekly (Fig. 1) and the infectious period to be at most 14 days (Fig. 2). Furthermore, akin to the assumption used in the work of Kermack and McKendrick (Brauer, 2008, 2010), we assume that new
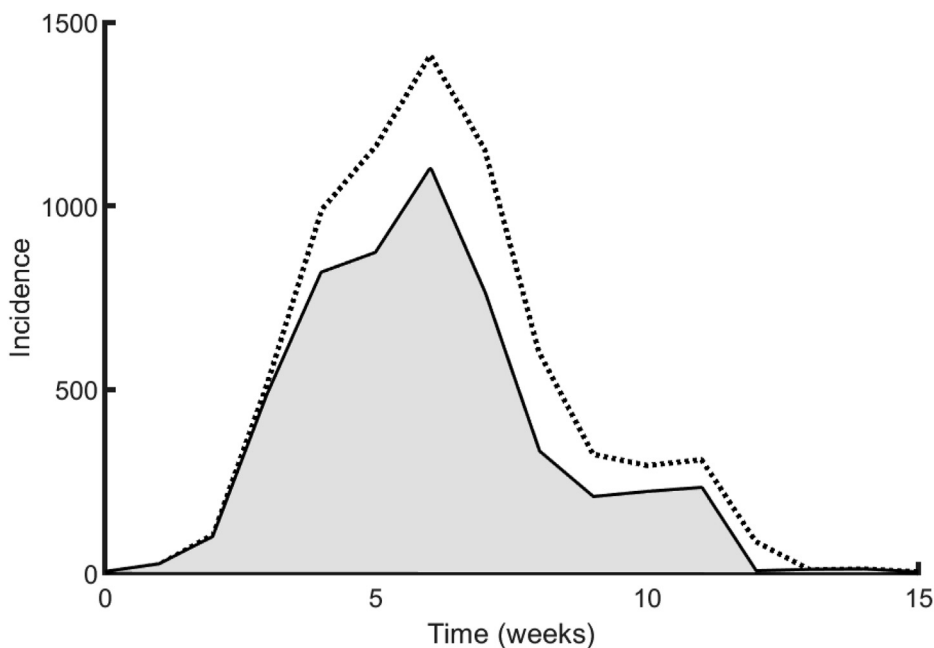

Fig. 1. New and total infectious individuals. New infectious individuals (solid line) and all infectious individuals (new and prior infectious individuals that are infectious) (dotted line).
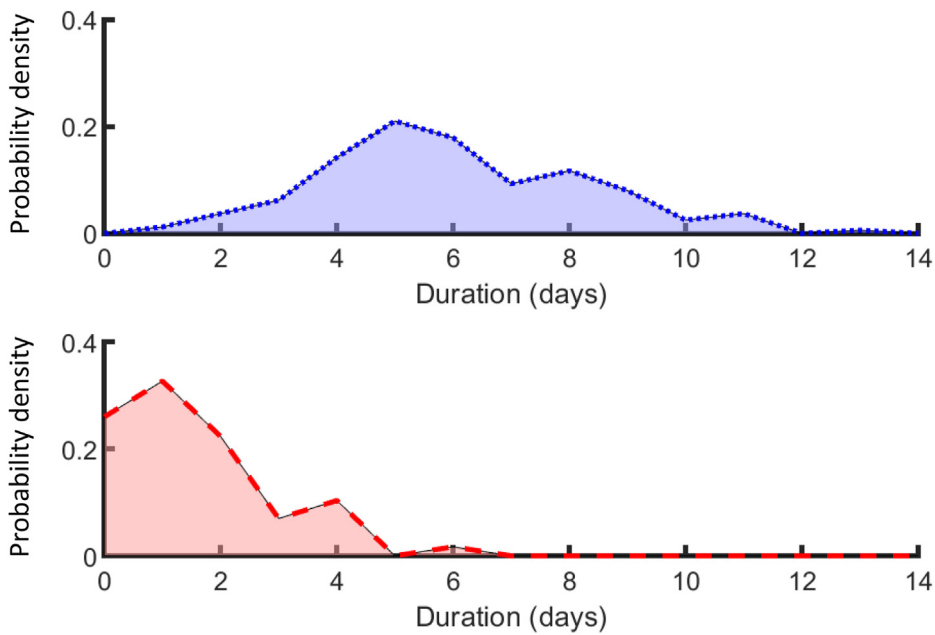
Fig. 2. Infectious periods. Infectious period distribution (top) and the conditional infectious period distribution for individuals whose infection lasts more than 7 days (bottom).
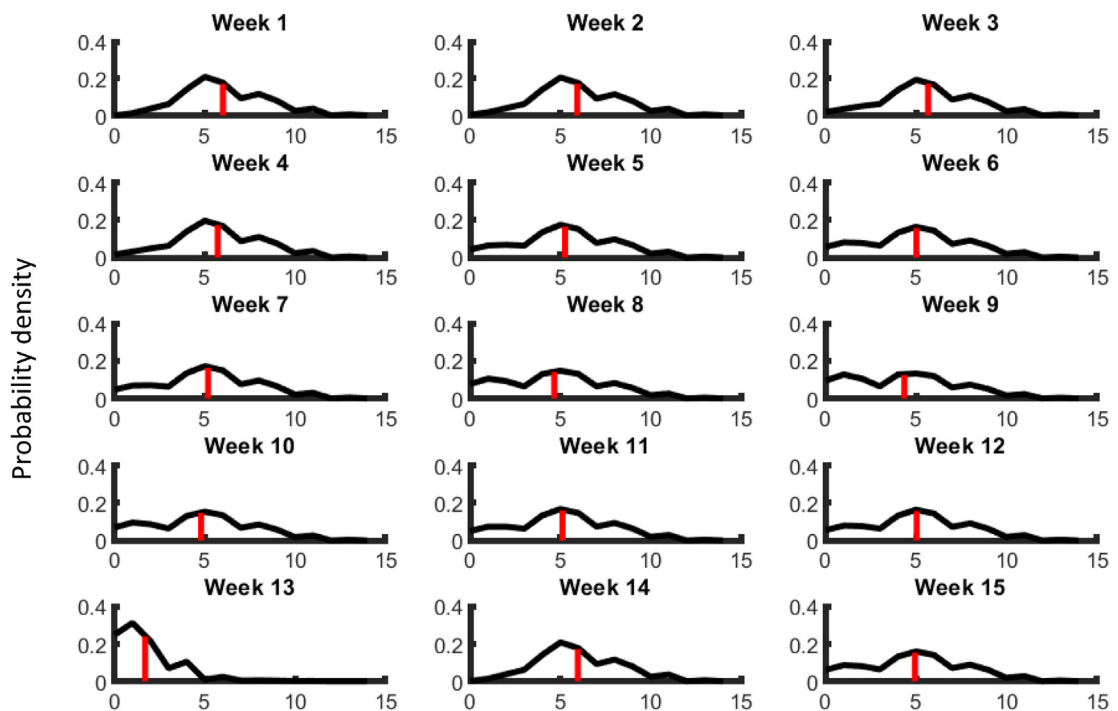

Fig. 3. Weekly duration of infection distribution. The duration of infection distribution and the average duration of infection (red bar) for the weeks 1–15. Each duration of infection distribution is a weighted sum of the infectious period distribution and conditional infectious period distribution, where the weights are based on the number of people currently and priorly infected with disease.

infections occur instantaneously at the end of each week. Given this setup, infectious individuals in a given week consist of newly infected individuals, and individuals that remain infectious from the prior week (Figs. 1 and 2).

Under this setup, we estimate the duration of infection distribution $\frac{dP}{dt}(t, t_i)$ (Fig. 3) from a discretized version of (1),

$$P(t,t_i) \approx \frac{\widehat{I}(t_i,0)}{I(t_i)} Q\left(t|t_i\right) + \sum_{j=i-1}^{i} \frac{\widehat{I}(t_i,t_j)}{I(t_i)} Q\left(t-t_j|t_i-t_j\right),$$

where $t_i = i\Delta t$, $\Delta t = 1$ week, and $t \geq t_i$.

Given the duration of infection distributions for each week the person-days of infection are calculated as the area enclosed by the survival of the infected individuals from time t_i (Fig. 4),

$$\int_{t_i}^{\infty} I(t_i)P(z,t_i)dz = I(t_i)m(t_i),$$

where $m(t_i)$ is the mean residual waiting-time of infection (Gupta & Bradley, 2003) in the ith week.

One alternative to representing person-days of infection as the area enclosed by the survival of the infected individuals at time t_i is by the product of the number of infected individuals and their mean residual waiting-time of infection at time t_i (Fig. 5). The product $I(t_i)m(t_i)$ serves as the basis for constructing GDECM.

2.3. Traditional compartmental models

Consider the general compartmental model of Kermack and McKendrick, as formulated by Brauer et al. (Brauer, 2010; Kermack and McKendrick, 1932, 1933, 1991). For this general compartmental model, the number of susceptible individuals are denoted as S, and the total infectivity at time t is denoted as $\varphi(t)$ (Brauer, 2008, 2010). We also define $\varphi_0(t)$ as the total infectivity at time t of the individuals initially infected with the disease at the beginning of the epidemic. Furthermore, the mean infectivity of an infected individual $t - x$ units of time after the start of the infection is $\pi(t - x)$, where $0 \leq \pi(t - x) \leq 1$, and the mean infectivity of individuals in the population at time t who where initially infected at time x as

$$A(t,x) = \pi(t-x)P(t,x).$$

For simplicity, we assume a constant mean infectivity, $\pi(t - x) = 1$. Given that transmission is a function of both shedding rate and behavior, it would be difficult to quantify changes in mean infectivity throughout an infection. Therefore, the progression of an epidemic throughout a population is described with the integral equations,
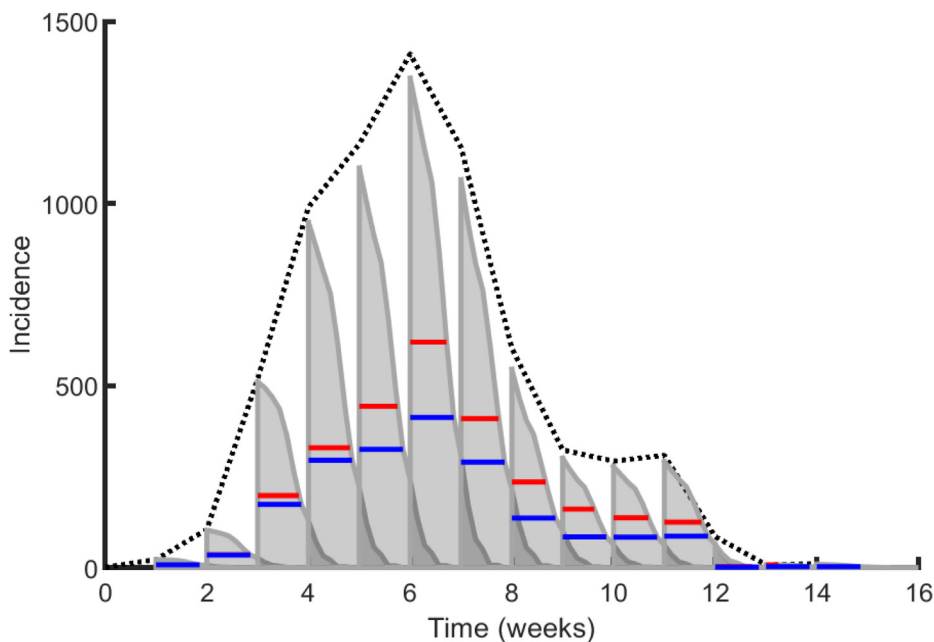

Fig. 4. Person-days of infection. Grey regions correspond to the total days infected people spread disease over the course of an epidemic. Dark grey regions correspond to infectious individuals that remain infectious after one week. Mean residual waiting-time for the time spent infectious for each week (red line), average duration of infection (blue line).
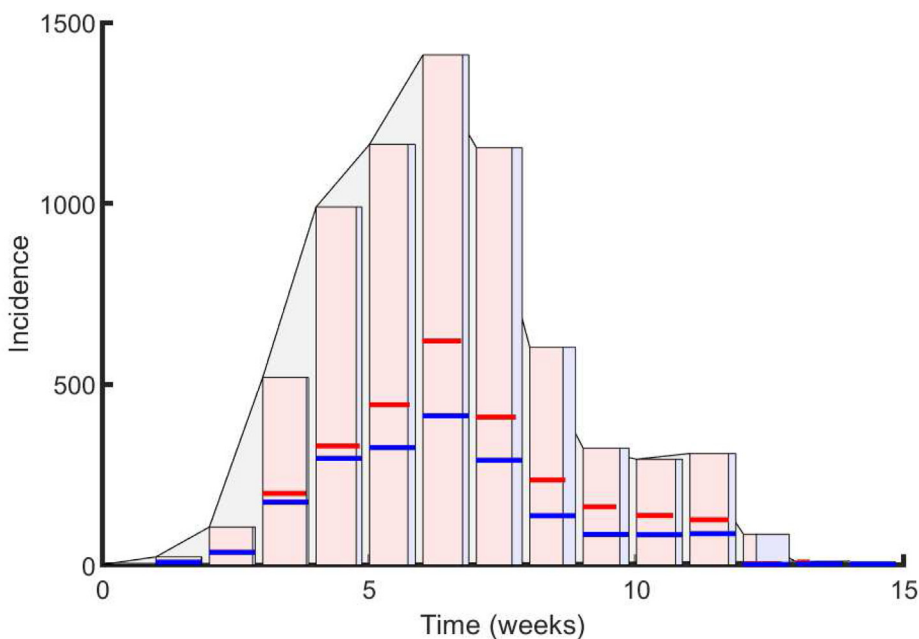
Fig. 5. Person-days of infection. Rectangular red regions correspond to the same number of total days infected people potentially spread disease over the course of an epidemic. Blue rectangular regions illustrate over estimation of person-days of infection using average druation of infection. Mean residual waiting-time for the time spent infectious for each week (red line), average duration of infection (blue line).

$$S(t) = S_0 - \int_0^t \lambda(x)S(x)dx,$$

$$\varphi(t) = \varphi_0(t) + \int_0^t \lambda(x)S(x)P(t,x)dx. \tag{2}$$

Here λ is the force of infection, which we assume to be

$$\lambda(t) = \beta \frac{\varphi(t)}{N},$$

where β is the average number of contacts individuals in a population make per unit of time and S_0 is the initial number of susceptible individuals at the start of the epidemic.

Traditionally, to reduce (2) to a system of differential equations requires that the duration of infection is exponentially distributed, which yields that the conditional survival function is $P(t,x) = e^{-\gamma(t-x)}$, where γ is the recovery rate, or equivalently $1/\gamma$ is the average duration of infection, and 2) that the total infectivity at time t corresponds to the number of infected individuals, $\varphi(t) = I(t)$. Combining these assumptions, along with an additional compartment to track recovered individuals, R, transforms system (2) into

$$S(t) = S_0 - \int_0^t \beta \frac{I(x)}{N} S(x)dx,$$

$$I(t) = I_0 e^{-\gamma t} + \int_0^t \beta \frac{I(x)}{N} S(x) e^{-\gamma(t-x)}dx, \tag{3}$$

$$R(t) = \left(R_0 + I_0 - I_0 e^{-\gamma t}\right) + \int_0^t \beta \frac{I(x)}{N} S(x)\left(1 - e^{-\gamma(t-x)}\right)dx.$$

Here I_0 and R_0 represent the total number of infected individuals and removed individuals at the start of the epidemic. An important feature of system (3) is that it conserves the total population:

$$S(t) + I(t) + R(t) = S_0 + I_0 + R_0 = N.$$

Differentiating system (3) with respect to t, and substituting the integral equation for $I(t)$ for the remaining integrals yields the classic SIR system (Fig. 6):

$$
\begin{aligned}
S'(t) &= -\beta \frac{I(t)}{N} S(t), \\
I'(t) &= \beta \frac{I(t)}{N} S(t) - \gamma I(t), \\
R'(t) &= \gamma I(t).
\end{aligned}
\tag{4}
$$

Alternatively, (2) reduces to a system of differential equations when the duration of infection follows an Erlang distribution, with the conditional survival function

$$P(t,x) = \sum_{j=0}^{k-1} \frac{(\gamma(t-x))^j}{j!} e^{-\gamma(t-x)}, \ t \geq x,$$

where k is a shape parameter that determines the total number of infection stages, $\frac{1}{\gamma}$ is the average duration spent in each stage, and 2) the total infectivity of the disease corresponds to k identical stages (in terms of the average duration spent in each stage) of infected individuals,

$$\varphi = I = \sum_{j=1}^{k} I_j.$$

Combining these assumptions, along with an additional compartment to track recovered individuals, R, transforms system (1) into
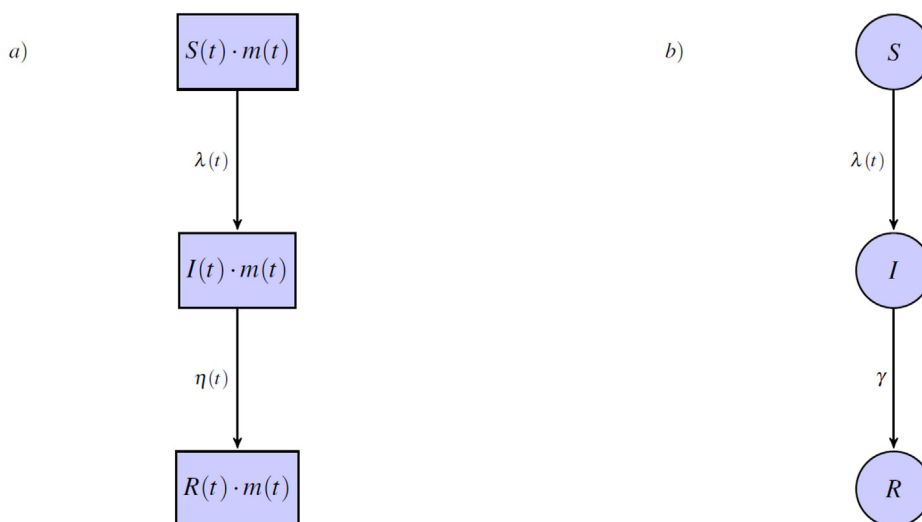

Fig. 6. Compartmental diagrams. a) Flow of suscpetible person-days to infected person-days and recovered person days, and b) Flow of suscpetible persons to infected persons and recovered persons.

$$S(t) = S_0 - \int_0^t \beta \frac{I(x)}{N} S(x) dx,$$

$$I(t) = I_0 \sum_{j=0}^{k-1} \frac{(\gamma t)^j}{j!} e^{-\gamma t} + \int_0^t \beta \frac{I(x)}{N} S(x) \sum_{j=0}^{k-1} \frac{(\gamma(t-x))^j}{j!} e^{-\gamma(t-x)} dx, \tag{5}$$

$$R(t) = R_0 + I_0 - I_0 \sum_{j=0}^{k-1} \frac{(\gamma t)^j}{j!} e^{-\gamma t} + \int_0^t \beta \frac{I(x)}{N} S(x) \left(1 - \sum_{j=0}^{k-1} \frac{(\gamma(t-x))^j}{j!} e^{-\gamma(t-x)} \right) dx.$$

Equivalently, if the linear chain trick is applied (see supplemental materials for details) (Lloyd, 2001), we have that

$$I_j(t) = I_j(0) \frac{(\gamma t)^{j-1}}{(j-1)!} e^{-\gamma t} + \int_0^t \beta \frac{I(x)}{N} S(x) \frac{(\gamma(t-x))^{j-1}}{(j-1)!} e^{-\gamma(t-x)} dx,$$

where $I(t) = \sum_j I_j(t)$.

An important feature of system (5) is that it conserves the total population:

$$S(t) + I(t) + R(t) = S_0 + \sum_j I_{0,j} + R_0 = N.$$

Differentiating system (5) with respect to t, and substituting $\gamma I_j(t)$ as needed yields the classic SI^{k}R system:

$$S'(t) = -\beta \frac{I(t)}{N} S(t),$$

$$I_1'(t) = \beta \frac{I(t)}{N} S(t) - \gamma I_1(t),$$

$$I_j'(t) = \gamma I_{j-1}(t) - \gamma I_j(t), \tag{6}$$

$$R'(t) = \gamma I_k(t),$$

where $I(t) = \sum_{j=1}^{k} I_j(t)$.

2.4. Generalized differential equation compartmental models

Before presenting the GDECM, we briefly contextualize its differences in comparison to traditional compartmental models. The traditional compartmental models of (4) and (6) track the changes in the number of individuals susceptible, infected, and recovered from disease, for a disease circulating in a population. For traditional compartmental models, the individuals infected with disease recover from infection at a constant rate, γ, which is analgous to a constant average duration of infection of $1/\gamma$ (Fig. 1). The premise of GDECM is to consider tracking the change in the number of person-days of infection (Fig. 2). Specifically, we characterize person-days of infection as the individuals infected with a disease multiplied by their average duration of time spent in the infected state at time t (Fig. 1).

To obtain GDECM, we generalize the assumptions used to formulate system (5). We first assume the total infectivity at time t is the number of infected individuals at time t :

$$\varphi(t) = I(t),$$

In addition, we assume that the duration of infection distribution corresponds to a non-homogeneous analog of the exponential distribution, as given by the conditional survival function

$$P(t,x) = e^{-\int_x^t \eta(z) dz}, \tag{7}$$

where $t \geq x$.

The rate of removal from the infectious state, $\eta(t)$, is directly linked to the mean residual waiting-time $m(t)$. Specifically, one can determine $\eta(t)$ from $m(t)$ through

$$\eta(t) = \frac{m'(t) + 1}{m(t)},$$

and the reciprocal relationship by

$$m(t) = \mu - t + g(t)\eta(t), \tag{8}$$

provided that $g(t)$ satisfies (Gupta & Bradley, 2003)

$$g'(t)\frac{dP}{dt} + g(t)\frac{d^2P}{dt^2} = (\mu - t)\frac{dP}{dt}.$$

The functions $\eta(z)$ and $m(z)$ possess the standard properties of hazard rates and mean residual waiting-time functions, including $\eta(z) \geq 0$, $m(z) \geq 0, m'(z) \leq 0,$ and $m(0) = \mu$ (Finkelstein, 2002a, 2002b).

Under this setup, we consider person-days susceptible to disease, person-days infectious with disease, and person-days recovered from disease to be governed by,

$$S(t)m(t) = Nm(t) - (N - S_0)\mu - \int_0^t \lambda(x)m(x)S(x)dx,$$

$$I(t)m(t) = I_0 P(t, 0)\mu + \int_0^t \lambda(x)m(x)S(x)P(t, x)dx, \tag{9}$$

$$R(t)m(t) = (R_0 + I_0 - I_0 P(t, 0))\mu + \int_0^t \lambda(x)m(x)S(x)(1 - P(t, x))dx,$$

where $m(0) = \mu$ by definition (Finkelstein, 2002a, 2002b), $\lambda(x)$ is the force of infection, and $P(t, x)$ is given by (5).

Adding the equations of system (9) together, it follows that

$$(S(t) + I(t) + R(t) - N)m(t) = (S_0 + I_0 + R_0 - N)\mu.$$

Thus, given $N = S_0 + I_0 + R_0$ and $m(t) \neq 0$ for all t, we have that the total population is conserved:

$$S(t) + I(t) + R(t) = N.$$

Imposing (7) on (9), taking the time derivative of system (9), assuming $\lambda(x) = \beta\frac{I(x)}{N}$ and applying Leibniz rule for the derivatives of integrals as needed, we obtain

$$S'(t)m(t) + S(t)m'(t) = Nm'(t) - \beta\frac{I(t)}{N}m(t)S(t),$$

$$I'(t)m(t) + I(t)m'(t) = -\eta(t)I_0 P(t, 0)\mu - \eta(t)\int_0^t \beta\frac{I(x)}{N}m(x)S(x)P(t, x)dx$$

$$+\beta\frac{I(t)}{N}m(t)S(t), \tag{10}$$

$$R'(t)m(t) + R(t)m'(t) = \eta(t)I_0 P(t, 0)\mu + \eta(t)\int_0^t \beta\frac{I(x)}{N}m(x)S(x)P(t, x)dx,$$

where $P(t, x)$ is given by (7).

Substituting (9) into (10) to eliminate the integrals, and isolating for each of $S'm + Sm'$, $I'm + Im'$, and $R'm + Rm'$, we obtain the compartmental model (Fig. 6):

$$S'(t)m(t) + S(t)m'(t) = Nm'(t) - \beta\frac{I(t)}{N}m(t)S(t),$$

$$I'(t)m(t) + I(t)m'(t) = \beta\frac{I(t)}{N}m(t)S(t) - \left(\frac{m'(t)+1}{m(t)}\right)I(t)m(t), \tag{11}$$

$$R'(t)m(t) + R(t)m'(t) = \left(\frac{m'(t)+1}{m(t)}\right)I(t)m(t).$$

By construction, system (11) is similar to its TDECM counterpart. However, there are some subtle differences. The terms involving $m'(t)$ on the left-hand side do not represent inflows that increase the number of susceptible, infected, or recovered individuals. These terms, which are non-positive due to $m'(t) \leq 0$ (Finkelstein, 2002a, 2002b; Gupta & Bradley, 2003), account for the change in person-days caused by a decrease in the average duration of infection. In a similar manor, the term $Nm'(t)$ on the right-hand side does not represent an inflow of susceptible individuals, as it accounts for the change in scale of total person-days over time.

2.5. Behavior in the absence of transmission

In the absence of transmission ($\beta = 0$), the infectious equation in (11) reduces to

$$I'(t)m(t) + I(t)m'(t) = -\left(\frac{m'(t)+1}{m(t)}\right)I(t)m(t).$$

Solving for $I(t)$, it follows that

$$I(t) = I_0 \exp\left(-\int_0^t \frac{2m'(z)+1}{m(z)}dz\right) = I_0 P(t,0)\exp\left(-\int_0^t \frac{m'(z)}{m(z)}dz\right).$$

Note, the extra factor $\exp\left(-\int_0^t \frac{m'(z)}{m(z)}dz\right)$ originates from $I(t)m'(t)$, and thus reflects the time variability in the mean residual waiting-time of the initially infected individuals.

2.6. Equilibria, the basic reproductive number $\mathscr{R}_0$, and the effective reproductive number $\mathscr{R}_e$

The compartmental model (11) potentially have several equilibria. As is standard, (11) possess the standard disease-free equilibrium:

$$S^* = N, \ I^* = 0, \ R^* = 0. \tag{12}$$

In addition, an equilibrium where the infection was exhausted, leaving susceptible individuals that escaped infection, and recovered individuals:

$$\widehat{S} = S_\infty, \ \widehat{I} = 0, \ \widehat{R} = R_\infty. \tag{13}$$

Turning our attention to the reproductive numbers of the disease, we consider the basic reproductive number using the next-generation method (van den Driessche & Watmough, 2002; van den Driessche, 2017), where $\mathscr{F} = \left(\frac{\beta}{N}S\right)$ and $\mathscr{V}^{-1} = \left(\frac{m(t)}{m'(t)+1}\right)$ It follows that

$$\mathscr{R}_0 = \rho\left(\mathscr{F}|_{DFE} \cdot \mathscr{V}^{-1}\Big|_{DFE}\right) = \beta\frac{m_{DFE}}{m'_{DFE}+1}. \tag{14}$$

Here $\rho(\cdot)$ is the spectral radius, $\mathscr{F}|_{DFE} \cdot \mathscr{V}^{-1}_{DFE}$ is the next-generation matrix evaluated at the disease-free equilibrium, and $m_{DFE} \in \left[\lim_{t\to\infty} m(t), m(0)\right]$ is the constant value of $m(t)$ at the disease-free equilibrium.

Following a similar setup, we consider the effective reproductive number, in the face of time variability of $m(t)$, to be

$$\mathscr{R}_e(t) = \beta \frac{m(t)}{m'(t) + 1} \frac{S(t)}{N}. \tag{15}$$

To summarize, in this section we illustrated that GDECM possesses similar biological properties to TDCEM. Specifically, GDECM's behavior in the absence of transmission, its ability to conserve the total population, equilibria, and reproductive numbers all closely resemble similar features in TDCEM. Therefore, to further compare GDECM with TDCEM, in the following section we consider the GDECM based on the duration of infection distributions of TDECM, namely the exponential or Erlang distributions, before demonstrating that GDECM exists for infectious periods that are Pearson distributed.

3. Special cases of GDECM

We now consider the duration of infection distributions used in traditional differential equation compartmental models to obtain systems of differential equations, namely the exponential distribution, and the Erlang distribution. In addition, we illustrate a GDECM that accounts for any mean, standard deviation, skewness, and excess kurtosis, by assuming that the duration of infection is Pearson distributed.

3.1. The exponential distribution

Assuming the infectious period is exponentially distributed, $Q(t) = e^{-\gamma t}$, then the duration of infection is also exponentially distributed, and the conditional survival function for the epidemic is $P(t,x) = e^{-\gamma(t-x)}$. Under this assumption, we have that (Gupta & Bradley, 2003),

$$\eta(t) = -\frac{1}{P(t,x)} \frac{dP(t,x)}{dt} = \gamma, \tag{16}$$

and

$$g(t) = \frac{1}{\gamma} t. \tag{17}$$

Solving (8) under the assumption of (17) yields,

$$m(t) = \frac{1}{\gamma}. \tag{18}$$

Substituting (18) into (11), and multiplying through by γ, we arrive at the traditional compartmental models:

$$\begin{aligned}
S'(t) &= -\beta \frac{I(t)}{N} S(t), \\
I'(t) &= \beta \frac{I(t)}{N} S(t) - \gamma I(t), \\
R'(t) &= \gamma I(t).
\end{aligned} \tag{19}$$

Finally, because $m(t) = \frac{1}{\gamma} = m(0) = \mu$, the basic reproductive number obtained by next-generation method (van den Driessche, 2017) is

$$\mathscr{R}_0 = \frac{\beta}{\gamma} \tag{20}$$

3.2. The Erlang distribution

If the infectious period is Erlang distributed with survival function $Q(t) = \sum_{j=0}^{k-1} \frac{(\gamma t)^j}{j!} e^{-\gamma t}$, then necessarily $P(t,0) = Q(t)$. It follows that the associated conditional survival function of the duration of infection distribution is

$$P(t,x) = \left(\sum_{j=0}^{k-1} \frac{(\gamma x)^j}{j!} e^{-\gamma x}\right)^{-1} \sum_{j=0}^{k-1} \frac{(\gamma t)^j}{j!} e^{-\gamma t}, \quad t \geq x. \tag{21}$$

Note, this formulation of $P(t,x)$ differs from that used in the derivation of TDECM, which bases $P(t,x)$ on the convolution of exponential random variables.

From (9) and (21), we have that

$$\eta(t) = \left(\sum_{j=0}^{k-1} \frac{(\gamma t)^j}{j!} e^{-\gamma t}\right)^{-1} \frac{\gamma(\gamma t)^{k-1}}{(k-1)!} e^{-\gamma t} = \left(\sum_{j=0}^{k-1} \frac{(\gamma t)^j}{j!}\right)^{-1} \frac{\gamma(\gamma t)^{k-1}}{(k-1)!}. \tag{22}$$

Using $g(t) = \frac{1}{\gamma} t$ for an Erlang distribution (Gupta & Bradley, 2003), it follows from (8) that the mean residual waiting-time:
$m(t) = \frac{k}{\gamma} - t + \frac{1}{\gamma} t \left(\sum_{j=0}^{k-1} \frac{(\gamma t)^j}{j!}\right)^{-1} \frac{\gamma(\gamma t)^{k-1}}{(k-1)!}.$

For the ease of presentation, we rewrite $m(t)$ and its derivative with respect to t as

$$m(t) = \frac{k}{\gamma} - t + \frac{1}{\gamma} \frac{(\gamma t)^k}{\Gamma(k, \gamma t)} e^{-\gamma t}$$

and

$$m'(t) = -1 + \left(\frac{(\gamma t)^k e^{-\gamma t}}{\Gamma(k, \gamma t)} + \gamma t - k\right) \frac{(\gamma t)^{k-1}}{\Gamma(k, \gamma t)} e^{-\gamma t}. \tag{23}$$

Using (23) in system (11), where the left-hand sides are left as $\frac{d}{dt}(S(t)m(t))$ and $\frac{d}{dt}(I(t)m(t))$ for the ease of presentation, we obtain a compartmental model comprised of three equations, regardless of the value of k from the Erlang distribution:

$$\frac{d}{dt}\left(S(t)\left(\frac{k}{\gamma} - t + \frac{1}{\gamma} \frac{(\gamma t)^k}{\Gamma(k, \gamma t)} e^{-\gamma t}\right)\right) = N\left(-1 + \left(\frac{(\gamma t)^k e^{-\gamma t}}{\Gamma(k, \gamma t)} + \gamma t - k\right) \frac{(\gamma t)^{k-1}}{\Gamma(k, \gamma t)} e^{-\gamma t}\right) - \beta \frac{I(t)}{N}\left(\frac{k}{\gamma} - t + \frac{1}{\gamma} \frac{(\gamma t)^k}{\Gamma(k, \gamma t)} e^{-\gamma t}\right) S(t),$$

$$\frac{d}{dt}\left(I(t)\left(\frac{k}{\gamma} - t + \frac{1}{\gamma} \frac{(\gamma t)^k}{\Gamma(k, \gamma t)} e^{-\gamma t}\right)\right) = \beta \frac{I(t)}{N}\left(\frac{k}{\gamma} - t + \frac{1}{\gamma} \frac{(\gamma t)^k}{\Gamma(k, \gamma t)} e^{-\gamma t}\right) S(t) - \left(\frac{\left(\frac{(\gamma t)^k e^{-\gamma t}}{\Gamma(k, \gamma t)} + \gamma t - k\right)}{\left(\frac{k}{\gamma} - t\right) e^{\gamma t} \frac{\Gamma(k, \gamma t)}{(\gamma t)^{k-1}} + t}\right) I(t),$$

and
$$R(t) = N - S(t) - I(t),$$

$$\tag{24}$$

Alternatively, by applying the 'linear chain trick' (Hurtado & Kirosingh, 2019; Krylova & Earn, 2013; Lloyd, 2001) (supplemental materials) to the system of three equations, we equivalently characterize $\frac{d}{dt}(I(t)m(t))$ as:

$$I'_1(t)m(t) + I_1(t)m'(t) = -\gamma I_1(t)m(t) + \frac{\beta}{N}\left(\sum_{i=1}^{k} I_i(t)\right) m(t)S(t) \left(\sum_{j=0}^{k-1} \frac{(\gamma t)^j}{j!}\right)^{-1}, \tag{25}$$

and

$$I'_j(t)m(t) + I_j(t)m'(t) = -\gamma I_j(t)m(t) + \gamma I_{j-1}(t)m(t) + \frac{\beta}{N}\left(\sum_{i=1}^{k} I_i(t)\right) m(t)S(t) \left(\sum_{i=0}^{k-1} \frac{(\gamma t)^i}{i!}\right)^{-1} \frac{(\gamma t)^{j-1}}{(j-1)!}. \tag{26}$$

Thus, through (21), (22), and (24), we either reduce system (11) to a system of 3 differential equations based on an infectious period that is Erlang distributed, or through the 'linear chain trick' to a system of $k+2$ differential equations (i.e. (25)-(26), S' and R').

Note, an important distinction between (25)–(26) and the TDCEM analog is that new infections enter into any given infectious state because of the terms $\frac{\beta}{N}(\sum_{i=1}^{k} I_i(t))m(t)S(t)\left(\sum_{i=0}^{k-1}\frac{(\gamma t)^i}{i!}\right)^{-1}\frac{(\gamma t)^{j-1}}{(j-1)!}$, whereas the TDECM (6) requires all new infections to enter the first stage of infection.

In regards to the basic reproductive numbers, we have that

$$\mathcal{F} = \frac{\beta}{N}m(t)S(t)\left(\sum_{i=0}^{k-1}\frac{(\gamma t)^i}{i!}\right)^{-1}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \gamma t & \gamma t & \cdots & \gamma t \\ \vdots & \vdots & \vdots & \vdots \\ \frac{(\gamma t)^{k-1}}{(k-1)!} & \frac{(\gamma t)^{k-1}}{(k-1)!} & \frac{(\gamma t)^{k-1}}{(k-1)!} & \frac{(\gamma t)^{k-1}}{(k-1)!} \end{pmatrix},$$

and

$$\mathcal{V} = \gamma m(t)\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & -1 & 1 & 0 \\ 0 & \cdots & \cdots & -1 & 1 \end{pmatrix}$$

Evaluating $\mathcal{F}$ and $\mathcal{V}$ at the disease-free equilibrium, and when $t = 0$, it follows that

$$\rho\left(\mathcal{F}\Big|_{\substack{DFE \\ t=0}} \cdot \mathcal{V}^{-1}\Big|_{\substack{DFE \\ t=0}}\right) = \beta\frac{k}{\gamma}.$$

and as $t \to \infty$, it follows that

$$\rho\left(\lim_{t\to\infty}\mathcal{F}|_{DFE} \cdot \mathcal{V}^{-1}\Big|_{DFE}\right) = \beta\frac{1}{\gamma}.$$

Thus, depending on how the value of m_{DFE} is defined, we have the basic reproductive number

$$\mathcal{R}_0 = \beta\frac{m_{DFE}}{m'_{DFE}+1} \in \left[\beta\frac{1}{\gamma}, \ \beta\frac{k}{\gamma}\right]. \tag{29}$$

3.3. The Pearson distribution

If the infectious period is Pearson distributed, then the associated conditional survival function of the duration of infection distribution does not possess a general closed-form. Thus, we define the Pearson distribution by

$$\frac{d^2}{dt^2}P(t,x) = h(t;\mu,\sigma,\gamma_1,\gamma_2)\frac{d}{dt}P(t,x), \tag{30}$$

where

$$h(t;\mu,\sigma,\gamma_1,\gamma_2) = \frac{\left(-12\gamma_1^2+10\gamma_2+12\right)(t-\mu)+(\gamma_1\gamma_2+6\gamma_1)\sigma}{\left(3\gamma_1^2-2\gamma_2\right)(t-\mu)^2-\sigma\gamma_1(\gamma_2+6)(t-\mu)+\left(3\gamma_1^2-12-4\gamma_2\right)\sigma^2}$$

$$= \frac{t-\mu-d}{c(t-\mu)^2+b(t-\mu)+a}, \tag{31}$$

with $a = \frac{\sigma^2(3\gamma_1^2-4\gamma_2-12)}{12\gamma_1^2-10\gamma_2-12}$, $b = \frac{-\sigma\gamma_1(\gamma_2+6)}{12\gamma_1^2-10\gamma_2-12}$, $c = \frac{3\gamma_1^2-2\gamma_2}{12\gamma_1^2-10\gamma_2-12}$, and $d = \frac{(\gamma_1\gamma_2+6\gamma_1)\sigma}{12\gamma_1^2-10\gamma_2-12}$.

Note, μ, σ, γ_1, and γ_2 are the mean, standard deviation, skewness, and excess kurtosis of the infectious period distribution (Gupta & Bradley, 2003; Pearson, 1893, 1895).

To obtain the mean residual waiting-time in terms of the hazard function, we have that (Gupta & Bradley, 2003; Papathanasiou, 1995):

$$g(t) = \frac{1}{1 - 2c}\left(a + b(t - \mu) + c(t - \mu)^2\right). \tag{32}$$

It follows from (18) that the mean residual waiting-time is:

$$m(t) = \mu - t + \left(\frac{1}{1 - 2c}\left(a + b(t - \mu) + c(t - \mu)^2\right)\right)\eta(t), \tag{33}$$

where $c \neq 1/2$.

To ensure the mean residual waiting-time is finite we apply l'Hôpital's rule, using the facts that $\int_0^t P(z, 0)dz \to \mu$, $P(t, 0) \to 0, \frac{dP}{dt} \to 0, \frac{d^2P}{dt^2} \to 0$ as $t \to \infty$, to obtain

$$\lim_{t \to \infty} m(t) = \lim_{t \to \infty} \frac{\mu - \int_0^t P(z, 0)dz}{P(t, 0)} = \lim_{t \to \infty} -\frac{P(t, 0)}{\frac{dP}{dt}} = \lim_{t \to \infty} -\frac{\frac{dP}{dt}}{\frac{d^2P}{dt^2}} = \lim_{t \to \infty} -\frac{1}{h(t; \mu, \sigma, \gamma_1, \gamma_2)}. \tag{34}$$

Thus, we require that $3\gamma_1^2 - 2\gamma_2 = 0$ to force the terms involving t^2 to drop out. Under this assumption, we have that

$$\lim_{t \to \infty} -\frac{1}{h\left(t; \mu, \sigma, \gamma_1, \frac{3}{2}\gamma_1^2\right)} = \frac{1}{2}\sigma\gamma_1, \tag{35}$$

Given (35) we evaluate $\mathscr{F}$ and $\mathscr{V}$ at the disease-free equilibrium, as $t \to \infty$, yielding

$$\rho\left(\lim_{t \to \infty} \mathscr{F}|_{DFE} \cdot \mathscr{V}^{-1}\Big|_{DFE}\right) = \beta\frac{1}{2}\sigma\gamma_1. \tag{36}$$

Similarly, evaluating $\mathscr{F}$ and $\mathscr{V}$ at the disease-free equilibrium when $t = 0$, it follows that
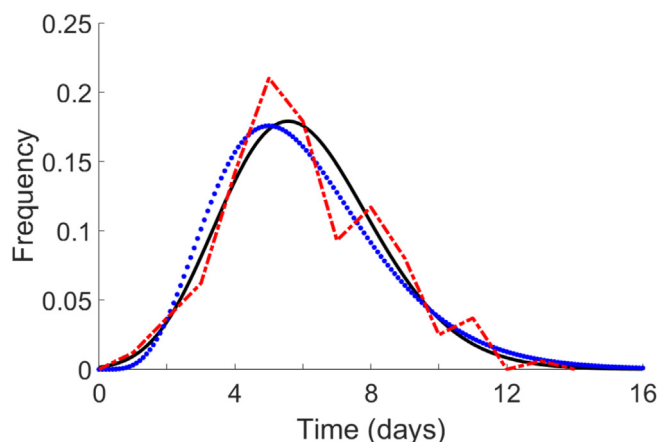

Fig. 7. Measles infectious period. Infectious period distribution based on observed incidence data (red dash-dot line), fit Pearson distribution (black solid line) based on sample moments $\bar{x} = 6.01, s = 2.26, \hat{\gamma}_1 = 0.39, \hat{\gamma}_2 = 3.22$, and fit Erlang distribution (blue dotted line) based a shape parameter of $n = 6$, and rate parameter of $\bar{\gamma} = 0.999$.
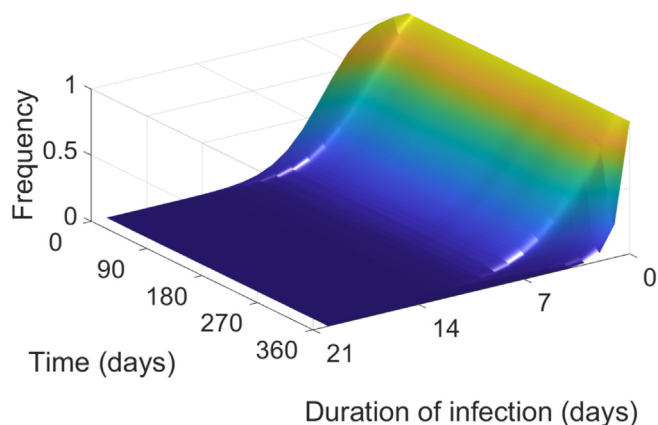
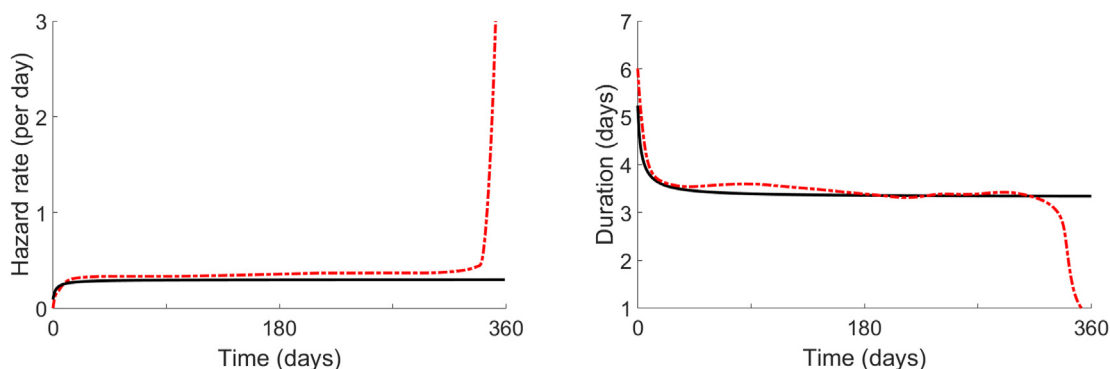**Fig. 8.** Survival function for individuals initially infected at a given day throughout the epidemic.

Fig. 9. Hazard rate $\eta(t)$ and mean residual waiting-time $m(t)$.

$$\rho\left(\mathscr{F}\Big|_{\substack{DFE \\ t=0}} \cdot \mathscr{V}^{-1}\Big|_{\substack{DFE \\ t=0}} \right) = \beta\mu. \tag{37}$$

It follows that the basic reproductive number is

$$\mathscr{R}_0 = \beta\,\frac{m_{DFE}}{m'_{DFE}+1} \in \left[\beta\frac{1}{2}\sigma\gamma_1, \beta\mu\right], \tag{38}$$

depending on the value of m_{DFE}.

Note, the lower and upper bounds of (38) are consistent with the next-generation method estimate for the basic reproductive numbers (20) and (29), as the Erlang distribution has moments $\mu = \frac{k}{\gamma}$, $\sigma = \frac{\sqrt{k}}{\gamma}$ and $\gamma_1 = \frac{2}{\sqrt{k}}$, which implies $\frac{1}{2}\sigma\gamma_1 = \frac{1}{\gamma}$.

To summarize, in this section, we illustrated GDECM is equivalent to TDECM when the duration of infection is exponentially distributed. In addition, GDECM is similar to TDECM when the duration of infection is Erlang distributed. Although, the GDECM advantageously possesses formulations as either a system of $k+2$ differential equations, as would be required from TDECM, or as a system of 3 differential equations. We concluded by demonstrating GDECM exist for non-Erlang distributions, namely the family of Pearson distribution, and now provide an application of such a GDECM to describe measles transmission in Iceland in the next section.
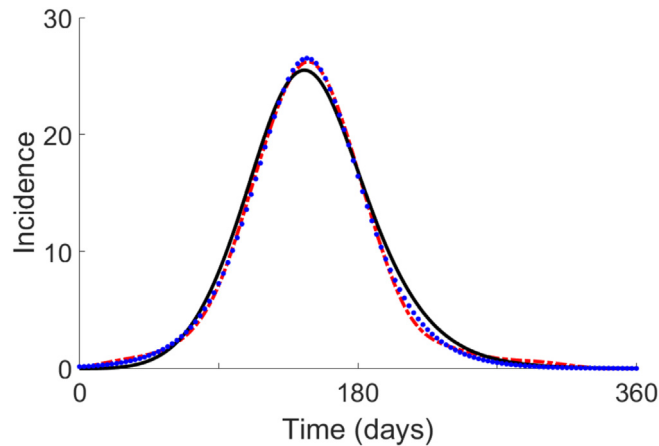
Fig. 10. Predicted measles incidence in Reykjavik Iceland during 1924. Trajectories of incidence are based on interpolation of data (dash-dotted red line) (Cliff et al., 1981), and simulated trajectory with the GDECM (black solid line), and simulated trajectory with SI^{k}R model (blue dotted line).

4. Application of methodology

To illustrate the utility of our approach, we model measles outbreaks based on time series data of measles incidence in Iceland from 1924 to 1928 (Cliff et al., 1981), taking into account classical data on the infectious period of measles (Simpson, 1952).

We use sample moments of the mean, standard deviation, skewness, and kurtosis to estimate the infectious period distribution of measles, under the assumption that it is Pearson distributed, and compare the GDECM to a traditional compartmental model that features an infectious period that is Erlang distributed (Fig. 7).

For the GDECM we determine the duration of infection distribution (1). Assuming that $P(\tau, t)$ is Pearson distributed with $g_2 = \frac{3}{2}g_1^2$ (Fig. 8), it follows that the mean, standard deviation, skewness, and kurtosis are 5.98, 4.15, 1.62, and 3.93 days, respectively. Given these moments, the hazard rate and mean residual waiting-time naturally follow (Fig. 9). Finally, fitting system (11) to data using a least-squares procedure yields $\beta \approx 1.9061$ (Fig. 10).

When $P(\tau, t)$ is Erlang distributed, we have that the mean and standard deviation are 5.98, and 2.44 days, respectively. Given these values, it follows that the classical SIR model, with a duration of infection that is Erlang distributed, is given by

$$\frac{dS}{dt} = -\frac{\tilde{\beta}}{N} \sum_{j=1}^{n} I_j S, \ \frac{dI_1}{dt} = \frac{\beta}{N} \sum_{j=1}^{n} I_j S - \tilde{\gamma} I_1, \ \frac{dI_k}{dt} = \tilde{\gamma} I_{k-1} - \tilde{\gamma} I_k, \tag{39}$$

where $n = 6, \mu = 5.98$ days, $\tilde{\gamma} = \frac{n}{\mu}$. From a least-squares fit (Fig. 10) we have that $\tilde{\beta} \approx 3.09$, $k \in \{1, ..., 6\}$, where the total population of Reykjavik (in 1924) is estimated to be $N = 24134$ (Greenhalgh & Day, 2017).

Given the model parameters, the possible range of basic reproductive number (36) of measles for the GDECM (11) given by the next-generation method is

$$\widehat{\mathscr{R}}_0 \in \left[\frac{1}{2}\beta\sigma\gamma_1, \ \beta\mu \right] = [6.3291, \ 11.3946],$$

and for the traditional compartmental model (39), we have that

$$\widehat{\mathscr{R}}_0 = \tilde{\beta}\frac{n}{\tilde{\gamma}} = 18.54.$$

5. Discussion

In this work, we presented a class of GDECM by modifying the classical assumptions that simplify nonlinear Volterra integral equations into systems of differential equations. To do this, we generalize the notion of compartmental models to describe person-days of infection, rather than infected persons. We illustrate the consistency of our class of models with traditional compartmental models for exponential and Erlang distributed infectious periods, present a new class of differential equation compartmental models based on a Pearson distributed infectious period, and provide equilibria and basic reproductive numbers for our approach.

The requirement that the duration of infection follows an exponential distribution is often a source of weakness with regards to the biological validity of differential equation compartmental models. While the extension of such models through the linear chain trick (Hurtado & Kirosingh, 2019; Krylova & Earn, 2013; Lloyd, 2001) to a duration of infection that follows the Erlang distribution alleviates this weakness to some degree, it does so at the cost of inflating the size of the compartmental model and thereby increasing the computational complexity of the system. Our new class of models avoids this inflation while retaining the benefits of having an infectious period that follows an Erlang distribution. Thereby our new class of models offers an approach to reduce model complexity in an era when the complexity of compartmental models is ever increasing. Furthermore, if one generalizes the concept of a function to include the survival function of the Gamma distribution, our new class of models advantageously accounts for any Gamma distribution parameter values, including non-integer cases.

A substantial advantage of the new class of models is that they are ODE based. Therefore, like traditional compartmental models, they do not require specialist knowledge to use and possess well-developed numerical methods for their simulation. Furthermore, the theoretical extensions of GDECM to include alternative formulations of the force of infection, state-dependent recovery rates, in addition to the inclusion of additional disease compartments are easy to implement. In addition, many of the applications of the traditional models, such as the study of multi-strain dynamics, health benefit analysis, and cost-effectiveness analysis, should naturally carry over without the need to reinvent the procedures of each analysis.

A potentially fruitful avenue for future applications of our new class of models is in the study of virulence and disease evolution. To elaborate, because GDECM includes both the quantity of infected individuals and their duration of infection, it may serve as a better paradigm for investigating selective pressures that pathogens face, at least relative to traditional compartmental models. Similarly, our idea to track both the quantity of infected individuals along with their duration of infection could be adapted to study species competition, as a similar modification should provide stronger intuition on species fitness.

A surprising outcome of our work is the discovery that the lower bound provided by the next-generation method estimate of the basic reproductive number for the Pearson distributed example depends on standard deviation and skewness, instead of the mean. As standard deviation and skewness indicate the spread and lean of a distribution, it seems reasonable that their combination makes for a decent proxy for the location of the middle of a distribution. While this could be a consequence of assuming that the duration of infection follows the Pearson distribution, it also highlights a potentially new approach to bound the basic reproductive number for a disease directly from data.

The use of an infectious period that is Pearson distributed may also open up an interesting avenue into bifurcation analysis. In particular, with modifications to incorporate demographic turnover or loss of immunity and the use of the Pearson distribution, one could investigate if a relationship exists between the occurrence of periodic cycles and the first four moments of the infectious period distribution. Thereby, one may be able to gain intuition as to whether bifurcations are likely to occur simply by examining statistical moments and the disease transmission rate. In addition, through such modified models, it may be possible to determine the existence (or non-existence) of Hopf bifurcations by examining whether $m(t)$ is periodic, as this seems like it would be a requirement for such behavior in reality. Additional avenues for investigation include the direct extension to periodic compartmental models (Wang & Zhao, 2008) through the use of some form of double truncated mean residual waiting-time (Khorashadizadeh et al., 2012), and a thorough comparison to renewal equations (Champredon et al., 2018), as such work would likely prove valuable for the study infectious disease dynamics.

As GDECMs are based on the integral equation versions of the Kermack and McKendrick model, they share this model's limitations. Namely, the assumptions of a sufficiently large and well-mixed population, the compartmentalization of diseases into distinct stages, and the transmission assumption of the law of mass-action. While these limitations may seem numerous, they do not impede research on traditional models, and thereby should not inhibit the theoretical extension and application of the broader class of models proposed here.

The traditional assumptions that reduce the integral equation version of the Kermack and McKendrick model to a system of differential equations provide disease modelers with a rich source for mathematical and scientific discovery. Here, we proposed a generalization of these traditional assumptions to a biologically more accurate description of the total quantity of a disease in a population. By imposing these new assumptions, we provide a more descriptive picture of how disease propagates throughout a population while retaining the convenience and simplicity of differential equation compartmental models.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.idm.2021.08.007.

Author's contributions

Scott Greenhalgh: Conceptualization, Formal analysis, Writing − Original draft preparation, Reviewing and Editing. Carly Rozins: Formal analysis, Writing − Original draft preparation, Reviewing and Editing

References

Alexander, M. E., & Moghadas, S. M. (2005). Bifurcation analysis of an SIRS epidemic model with generalized incidence. *SIAM Journal on Applied Mathematics, 65*, 1794–1816. https://doi.org/10.1137/040604947

Bailey, N. T. J. (1956). On estimating the latent and infectious periods of measles: I. Families with two susceptibles only. *Biometrika, 43*, 15. https://doi.org/10.2307/2333574

Bailey, N. T. J., & Others. (1975). *The mathematical theory of infectious diseases and its applications*. Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE.

Blackwood, J. C., & Childs, L. M. (2018). An introduction to compartmental modeling for the budding infectious disease modeler. *Lett Biomath, 5*, 195–221. https://doi.org/10.1080/23737867.2018.1509026

Brauer, F. (2008). Age-of-infection and the final size relation. *Mathematical Biosciences and Engineering, 5*, 681–690. https://doi.org/10.3934/mbe.2008.5.681

Brauer, F. (2010). General compartmental epidemic models. *Chinese Ann Math Ser B, 31*, 289–304. https://doi.org/10.1007/s11401-009-0454-1

Casagrandi, R., Bolzoni, L., Levin, S. A., & Andreasen, V. (2006). The SIRC model and influenza A. *Mathematical Biosciences, 200*, 152–169. https://doi.org/10.1016/j.mbs.2005.12.029

Champredon, D., Dushoff, J., & Earn, D. J. D. (2018). Equivalence of the erlang-distributed SEIR epidemic model and the renewal equation. *SIAM Journal on Applied Mathematics, 78*, 3258–3278. https://doi.org/10.1137/18M1186411

Clark, T. G., Bradburn, M. J., Love, S. B., & Altman, D. G. (2003a). Survival Analysis Part I: Basic concepts and first analyses. *British Journal of Cancer, 89*, 232–238. https://doi.org/10.1038/sj.bjc.6601118

Clark, T. G., Bradburn, M. J., Love, S. B., & Altman, D. G. (2003b). Survival Analysis Part IV: Further concepts and methods in survival analysis. *British Journal of Cancer, 89*, 781–786. https://doi.org/10.1038/sj.bjc.6601117

Cliff, A. D., Haggett, P., Ord, J. K., & Versey, G. R. (1981). *Spatial diffusion: An historical geography of epidemics in an island community*. Cambridge University Press.

Day, T., & Proulx, S. R. (2004). A general theory for the evolutionary dynamics of virulence. *The American Naturalist, 163*, E40–E63. https://doi.org/10.1086/382548

Diekmann, O., Metz, H., & Heesterbeek, H. (1995). The legacy of Kermack and McKendrick. In D. Mollison (Ed.), *Epidemic models: Their structure and relation to data* (pp. 95–115). Cambridge, UK: Cambridge University Press.

van den Driessche, P. (2017). Reproduction numbers of infectious disease models. *Infect Dis Model, 2*, 288–303. https://doi.org/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*, 29–48. https://doi.org/10.1016/S0025-5564(02)00108-6

Earn, D. J. (2000). A simple model for complex dynamical transitions in epidemics. *Science, 287*, 667–670. https://doi.org/10.1126/science.287.5453.667, 80-.

Feng, Z., Xu, D., & Zhao, H. (2007). Epidemiological models with non-exponentially distributed disease stages and applications to disease control. *Bulletin of Mathematical Biology, 69*, 1511–1536. https://doi.org/10.1007/s11538-006-9174-9

Finkelstein, M. S. (2002a). On the reversed hazard rate. *Reliability Engineering & System Safety, 78*, 71–75. https://doi.org/10.1016/S0951-8320(02)00113-8

Finkelstein, M. S. (2002b). On the shape of the mean residual lifetime function. *Applied Stochastic Models in Business and Industry, 18*, 135–146. https://doi.org/10.1002/asmb.461

Gough, K. J. (1977). The estimation of latent and infectious periods. *Biometrika. Biometrika Trust, 64*, 559–565.

Greenhalgh, S., & Day, T. (2017). Time-varying and state-dependent recovery rates in epidemiological models. *Infect Dis Model*. https://doi.org/10.1016/j.idm.2017.09.002

Greenhalgh, S., Galvani, A. P., & Medlock, J. (2015a). Disease elimination and re-emergence in differential-equation models. *Journal of Theoretical Biology, 387*, 174–180. https://doi.org/10.1016/j.jtbi.2015.09.037

Greenhalgh, S., Hobbs, C. V., & Parikh, S. (2018). Brief report: Antimalarial benefit of HIV antiretroviral therapy in areas of low to moderate malaria transmission intensity. *Journal of Acquired Immune Deficiency Syndromes, 79*, 249–254. https://doi.org/10.1097/QAI.0000000000001783

Greenhalgh, S., Ndeffo, M., Galvani, A. P., & Parikh, S. (2015b). The epidemiological impact of HIV antiretroviral therapy on malaria in children. *AIDS, 29*, 473–482. https://doi.org/10.1097/QAD.0000000000000550

Gupta, R. C., & Bradley, D. M. (2003). Representing the mean residual life in terms of the failure rate. *Mathematical and Computer Modelling, 37*, 1271–1280. https://doi.org/10.1016/S0895-7177(03)90038-0

Heffernan, J. M., Smith, R. J., & Wahl, L. M. (2005). Perspectives on the basic reproductive ratio. *Journal of The Royal Society Interface, 2*, 281–293. https://doi.org/10.1098/rsif.2005.0042

Hethcote, H. W., & Tudor, D. W. (1980). Integral equation models for endemic infectious diseases. *Journal of Mathematical Biology, 9*, 37–47. https://doi.org/10.1007/BF00276034

Hethcote, H. W., & van den Driessche, P. (1991). Some epidemiological models with nonlinear incidence. *Journal of Mathematical Biology, 29*, 271–287. Available: http://www.ncbi.nlm.nih.gov/pubmed/2061695.

Hurtado, P. J., & Kirosingh, A. S. (2019). Generalizations of the 'linear chain trick'': Incorporating more flexible dwell time distributions into mean field ODE models. *Journal of Mathematical Biology*. https://doi.org/10.1007/s00285-019-01412-w

Kelly-Hope, L. A., & McKenzie, F. E. (2009). The multiplicity of malaria transmission: A review of entomological inoculation rate measurements and methods across sub-Saharan Africa. *Malaria Journal, 8*, 19. https://doi.org/10.1186/1475-2875-8-19

Kermack, W. O., & McKendrick, A. G. (1932). Contributions to the mathematical theory of epidemics–II. The problem of endemicity, 1991 *Bulletin of Mathematical Biology, 53*, 57–87. Available: http://www.ncbi.nlm.nih.gov/pubmed/2059742.

Kermack, W. O., & McKendrick, A. G. (1933). Contributions to the mathematical theory of epidemics–III. Further studies of the problem of endemicity, 1991 *Bulletin of Mathematical Biology, 53*, 89–118. Available: http://www.ncbi.nlm.nih.gov/pubmed/2059743.

Kermack, W. O., & McKendrick, A. G. (1991). Contributions to the mathematical theory of epidemics–I. 1927. *Bulletin of Mathematical Biology, 53*, 33–55. Available: http://www.ncbi.nlm.nih.gov/pubmed/2059741.

Khorashadizadeh, M., Rezaei Roknabadi, A. H., & Mohtashami Borzadaran, G. R. (2012). Characterizations of lifetime distributions based on doubly truncated mean residual life and mean past to failure. *Communications in Statistics - Theory and Methods, 41*, 1105–1115. https://doi.org/10.1080/03610926.2010.535626

Krylova, O., & Earn, D. J. D. (2013). Effects of the infectious period distribution on predicted transitions in childhood disease dynamics. *Journal of The Royal Society Interface, 10*, 20130098. https://doi.org/10.1098/rsif.2013.0098

Liu, W. M., Hethcote, H. W., & Levin, S. A. (1987). Dynamical behavior of epidemiological models with nonlinear incidence rates. *Journal of Mathematical Biology, 25*, 359—380. Available: http://www.ncbi.nlm.nih.gov/pubmed/3668394.

Liu, W. M., Levin, S. A., & Iwasa, Y. (1986). Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models. *Journal of Mathematical Biology, 23*, 187—204. Available: http://www.ncbi.nlm.nih.gov/pubmed/3958634.

Lloyd, A. L. (2001). Realistic distributions of infectious periods in epidemic models: Changing patterns of persistence and dynamics. *Theoretical Population Biology, 60*, 59—71. https://doi.org/10.1006/tpbi.2001.1525

Papathanasiou, V. (1995). A characterization of the Pearson system of distributions and the associated orthogonal polynomials. *Annals of the Institute of Statistical Mathematics, 47*, 171—176. https://doi.org/10.1007/BF00773421

Pearson, K. (1893). Contributions to the mathematical theory of evolution. *Proceedings of the Royal Society of London, 54*, 329—333. https://doi.org/10.1098/rspl.1893.0079

Pearson, K. (1895). Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. *Philos Trans R Soc A Math Phys Eng Sci, 186*, 343—414. https://doi.org/10.1098/rsta.1895.0010

Ponciano, J. M., & Capistrán, M. A. (2011). First principles modeling of nonlinear incidence rates in Seasonal epidemics. Pascual M, editor. *PLoS Computational Biology, 7*, Article e1001079. https://doi.org/10.1371/journal.pcbi.1001079

Roberts, M., Andreasen, V., Lloyd, A., & Pellis, L. (2015). Nine challenges for deterministic epidemic models. *Epidemics, 10*, 49—53. https://doi.org/10.1016/j.epidem.2014.09.006

Ruan, S., & Wang, W. (2003). Dynamical behavior of an epidemic model with a nonlinear incidence rate. *Journal of Differential Equations, 188*, 135—163. https://doi.org/10.1016/S0022-0396(02)00089-X

Simpson, R. E. H. (1952). Infectiousness of communicable diseases in the household (measles, chickenpox, and mumps). *Lancet, 2*, 549—554. Available: http://www.ncbi.nlm.nih.gov/pubmed/12981903.

Tchuenche, J., Garira, W., Gumel, A., & Mukandavire, Z. (2009). Mathematical analysis of a model for HIV-malaria co-infection. *Mathematical Biosciences and Engineering, 6*, 333—362. https://doi.org/10.3934/mbe.2009.6.333

Wang, W., & Zhao, X.-Q. (2008). Threshold dynamics for compartmental epidemic models in periodic environments. *Journal of Dyn Differ Equations, 20*, 699—717. https://doi.org/10.1007/s10884-008-9111-8