Article

Analytical Estimation of Data-Motivated Time-Dependent Disease Transmission Rate: An Application to Ebola and Selected Public Health Problems

Anuj Mubayi [1,2,3,4] , **Abhishek Pandey** [5] , **Christine Brasic** [6] , **Anamika Mubayi** [7] , **Parijat Ghosh** [8] and **Aditi Ghosh** [6,9]*

1   PRECISIONheor, Los Angeles, CA 90025, USA; amubayi@ilstu.edu
2   Department of Mathematics, Illinois State University, Normal, IL 61790, USA
3   Center for Collaborative Studies in Mathematical Biology, Illinois State University, Normal, IL 61790, USA
4   College of Health Solutions, Arizona State University, Tempe, AZ 28608, USA
5   Yale School of Public Health, Yale University, New Haven, CT 06520, USA; abhishek.pandey@yale.edu
6   Department of Mathematics and Statistics, University of Wisconsin-Whitewater, Whitewater, WI 53190, USA; brasiccs23@uww.edu
7   Department of Chemistry, University of Allahabad, Allahabad 211002, India; anamika.mubayi@gmail.com
8   School of Medicine-Family and Community Medicine, University of Missouri, Columbia, MO 65201, USA; phool.ghosh@gmail.com
9   Department of Mathematics, Texas A & M-Commerce, Commerce, TX 75429, USA
*   Correspondence: ghosha@uww.edu

Abstract: Obtaining reasonable estimates for transmission rates from observed data is a challenge when using mathematical models to study the dynamics of ?infectious? diseases, like Ebola. Most models assume the transmission rate of a contagion either does not vary over time or change in a fixed pre-determined adhoc ways. However, these rates do vary during an outbreak due to multitude of factors such as environmental conditions, social behaviors, and public-health interventions deployed to control the disease, which are in-part guided by changing size of an outbreak. We derive analytical estimates of time-dependent transmission rate for an epidemic in terms of either incidence or prevalence using a standard mathematical SIR-type epidemic model. We illustrate applicability of our method by applying data on various public health problems, including infectious diseases (Ebola, SARS, and Leishmaniasis) and social issues (obesity and alcohol drinking) to compute transmission rates over time. We show that time-dependent transmission rate estimates can have a large variation, depending on the type of available data and other epidemiological parameters. Time-dependent estimation of transmission rates captures the dynamics of the problem better and can be utilized to understand disease progression more accurately.

Keywords: transmission coefficient; infectious disease dynamics; compartmental model; parameter estimation; epidemic modeling

1. Introduction

An epidemic is a function of environmental factors and a contact structure that varies over time, which in turn leads to varying transmission potential of an "infection". We also refer the word "infection" to describe social influences exerted by a typical influential individual with a particular social problem that results in a naive (to the social problem) individual becoming involved in the problem. For example, an alcohol drinker might influence an abstainer into imitating drinking behavior and initiating alcohol drinking. Many authors have studied outbreaks of social problems and infectious diseases using compartmental transmission/influence model. Qualitative aspects of homogeneous mixing models with constant transmission potential of an infection are well understood for various applications. These models are relatively easy to analyze and can answer questions, at the

population level, with good precision. Homogeneous mixing compartmental models have a long history; however, quantification of temporal transmission potential of an infectious agent, an input variable for this type of model, has been a challenge.

William Hamer first published a paper in 1906 containing an epidemic model for the transmission of measles where his observation included the incidence of new cases, in a time interval, is proportional to the product, SI, of the density of susceptibles (S) and the density I of infectives (I) in the population. The formulation of incidence can be explained by considering some epidemiological quantity. Consider a single susceptible individual in a homogeneously mixing population of size N. This individual contacts other members of the population at the rate c, per unit time, and a proportion I/N of these contacts are with individuals who are "infectious". If the probability of transmission of infection given contact is ρ, then the rate at which the infection is transmitted to a susceptible is $\rho c I/N$, per unit time, and the rate at which the susceptible population becomes infected is $\rho c S I/N$.

The contact rate is often a function of population density, reflecting the fact that contacts take time and saturation occurs. If c is assumed approximately proportional to N or equal to constant, incidence can be represented by terms such as βSI (referred as *mass action incidence*) or $\beta SI/N$ (referred as *standard incidence*), respectively. The parameter β, which includes the contact rate c, is known as a "transmission coefficient" (or "effective contact rate" or "transmission potential") with units as time^{-1}. At low population densities mass action is a reasonable approximation of a much more complex contact structure; however, in general, standard incidence is more appropriate for modeling transmission for human diseases or influences for social problems. The term $\beta I/N$ is sometimes referred as the force of infection, i.e., per-capita rate at which susceptible members of the host population are becoming infected. On the other hand, the transmission rate, represents the number of new infections per unit of time generated by an infected individual. The transmission rate is calculated by dividing incidence for a given time period by a disease prevalence for the same time interval.

Most infectious disease data are collected in form of incidence and/or prevalence. *Prevalence* of a "disease" in a population is defined as the total number of cases of the disease in the population at a given time, whereas *prevalence proportion* is computed by dividing the total number of cases in the population by the number of individuals in the population. It is used as an estimate of how common a condition is within a population over a certain period of time. *Incidence* is a measure of the risk of developing some new condition within a specified period of time. Incidence proportion (also known as cumulative incidence) is the number of new cases within a specified time period divided by the size of the population initially at risk. When the denominator is the sum of the person-time of the at-risk population, it is also known as the incidence density rate or person-time incidence rate. Using person-time rather than just time handles situations where the amount of observation time differs between people, or when the population at risk varies with time. Prevalence is a measurement of all individuals affected by the disease within a particular period of time, whereas incidence is a measurement of the number of new individuals who contract a disease during a particular period of time. So, prevalence and incidence proportion at the time t is given by $I(t)/N(t)$ and $\beta \times (S(t)/N(t)) \times (I(t)/N(t))$, respectively.

In compartmental mathematical models, varied assumptions are made based on characteristics of a modeling disease which lead modelers to focus on more important aspects of the epidemic. For example, an epidemic that occurs on a timescale that is much shorter than that of the population replenishment (that is, epidemic occurs at a much faster rate than births and deaths in the population), constant population size can be assumed. Additional common features of these models might include temporary or permanent recovery of infected individuals and a birth rate into infective class. Whether establishment or a major outbreak of an infectious disease or a social problem will occur in a population, requires extensive experience or a mathematical model of disease dynamics and estimates of the parameters of the disease model. Here, we provide a method for estimating the transmission coefficient (β), which is a key parameter in shaping the epidemic dynamics

generated from the model because of nonlinearity associated with the term containing it. A suitable set of data for estimation of β includes prevalence and incidence of the outbreak in question. There are many different methods for estimating β but most of them results in an aggregate value over time. The methods in the literature include estimation using regression of prevalence and time since start of an epidemic [1], estimating from equation for basic reproductive rate when threshold density is known [2], estimating from equilibrium prevalence [3,4], using age prevalence curves [5], inferring from behavior or contact data [6], and iterative comparison of field prevalence data with model predictions [7].

Some researchers have modeled time-varying transmission coefficients for diseases that follow seasonal patterns but using a predefined functional form [8]. On the other hand, a study by Finkenstadt and Grenfell [9] uses a discrete time model that allows for a temporally varying transmission parameter with a period of one year with no assumption on functional form. However, their estimation is computationally intensive and assumes that reporting interval of the available data must be an integer fraction of the serial interval of the disease. Another study by Pollicott et al. (2012), suggested first to fit the data with a pre-defined continuous function and then provide an analytical estimate of the transmission rate. However, their method was only applicable to prevalence data, with some restrictive assumptions on the initial number of susceptible or vital rates [10]. In the current study, we provide an analytical estimation of transmission coefficient using distinct and novel mathematical approach that is not only applicable to both prevalence and incidence data but also has its applicability to wide public-health problems including social issues. Table 1 provides a brief comparison of the estimation procedures in the Pollicott et al and the present study.

Table 1. Some common differences and similarity in estimation methodology between the current study and that in Pollicott et al. (2012).

	Pollicott et al. (2012) [10]	Mubayi et al. (Current Study)
Data Modified (fitted to a functional form)	✓	✗
Applicable to Prevalence	✓	✓
Applicable to Incidence	✗	✓
Requires $\beta(0)$	✓	✗
Requires S(0)	✗	✓
Multiple Applications	✗	✓

Examples of social problems such as alcohol drinking and obesity and infectious diseases such as Ebola, Visceral Leishmaniasis (or Kala-azar), and SARS are used to show relevance of the analytical work. The available data of US college alcohol drinking and obesity outbreak in US include prevalence trends, whereas incidence data of Ebola outbreak in West Africa (Guinea, Sierra Leone, and Liberia), Kala-azar outbreak in Bihar, and SARS epidemic in Hong Kong are used for the estimation.

In this paper, we compute time-dependent and -independent transmission coefficient of Ebola virus disease along with other health care problems such as college alcohol drinking, the obesity epidemic in United States, the spread of Visceral Leshmaniasis, and the spread of the 2003 SARS Outbreak in Hong Kong. The remaining paper is stratified as follows: Section 2 provides a compartmental SIR model and two analytical expressions of transmission coefficients based on prevalence and incidence data; examples for computing coefficient over time using each of the two expressions and field data are shown in Section 3; and finally, the results are discussed in Section 4. Figure 1 represents the overview of this paper.
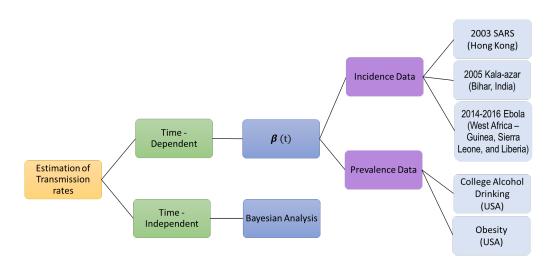
Figure 1. Overview of the paper.

2. Materials and Methods

2.1. Formulation for Time-Dependent Estimation

Epidemics in a population are typically captured via an SIR-type (Susceptible-Infectious-Recovered) epidemic models. It is assumed that in a well-mixed population, individuals interact with each other at random. The model considers a population of size N, where $S(t)$, $I(t)$ and $R(t)$ represents number of susceptible, infectious, and removed individuals at time t. Individuals are recruited in the population at the rate $b(t)$, die at the μ constant per-capita mortality rate and recover from infection at a α constant per-capita rate. The model assumes that the recruitment rate is governed by immigration, emigration and natural births, and the recovered individuals are immune to the infection but after temporary immunity period a recovered individual may lose immunity and move to S class. Hence, a "disease" outbreak in population can be captured by the following system of differential equations:

$$\frac{dS}{dt} = pb(t) - \beta(t)SI + \gamma(t)R - \mu(t)S \tag{1}$$

$$\frac{dI}{dt} = \bar{p}b(t) + \beta(t)SI - \alpha(t)I - \mu(t)I \tag{2}$$

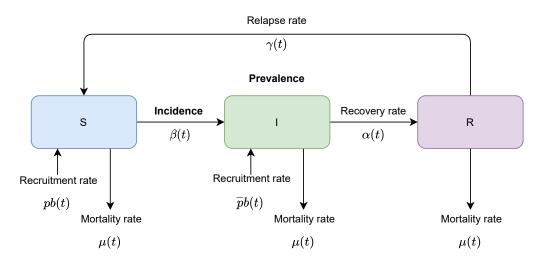where $R(t) = 1 - S(t) - I(t)$ and parameters are defined in Tables 2 and 3, Figure 2.

Figure 2. Schematic diagram for the SIR model.

Following steps carried out in Hadeler [11] and using Equations (1) and (2), we derive two explicit expressions for $\beta(t)$: one based on prevalence data and other on the incidence of the disease. The main derivation steps for are mentioned below.

Table 2. Definition of variables and parameters in the model given by Equations (1) and (2).

Variables	Definitions
S	Proportion of susceptibles
I	Proportion of "infected and infectious"
R	Proportion of recovered individuals
Parameters	**Definitions**
b(t)	Rate of recruitment in the population
p	Proportion of new recruits that are susceptibles
$\bar{p} = 1 - p$	Proportion of new recruits that are infectious
$\beta(t)$	Transmission or influence coefficient
$\alpha(t)$	Per-capita recovery rate (its reciprocal is infectious period)
$\gamma(t)$	Per-capita rate of losing immunity or relapse rate
$\mu(t)$	Per-capita mortality or departure rate

Table 3. Definition of variables and parameters in the model given by Equations (1) and (2).

Parameters	── Estimates ──				
	Ebola	**Alcohol**	**Obesity**	**Kala-Azar**	**SARS**
p	1.0	0.35 [12,13]	0.94 [14]	1.0	1.0
	(per month)	(per year)	(per year)	(per month)	(per day)
b	0.0	0.29 [15]	0.01	0.003 [16]	0.0
α	0.003 [17]	0.17 [4]	0.22 [18]	0.211 [7]	0.04 [19]
γ	0.008 [20]	0.0	0.14 [18]	0.0	0.0
μ	0.0	0.27 [4]	0.013	0.001 [7]	0.0

2.1.1. Derivation of $\beta(t)$ Expression in Terms of Prevalence

Suppose prevalence data are available. Derivation of $\beta(t)$ as a function of prevalence is carried out as follows. Adding Equations (1) and (2) we obtain

$$(S + I)' = (b(t) + \gamma(t)) - (\gamma(t) + \mu(t))(S + I) - \alpha(t)I \tag{3}$$

Setting $c(t) = b(t) + \gamma(t)$ and $d(t) = \gamma(t) + \mu(t)$ in Equation (3) and solving it we obtain

$$S(t) = (S(0) + I(0))Z(0,t) + \int_0^t Z(s,t)[c(s) - \alpha(s)I(s)]ds - I(t) \tag{4}$$

where $Z(a,b) = exp\left(-\int_a^b d(s)ds\right)$.

Isolating $\beta(t)$ from Equation (2) we obtain $\beta(t)$ as function of prevalence (*I*)

$$\beta(t) = \frac{I' + (\alpha(t) + \mu(t))I - \bar{p}b(t)}{SI} \tag{5}$$

where $S(t)$ is given by Equation (4).

Note, beside prevalence (*I*), we also need I' to compute $\beta(t)$ using Formula (5). However, I' can be approximated using prevalence data.

2.1.2. Derivation of $\beta(t)$ Expression in Terms of Incidence

On the other hand, suppose incidence data are available. To calculate expression of $\beta(t)$ as a function of incidence ($w(t) = \beta(t)SI$) we first solve Equation (2) for *I* with initial

condition $I(T)$ (where $T \in [0, L]$ is a time at which the prevalence proportion, $I(T)$, is available) and obtain

$$I(t) = I(T)H(T,t) - \int_t^T H(s,t)[w(s) + \bar{p}b(s)]ds \qquad (6)$$

where $H(a,b) = exp\left(-\int_a^b (\alpha(s) + \mu(s))ds\right)$.

Using this expression of $I(t)$ in Equation (1) and solving the resultant equation for S with initial condition $S(0)$ we obtain

$$\begin{aligned} S(t) = &Z(0,t)S(0) + \int_0^t Z(u,t)[pb(u) - w(u)]du \\ &+ \int_0^t Z(u,t)\gamma(u)\left[1 - I(T)H(T,u) + \int_u^T H(s,u)[w(s) + \bar{p}b(s)]ds\right]du \end{aligned} \qquad (7)$$

Thus,

$$\beta(t) = \frac{w(t)}{SI} \qquad (8)$$

where $S(t)$ and $I(t)$ are given by Equations (6) and (7), respectively.

Note, we need prevalence at time point T, $I(T)$, to compute $\beta(t)$ using Formula (7). The time point T can be appropriately chosen, close to maximum of prevalence and not towards starting or end of epidemic.

2.2. Time-Independent Estimation: Bayesian Analysis

The Bayesian Monte Carlo Markov Chains (MCMC) approach can be used to quantify uncertainty around the transmission rates and compare our analytical estimates with it.

Let θ represents vector of our transmission parameters and $\boldsymbol{y} = (y_1, y_2, \ldots, y_T)^T$ is the available data set. We can take likelihood function in our Bayesian approach as

$$L(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{i=1}^T \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) \times exp\left(-\frac{[\,logit(y_i) - logit(f_i(\boldsymbol{\theta}))\,]^2}{2\sigma^2}\right) \qquad (9)$$

where T is the total number of data points in the data set, σ is the appropriately chosen variance and $f(\theta)$ is the model output function for which data are used. If more than one data sets are used then the likelihood can be modified as follows:

$$L(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_k L(\boldsymbol{y_k}|\boldsymbol{\theta})$$

Although a Bayesian approach can provide uncertainty around time-independent average transmission rate, it does not inform how the transmission rate varied over time and uncertainty itself is constant over time. Therefore, this approach, while assists in understanding uncertainty in disease progression, it does not address the challenge of capturing changing transmission rates over the progression of an epidemic with respect to time.

3. Results

We use four examples to show how to estimate β over time from the available epidemiological data. The examples provide a method to study social and public-health issues. To compute estimates of $\beta(t)$, we use first order discretization for derivatives and composite trapezoidal rule for integration as given below

$$f'(t) \approx \frac{f(t+h) - f(t)}{h}$$

$$\int_0^t f(x)dx \approx h\left(\frac{f(0) + f(t)}{2} + \sum_{k=1}^{n-1} f(kh)\right).$$

These discretizations are used in the formulas given in Equations (5) and (8).

We can avoid this discretization by choosing a function, for example, a polynomial that can be fitted to the prevalence and incidence temporal data. This fitted function can then be used directly in Equations (5) and (8). Additional demographical and epidemiological data that we require in the $\beta(t)$ estimation for both incidence and prevalence case are duration of infectious period, recruitment rate, natural mortality rate, and relapse rate.

3.1. Using Incidence Data

In this section, we apply available incidence data to three past epidemics: the 2014–2016 Ebola outbreak in West Africa, the 2005 outbreak Visceral Leishmaniasis in the Indian state of Bihar, and the 2003 SARS outbreak in Hong Kong.

3.1.1. 2014–2016 Ebola Outbreak in West Africa

In this section, we estimate the transmission coefficient, $\beta(t)$ for the 2014–2016 Ebola epidemic in West Africa using available incidence data. The number of reported cases per month were retrieved from the Center for Disease Control and Prevention (CDC) and are shown totaled as West Africa (Figure 3a), and individually for Guinea (Figure 3c), Sierra Leone (Figure 3e), and Liberia (Figure 3g) [21]. For these estimates, prevalence is taken as 31 May 2015, as this point is close to the maximum prevalence and not towards the start of the epidemic (see Section 2.1.2). Incidence is calculated by dividing these case counts by the 2016 population for each country, as reported by the United Nations (UN) [22]. We assume a constant recovery rate of 10 days ($\alpha(t) = \alpha$), a constant relapse rate of 10 years ($\gamma(t) = \gamma$), no vertical transmission ($p = 1$), and a constant population ($b(t) = \mu(t) = u = 0$); since the CDC data provides monthly case counts, these parameters are adjusted to per month rates. We estimate $\beta(t)$ by simplifying Equation (6) as follows:

$$\beta(t) = \frac{w(t)}{\left(S(0) - \int_0^t w(u)du\right)\left(I(T)e^{\alpha(T-t)} - \int_t^T e^{\alpha(s-t)}w(s)ds\right)} \tag{10}$$

On discretizing Equation (10) we obtain the following expressions. If $t \leq T$,

$$\beta(t) \approx \frac{w(t)}{a_1 b_1} \tag{11}$$

where

$$a_1 = S(0) - h\left[\frac{w(0) + w(t)}{2} + \sum_{k=1}^{n-1} w(kh)\right]$$

and

$$b_1 = I(T)e^{\alpha(T-t)} - he^{-t\alpha}\left[\frac{g_1(t) + g_1(T)}{2} + \sum_{k=1}^{n-1} g_1(kh)\right]$$

If $t > T$,

$$\beta(t) \approx \frac{w(t)}{a_1 b_2} \tag{12}$$

where

$$b_2 = I(T)e^{\alpha(T-t)} + he^{-t\alpha}\left[\frac{g_1(t) + g_1(T)}{2} + \sum_{k=1}^{n-1} g_1(kh)\right]$$
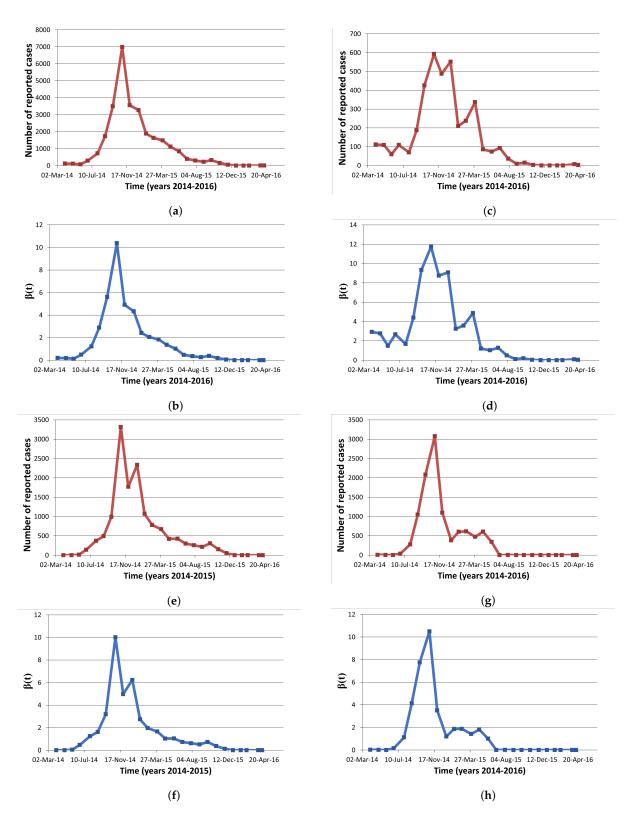
where $g_1(x) = e^{m_1 x}w(x)$ and $m_1 = \alpha$.

Figure 3. $\beta(t)$ estimates for the 2014–2016 Ebola outbreak: (**a**) Cases per month for West Africa; (**b**) $\beta(t)$ estimates for West Africa; (**c**) Cases per month for Guinea; (**d**) $\beta(t)$ estimates for Guinea; (**e**) Cases per month for Sierra Leone; (**f**) $\beta(t)$ estimates for Sierra Leone; (**g**) Cases per month for Liberia; (**h**) $\beta(t)$ estimates for Liberia.

For the estimation of $\beta(t)$ with regards to available incidence data, the estimates are found in Table A1 (see Appendix A) and are shown for West Africa (Figure 3b), Guinea (Figure 3d), Sierra Leone (Figure 3f), and Liberia (Figure 3h). Comparing the results for each region, we find the largest temporal estimate for both the mean and median $\beta(t)$ to be

that of Guinea (see Table 4 and Figure 4). Analyzing the estimates for transmission rate temporally, we observe that transmission rate follows the incidence pattern reflecting the effects of exponential incline in the beginning of epidemic as well as impacts of disease-acquired immunity as well as non-pharmaceutical interventions implemented over the course of epidemic (Figure 3).

Table 4. Summary statistics of Ebola results for $\beta(t)$.

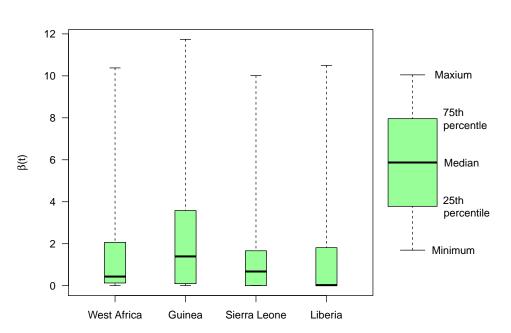|	**Minimum**	*Mean* $\pm$ *SD*	**25th Percentile**	**Median**	**75th Percentile**	**Maximum**
West Africa	0.00	1.57 ± 2.36	0.14	0.43	2.01	10.38
Guinea	0.00	2.73 ± 3.34	0.10	1.36	3.50	11.74
Sierra Leone	0.00	1.51 ± 2.30	0.02	0.68	1.65	6.23
Liberia	0.00	1.40 ± 2.52	0.00	0.03	1.71	10.50

Figure 4. Box-and-Whiskers plot for estimates of $\beta(t)$ using available incidence data for the 2014–2015 Ebola epidemic of West Africa.

3.1.2. 2005 Occurrence of Visceral Leishmaniasis in Bihar, India

Visceral Leishmaniasis (VL) is a vector borne infectious disease that is spread from person to person by a bite of the tiny insect, sandfly. Large population suffers from VL in some tropical and subtropical countries of the world. The highest burden of the VL is found in Indian state of Bihar. We obtained underreporting adjusted 2005 incidence data of Bihar from [7]. The data contain number of new cases during past month adjusted for underreporting. The Expression (13) is used to estimate $\beta(t)$ via two different models. The first model was for a single outbreak and hence demography was not considered whereas the second model assumed birth and death though with a same per-capita rate.

If $t \leq T$ then

$$\beta(t) \approx \frac{w(t)}{a_2 b_3} \tag{13}$$

where

$$a_2 = 1 - e^{-\mu t}(1 - S(0)) - he^{-tm_2}\left[\frac{g_2(0) + g_2(t)}{2} + \sum_{k=1}^{n-1} g_2(kh)\right]$$

and

$$b_3 = I(T)e^{(\alpha+\mu)(T-t)} - he^{-tm_3}\left[\frac{g_3(t) + g_3(T)}{2} + \sum_{k=1}^{n-1} g_3(kh)\right]$$

If $t > T$

$$\beta(t) \approx \frac{w(t)}{a_2 b_4} \tag{14}$$

where

$$a_3 = 1 - e^{-\mu t}(1 - S(0)) - he^{-tm_2}\left[\frac{g_2(0) + g_2(t)}{2} + \sum_{k=1}^{n-1} g_2(kh)\right]$$
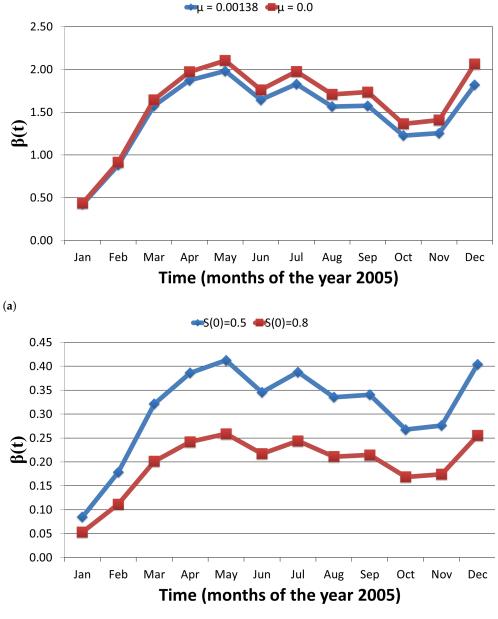
and

$$b_4 = I(T)e^{(\alpha+\mu)(T-t)} + he^{-tm_3}\left[\frac{g_3(t) + g_3(T)}{2} + \sum_{k=1}^{n-1} g_3(kh)\right]$$

where $g_i(x) = e^{m_i x} w(x)$ (for $i = 2, 3$), $m_2 = \mu$ and $m_3 = \alpha + \mu$.
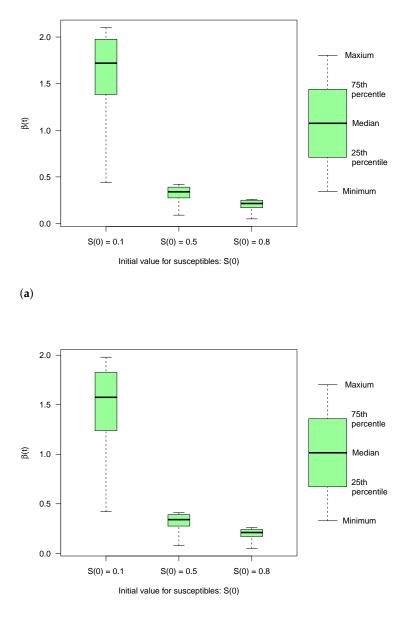
Since annual epidemic during 2005 started showing clear trend of decaying in the month of October, we took this time to compute the prevalence of VL in Bihar. Prevalence during October 2005 was computed under assumption that 25% of worldwide leishmaniasis prevalence is from VL cases whereas remaining is from other forms of Leishmaniasis. It also assumed 20% of global burden is in Bihar. Since some proportion of a population are naturally immune to the disease, we carried out estimation for three different values of initial proportion of susceptibles, namely 0.1, 0.5 and 0.8. Recovery rate of 0.21 per month and influx/outflux rate of the population of 0.00138 was computed using data from Mubayi et al. (2010) [7]. The other assumptions of the model include constant recovery (i.e., $\alpha(t) = \alpha$), no vertical transmission (i.e., $p = 1$), permanent recovery (i.e., $\gamma(t) = 0$) and same constant per-capita incoming and outgoing rates (i.e., $b(t) = \mu(t) = \mu$). We only model human population and do not take into account vector population explicitly. Thus, $\beta(t)$ could be interpreted as vectorial capacity of sandfly population transmitting infection between humans.

The obtained estimates of $\beta(t)$ are given in Table A2 (see Appendix A) and Figures 5a,b and 6a,b. The β estimates that we have computed here are comparable to corresponding estimates in [7] (in this reference the mean estimates are $\beta_h = 0.13$ (with median = 0.11, SD = 0.08, Q1 (25th percentile) = 0.07, Q3 (75th percentile) = 0.17) and $\beta_v = 0.12$ (with Median = 0.11, Std = 0.08, Q1 (25th percentile) = 0.07, Q3 (75th percentile) = 0.16) where around 75% of the population was susceptible).

(**a**)

(**b**)

Figure 5. Estimates of $\beta(t)$ for the 2005 outbreak of Visceral Leishmaniasis in the Indian state of Bihar, using available incidence data: (**a**) Estimates of $\beta(t)$ related to an outbreak of Visceral Leishmaniasis, when the initial value of susceptibles, $S(0) = 0.1$; (**b**) Estimates of $\beta(t)$ for two initial proportion of susceptibles in a population affected with Visceral Leishmaniasis. Estimates of $\beta(t)$ obtained for two different values of the mortality rate, μ, (0.0 and 0.00138) are almost same.

(**a**)

(**b**)

Figure 6. Box-and-Whiskers plot for estimates of $\beta(t)$ for two different values of μ in a Visceral Leishmaniasis outbreak: (**a**) Estimates for $\mu = 0.0$; (**b**) Estimates for $\mu = 0.00138$.

3.1.3. 2003 SARS Outbreak in Hong Kong

Severe acute respiratory syndrome (SARS) is a viral respiratory illness caused by a coronavirus. SARS epidemic in Hong Kong is shown in Figure 7a. We estimated transmission coefficient using a single outbreak model with parameters values given in Table A3. The formula used for estimating $\beta(t)$ is

$$\beta(t) = \frac{w(t)}{\left(S(0) - \int_0^t w(u)du \right) \left(I(T)e^{\alpha(T-t)} - \int_t^T e^{\alpha(s-t)}w(s)ds \right)} \tag{15}$$

(**a**)

(**b**)

Figure 7. The 2002–2003 SARS outbreak in Hong Kong: (**a**) Daily reported cases; (**b**) Estimation of $\beta(t)$ using available incidence data. Prevalence for 16 April 2003 was taken in the calculation when number of symptomatic cases started declining [23].

On discretizing Equation (15) we obtain following expressions. If $t \le T$,

$$\beta(t) \approx \frac{w(t)}{a_4 b_5} \tag{16}$$

where

$$a_4 = S(0) - h\left[\frac{w(0) + w(t)}{2} + \sum_{k=1}^{n-1} w(kh)\right]$$

and

$$b_5 = I(T)e^{\alpha(T-t)} - he^{-t\alpha}\left[\frac{g_4(t) + g_4(T)}{2} + \sum_{k=1}^{n-1} g_4(kh)\right] \tag{17}$$

If $t > T$,

$$\beta(t) \approx \frac{w(t)}{a_4 b_6} \tag{18}$$

where

$$b_6 = I(T)e^{\alpha(T-t)} + he^{-t\alpha}\left[\frac{g_4(t) + g_4(T)}{2} + \sum_{k=1}^{n-1} g_4(kh)\right] \tag{19}$$

where $g_4(x) = e^{m_4 x}w(x)$ and $m_4 = \alpha$.

The temporal estimates of $\beta(t)$ are shown in Table A3 (see Appendix A) and Figure 7b.

3.2. Using Prevalence Data

We use US national college alcohol drinking and obesity data as examples in this section. In Appendix A, We also present a hypothetical example with synthetic prevalence data and known time-varying transmission rate to illustrate the ability of our analytical expression to accurately capture the time-dependent transmission rate

3.2.1. College Alcohol Drinking

The available alcohol drinking data represent prevalence (proportion of cases at a certain time) and not incidence (new cases over time period). This is because the data are based on the survey where the drinking pattern estimates are obtained by asking individuals their drinking behavior during past one year. Hence, data can be interpreted as the number of individuals in certain drinking category at a particular time. Therefore, we use formula given in Equation (5) to estimate $\beta(t)$. We assume that drinking is a result of social influences exerted by drinkers (I) on susceptibles (S) or social drinkers. Individuals recovered from drinking at a constant rate α (i.e., $\alpha(t) = \alpha$). The recovery is assumed to be permanent (i.e., $\gamma(t) = 0$). The incoming and departure rates are same (i.e., $\mu(t) = b(t) = \mu$) and $p = 1$. These assumptions are reasonable in context of the type of data (college population) used here.

Alcohol drinking data, obtained from Engs et al., 1997 and 1999, is given in the Table A4 [12,13] that represent the trend observed in national college drinking surveys. The recovery rate, α is taken to be 0.17 [4]. We estimate $\beta(t)$ using simplified Equation (5) and above assumptions as follows

$$\beta(t) \approx \frac{\frac{I(t) - I(t-h)}{hI(t)} + \alpha + \mu - \frac{\bar{p}\mu}{I(t)}}{(S(0) + I(0))e^{-\mu t} - I(t) + e^{-\mu t}hc_1} \tag{20}$$

where

$$c_1 = \frac{f_2(0) + f_2(t)}{2} + \sum_{k=1}^{n-1} f_1(kh),$$

and

$$f_2(x) = e^{\mu x}[\mu - \alpha I(x)].$$

If $\mu = 0$, this equation can be reduced, where f_2 is $-\alpha I(x)$.

We found that mean estimate of β is 1.04 (std = 0.3; Table A4 see Appendix A and Figure 8) during 1982–1994 for the national college drinkers. The estimates of β are comparable to the estimates obtained in the [4]. These estimates of $\beta(t)$ are all contained in 95% CI of the estimates in the [4], which are $\beta_0 = 1.69$ (95% CI [0.63, 2.75]) and $\beta_2 = 0.75$ (95% CI [0.29, 1.21]).

Engs et al., 1994 and 1997 suggest that 65% of freshmen are drinkers during the start of Fall semester. Hence, we assumed that 0.65 proportion of incoming students are drinkers, i.e., $p = 0.35$. We assumed negligible change in size of a college population and consider rate of enrollments equal to combined rate of graduation and dropout rates (i.e., $b(t) = \mu(t) = \mu$).
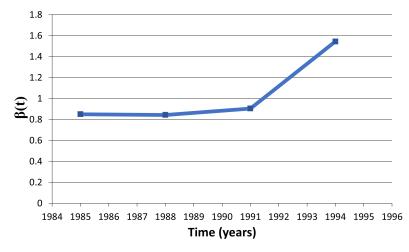
Figure 8. Estimates of $\beta(t)$ related to alcohol drinking college population, when $\mu = 0$.

3.2.2. Obesity Epidemic in US

We use model to see whether weight gain in one person is associated with weight gain in his or her family members and friends. Obese persons are an individual whose body-mass index (the weight in kilograms divided by the square of the height in meters) is greater than or equal to 30. It is found that there has been increasing number of obese persons in a community and a person's chances of becoming obese increases dramatically if he or she had a parent, sibling, friend or spouse who became obese in a given interval [24]. The most reasonable explanations for the obesity epidemic, include changes in which luxuries and food consumption are being promoted in the society and has not spared any socioeconomic class. An obesity is a result of individual's choice and behavior which is influenced by appearance and behavior of others in the community. Hence, it suggests that just as with the spread of drug-use or infectious diseases, weight gain in one person might influence weight gain in other person, i.e., it is not that obese or non-obese people simply find other similar people to hang out with. This influence could be direct or indirect, which can vary continuously over time and may depend on demographic and social factors of the community as well.

We used annual CDC data from references [14,18] to estimate parameters for our obese epidemic model. The data obtained from [25] include an age-adjusted prevalence of obesity in US using the projected 2000 U.S. population. The model assumes constant population and hence $b(t) = \mu(t) = \mu$. It is assumed that 6% of children are born obese [14]. The value of recovery rate is assumed to be equal to an average of rate at which an overweight individuals move on diet (4.068×10^{-3} per week [18]) and rate at which an obese individual stops or reduces bakery, fried meals and soft drinks consumption (4.4379×10^{-3} per week [18]). We assume obesity reduces life span by 6 to 7 years. Hence if average life span in US is 78.4 years than average life span of at-risk population for obese is $(78.4 - 6.5)$ years. The estimated β from [18] ranges from 0.02 to 0.04. These estimates are much lower than our estimated values in Table A5 (see Appendix A) with range of (0.36, 3.02) (Figure 9). This is because the region of our study differs from the region modeled by [18]. Our results suggest that estimates of transmission coefficient increase with increase in μ and decrease in initial size of susceptible population, $S(0)$.

$$\beta(t) \approx \frac{\frac{I(t) - I(t-h)}{hI(t)} + \alpha + \mu - \frac{\bar{p}\mu}{I(t)}}{(S(0) + I(0))e^{-(\gamma+\mu)t} - I(t) + e^{-(\gamma+\mu)t}hc_2} \tag{21}$$

where

$$c_2 = \frac{f_3(0) + f_3(t)}{2} + \sum_{k=1}^{n-1} f_3(kh)$$

and

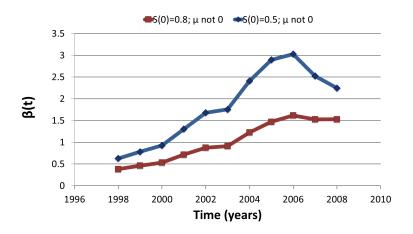$$f_3(x) = e^{(\gamma+\mu)x}[\gamma + \mu - \alpha I(x)].$$

Figure 9. $\beta(t)$ estimates of Obesity with initial (1997) prevalence of 19.5%.

3.3. Estimation of Time-Dependent Transmission Coefficient Using Synthetic Prevalence and Incidence Data

We demonstrate our method of using prevalence and incidence data to estimate time-dependent transmission coefficient using synthetic prevalence and incidence data generated with two choices of transmission coefficients and the model ((1) and (2)) with rest of parameters given by (Table 5).

Table 5. Parameters for generating synthetic prevalence data.

Parameters	p	$b(t)$	$\gamma(t)$	$\mu(t)$	$\alpha(t)$
Values	1	33/1000	0	33/1000	8

In the first case, we assumed transmission coefficient to be constant over time and in the second case, we consider a transmission coefficient that is seasonal. They are given by

1. $\beta(t) = 20$,
2. $\beta(t) = 20(1 - \epsilon \cos 2\pi t)$, with $\epsilon = 0.1$.

We generate daily prevalence and incidence data for two years and estimated monthly transmission coefficient using Equations (5) and (8) respectively. We used MATLAB's 'pchip' function to interpolate the synthetic prevalence and incidence data in the formulation and integrated using 'integral' function. The monthly estimates for time-dependent transmission coefficient were reasonably accurate and close to the true values of the transmission coefficients used to generate prevalence and incidence data in both the cases when transmission coefficient was constant and when it was periodic (Figure 10). As birth rates, mortality rates and recovery rates are often considered constant in models, we used constant terms for these variables to limit simulation time. However, if time-dependent information on these variables is available they can easily be incorporated and simulated.
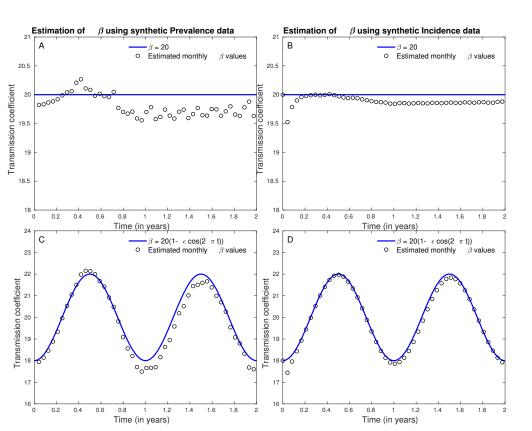
Figure 10. Monthly estimates of transmission coefficients and their true values used to generate synthetic prevalence data (**A**,**C**) and synthetic incidence data (**B**,**D**) for constant transmission coefficient (**A**,**B**) and periodic transmission coefficient (**C**,**D**).

4. Discussion

Compartmental models have provided valuable insights into the epidemiology of many infectious diseases. Transmission coefficient, a product of contact rate and probability of transmission given a contact, is a parameter in the compartmental model which naturally varies over time. This coefficient had the greatest effect on predictions of dynamics of disease or social problem and difficult to estimate. However, due to lack of detailed data as well as complexities involved in numerical estimating this parameter, most studies estimate it as a time-independent parameter averaging it over the course of epidemic. In this study, we present a method to estimate time-dependent transmission rate using two types of data commonly reported during infectious disease outbreaks: the time series of the number of infectives (or prevalence) and the number of new cases generated during a period of time (or incidence). By deriving an analytical method that uses a standard deterministic model and these data sets to directly estimate $\beta(t)$, this new approach resolves the computational challenges often involved with more complex model. By applying our approach to several infectious diseases, we illustrate applicability of our methods in various contexts. Moreover, similar approaches can be applied with any appropriate mathematical model to derive time-dependent transmission rate for diseases whose dynamics may need to incorporate other factors such as environment (for. e.g., role of waterbodies in cholera spread) or vector dynamics (for. e.g., impact of mosquito in dengue transmission).

Utility of approach presented in this manuscript is demonstrated using several public-health problems including Ebola, Visceral Leishmaniasis, US college alcohol drinking and obesity outbreak in the US. In particular, we estimated the temporal estimates of transmission rate for Ebola during 2014–2016 outbreak in West Africa (aggregated) as well as for individual countries of Liberia, Guinea and Sierra Leone. Our results though limited by the accuracy of data, demonstrated the wide-variability in transmission risks across the three countries. Moreover, we found that our temporal estimates of transmission risk followed

the pattern of incidence closely, but slightly delayed, reflecting the substantial contribution of transmission risk towards the nature of disease progression. During the times of public-health emergencies due to an infectious disease outbreaks such as Ebola outbreak in West Africa or ongoing COVID-19 pandemic, effective reproductive numbers are often estimated using incidence data to understand the progression of disease and inform strategies to curb the transmission. Although estimates of effective reproductive numbers are useful, combining it with estimation of time-varying transmission risk through our approach can be more informative to inform public-health decision-making. Transmission risk at a particular time is a product of contacts and probability of transmission. Thus, it can be used to make short term predictions about new infections as well as it can inform how much reduction in contact patterns or risk of transmission (through mask/vaccination/hygiene) can reduce the transmission parameter sufficiently to reverse the trend of an epidemic.

In the current study, we used simple deterministic model along with simple integration numerical techniques to show how commonly reported data (incidence and prevalence) can be used in informing temporal transmission risk, and thus manage public-health challenges more effectively. Practical application of our approach would improve with use of more complex models (appropriate) as well more sophisticated integration techniques. Moreover, analytical derivation can be used to understand the impact of changes in any other input parameter (such as smaller/longer quarantine periods) on transmission risk in a straight-forward way. Similarly, an area of future research can expand presented framework to understand how incomplete data may alter the quality of parameter estimation. Therefore, value of analysis reported here is as a beginning point for future research that will build on current approach to develop computational models that can inform policies in swift manner during public-health emergencies. We believe using our methods can provide good approximation of time-dependent transmission coefficients and goodness of approximation should increase with use of more sophisticated modeling techniques.

Abbreviations

The following abbreviations are used in this manuscript:

CDC	Center for Disease Control
MCMC	Monte Carlo Markov Chains
SD	Standard deviation
SIR	Susceptible–Infectious–Recovered
SARS	Severe Acute Respiratory Syndrome
USA	United States of America
US	United States
VL	Visceral Leishmaniasis

Appendix A. Tables

Table A1. 2014–2016 Ebola outbreak in West Africa. Estimates of $\beta(t)$ for incidence data.

	Date	West Africa Incidence: $\omega(t)$	$\beta(t)$	Guinea Incidence: $\omega(t)$	$\beta(t)$	Sierra Leone Incidence: $\omega(t)$	$\beta(t)$	Liberia Incidence: $\omega(t)$	$\beta(t)$
2014	31-Mar	5.1×10^{-6}	0.20	9.5×10^{-6}	2.93	0	0	1.7×10^{-6}	0.03
	30-Apr	4.8×10^{-6}	0.19	9.3×10^{-6}	2.78	0	0	1.1×10^{-6}	0.02
	28-May	3.2×10^{-6}	0.13	5.1×10^{-6}	1.50	2.2×10^{-6}	0.05	0	0
	24-Jun	1.2×10^{-5}	0.49	9.3×10^{-6}	2.68	1.9×10^{-5}	0.47	8.5×10^{-6}	0.16
	31-Jul	3.1×10^{-5}	1.23	6.0×10^{-6}	1.69	5.1×10^{-5}	1.24	6.1×10^{-5}	1.12
	28-Aug	7.3×10^{-5}	2.89	1.6×10^{-5}	4.41	6.7×10^{-5}	1.62	2.3×10^{-4}	4.14
	26-Sep	1.5×10^{-4}	5.62	3.6×10^{-5}	9.35	1.4×10^{-4}	3.20	4.5×10^{-4}	7.75
	31-Oct	3.0×10^{-4}	10.38	5.1×10^{-5}	11.74	4.5×10^{-4}	10.01	6.7×10^{-4}	10.50
	28-Nov	1.5×10^{-4}	4.92	4.2×10^{-5}	8.75	2.4×10^{-4}	4.98	2.4×10^{-4}	3.51
	31-Dec	1.4×10^{-4}	4.33	4.7×10^{-5}	9.08	3.2×10^{-4}	6.23	8.4×10^{-5}	1.20
2015	28-Jan	8.0×10^{-5}	2.42	1.8×10^{-5}	3.26	1.5×10^{-4}	2.74	1.3×10^{-4}	1.87
	25-Feb	6.9×10^{-5}	2.06	2.0×10^{-5}	3.58	1.1×10^{-4}	1.96	1.3×10^{-4}	1.88
	31-Mar	6.3×10^{-5}	1.84	2.9×10^{-5}	4.88	9.2×10^{-5}	1.66	1.0×10^{-4}	1.43
	29-Apr	4.7×10^{-5}	1.37	7.3×10^{-6}	1.21	5.8×10^{-5}	1.04	1.3×10^{-4}	1.81
	31-May	3.6×10^{-5}	1.03	6.3×10^{-6}	1.03	5.9×10^{-5}	1.04	7.5×10^{-5}	1.01
	30-Jun	1.7×10^{-5}	0.48	7.9×10^{-6}	1.29	4.1×10^{-5}	0.73	0	0
	31-Jul	1.3×10^{-5}	0.36	3.1×10^{-6}	0.50	3.5×10^{-5}	0.62	1.3×10^{-6}	0.02
	31-Aug	9.5×10^{-6}	0.27	7.7×10^{-7}	0.12	2.9×10^{-5}	0.52	0	0
	29-Sep	1.4×10^{-5}	0.39	1.3×10^{-6}	0.21	4.2×10^{-5}	0.74	0	0
	30-Oct	6.7×10^{-6}	0.19	2.6×10^{-7}	0.04	2.1×10^{-5}	0.37	0	0
	30-Nov	2.3×10^{-6}	0.07	0	0	7.5×10^{-6}	0.13	6.6×10^{-7}	0.01
	30-Dec	0	0	0	0	0	0	0	0
2016	27-Jan	8.5×10^{-8}	2.44×10^{-3}	0	0	2.7×10^{-7}	4.81×10^{-3}	0	0
	17-Feb	0	0	0	0	0	0	0	0
	30-Mar	3.0×10^{-7}	0.01	6.0×10^{-7}	0.10	0	0	0	0
	13-Apr	2.5×10^{-7}	0.01	2.6×10^{-7}	0.04	0	0	6.6×10^{-7}	0.01

Table A2. 2005 Visceral Leishmaniasis incidence data from Bihar, India. Estimates of $\beta(t)$ of models with and without demography.

Month	Incidence $(\omega(t))$	S(0):	0.1	0.5	0.8	0.1	0.5	0.8
			$\beta(t) : \mu = 0$			$\beta(t) : \mu = 0.00138$		
Jan	3.4×10^{-5}		0.44	0.09	0.05	0.42	0.08	0.05
Feb	6.2×10^{-5}		0.92	0.18	0.11	0.88	0.18	0.11
Mar	1.0×10^{-4}		1.65	0.33	0.21	1.57	0.32	0.20
Apr	1.2×10^{-4}		1.97	0.39	0.25	1.87	0.39	0.24
May	1.3×10^{-4}		2.10	0.42	0.26	1.98	0.41	0.26
June	1.0×10^{-4}		1.76	0.35	0.22	1.64	0.35	0.22
July	1.1×10^{-4}		1.98	0.39	0.25	1.83	0.39	0.24
Aug	9.5×10^{-5}		1.71	0.34	0.21	1.57	0.34	0.21
Sept	9.2×10^{-5}		1.73	0.34	0.22	1.58	0.34	0.17
Oct	6.8×10^{-5}		1.36	0.27	0.17	1.23	0.27	0.17
Nov	6.5×10^{-5}		1.41	0.28	0.17	1.25	0.28	0.17
Dec	9.2×10^{-5}		2.06	0.41	0.26	1.82	0.40	0.26

Table A3. Estimates of $\beta(t)$ using incidence data from the 2003 Severe Acute Respiratory Syndrome (SARS) epidemic in Hong Kong.

Date	Incidence $\omega(t)$	$\beta(t)$	Date	Incidence $\omega(t)$	$\beta(t)$	Date	Incidence $\omega(t)$	$\beta(t)$
14-Feb	1.5×10^{-7}	8.11×10^{-4}	23-Mar	1.6×10^{-5}	1.49×10^{-1}	29-Apr	1.5×10^{-6}	1.33×10^{-2}
15-Feb	0	0	24-Mar	1.6×10^{-5}	1.31×10^{-1}	30-Apr	2.9×10^{-7}	2.73×10^{-3}
16-Feb	0	0	25-Mar	1.2×10^{-5}	8.84×10^{-2}	1-May	7.3×10^{-7}	7.06×10^{-3}
17-Feb	1.5×10^{-7}	9.07×10^{-4}	26-Mar	9.1×10^{-6}	6.34×10^{-2}	2-May	1.0×10^{-6}	1.02×10^{-2}
18-Feb	0	0	27-Mar	7.8×10^{-6}	5.25×10^{-2}	3-May	7.3×10^{-7}	7.45×10^{-3}
19-Feb	1.5×10^{-7}	9.78×10^{-4}	28-Mar	5.9×10^{-6}	3.90×10^{-2}	4-May	4.4×10^{-7}	4.61×10^{-3}
20-Feb	1.5×10^{-7}	1.01×10^{-3}	29-Mar	4.7×10^{-6}	3.10×10^{-2}	5-May	7.3×10^{-7}	7.92×10^{-3}
21-Feb	1.5×10^{-7}	1.05×10^{-3}	30-Mar	7.8×10^{-6}	5.08×10^{-2}	6-May	7.3×10^{-7}	8.15×10^{-3}
22-Feb	1.5×10^{-7}	1.09×10^{-3}	31-Mar	5.9×10^{-6}	3.78×10^{-2}	7-May	1.3×10^{-6}	1.50×10^{-2}
23-Feb	1.5×10^{-7}	1.13×10^{-3}	1-Apr	3.8×10^{-6}	2.46×10^{-2}	8-May	4.4×10^{-7}	5.14×10^{-3}
24-Feb	4.4×10^{-7}	3.51×10^{-3}	2-Apr	5.1×10^{-6}	3.32×10^{-2}	9-May	5.9×10^{-7}	7.06×10^{-3}
25-Feb	1.5×10^{-7}	1.21×10^{-3}	3-Apr	5.0×10^{-6}	3.22×10^{-2}	10-May	2.9×10^{-7}	3.64×10^{-3}
26-Feb	1.5×10^{-7}	1.26×10^{-3}	4-Apr	6.2×10^{-6}	3.95×10^{-2}	11-May	7.3×10^{-7}	9.38×10^{-3}
27-Feb	1.5×10^{-7}	1.30×10^{-3}	5-Apr	6.5×10^{-6}	4.10×10^{-2}	12-May	4.4×10^{-7}	5.79×10^{-3}
28-Feb	2.9×10^{-7}	2.70×10^{-3}	6-Apr	5.9×10^{-6}	3.69×10^{-2}	13-May	4.4×10^{-7}	5.97×10^{-3}
1-Mar	5.9×10^{-7}	5.58×10^{-3}	7-Apr	4.1×10^{-6}	2.58×10^{-2}	14-May	4.4×10^{-7}	6.16×10^{-3}
2-Mar	7.3×10^{-7}	7.19×10^{-3}	8-Apr	3.5×10^{-6}	2.24×10^{-2}	15-May	4.4×10^{-7}	6.35×10^{-3}
3-Mar	2.9×10^{-7}	2.97×10^{-3}	9-Apr	3.1×10^{-6}	1.98×10^{-2}	16-May	1.5×10^{-7}	2.19×10^{-3}
4-Mar	4.4×10^{-7}	4.60×10^{-3}	10-Apr	4.7×10^{-6}	3.04×10^{-2}	17-May	2.9×10^{-7}	4.52×10^{-3}
5-Mar	4.4×10^{-7}	4.76×10^{-3}	11-Apr	3.4×10^{-6}	2.20×10^{-2}	18-May	0	0
6-Mar	2.5×10^{-6}	2.74×10^{-2}	12-Apr	3.1×10^{-6}	2.03×10^{-2}	19-May	1.5×10^{-7}	2.43×10^{-3}
7-Mar	4.1×10^{-6}	4.49×10^{-2}	13-Apr	2.6×10^{-6}	1.77×10^{-2}	20-May	2.9×10^{-7}	5.02×10^{-3}
8-Mar	5.0×10^{-6}	5.34×10^{-2}	14-Apr	4.1×10^{-6}	2.78×10^{-2}	21-May	7.3×10^{-7}	1.29×10^{-2}
9-Mar	5.0×10^{-6}	5.21×10^{-2}	15-Apr	3.4×10^{-6}	2.30×10^{-2}	22-May	1.5×10^{-7}	2.65×10^{-3}
10-Mar	1.8×10^{-6}	1.83×10^{-2}	16-Apr	2.6×10^{-6}	1.82×10^{-2}	23-May	0	0
11-Mar	3.1×10^{-6}	3.23×10^{-2}	17-Apr	3.4×10^{-6}	2.36×10^{-2}	24-May	1.5×10^{-7}	2.85×10^{-3}
12-Mar	2.9×10^{-6}	3.07×10^{-2}	18-Apr	2.5×10^{-6}	1.77×10^{-2}	25-May	1.5×10^{-7}	2.95×10^{-3}

Table A3. *Cont.*

Date	Incidence $\omega(t)$	$\beta(t)$	Date	Incidence $\omega(t)$	$\beta(t)$	Date	Incidence $\omega(t)$	$\beta(t)$
13-Mar	2.4×10^{-6}	2.47×10^{-2}	19-Apr	1.9×10^{-6}	1.38×10^{-2}	26-May	1.5×10^{-7}	3.05×10^{-3}
14-Mar	3.8×10^{-6}	4.01×10^{-2}	20-Apr	1.0×10^{-6}	7.60×10^{-3}	27-May	1.5×10^{-7}	3.16×10^{-3}
15-Mar	3.7×10^{-6}	3.82×10^{-2}	21-Apr	1.8×10^{-6}	1.34×10^{-2}	28-May	1.5×10^{-7}	3.27×10^{-3}
16-Mar	5.0×10^{-6}	5.11×10^{-2}	22-Apr	2.2×10^{-6}	1.70×10^{-2}	29-May	0	0
17-Mar	4.0×10^{-6}	3.99×10^{-2}	23-Apr	1.6×10^{-6}	1.27×10^{-2}	30-May	2.9×10^{-7}	7.00×10^{-3}
18-Mar	1.9×10^{-6}	1.93×10^{-2}	24-Apr	1.9×10^{-6}	1.54×10^{-2}	31-May	0	0
19-Mar	3.7×10^{-6}	3.72×10^{-2}	25-Apr	1.9×10^{-6}	1.57×10^{-2}	1-Jun	0	0
20-Mar	2.8×10^{-6}	2.82×10^{-2}	26-Apr	1.2×10^{-6}	9.87×10^{-3}	2-Jun	0	0
21-Mar	3.8×10^{-6}	3.85×10^{-2}	27-Apr	1.9×10^{-6}	1.64×10^{-2}	3-Jun	0	0
22-Mar	4.6×10^{-6}	4.53×10^{-2}	28-Apr	4.4×10^{-7}	3.88×10^{-3}			

Table A4. Alcohol drinking data and estimates of $\beta(t)$.

Years	1982	1985	1988	1991	1994
Drinkers ($I(t)$)	52%	51%	49%	46%	47%
$\beta(t)$ when $\mu = 0$	−	0.85	0.84	0.90	1.54

Table A5. Age-adjusted Prevalence of Obesity in US using the projected 2000 U.S. population [25].

Year	1997	1998	1999	2000	2001	2002
Prev. (%)	19.5	20.6	21.5	21.8	22.9	23.8
Year	2003	2004	2005	2006	2007	2008
Prev (%)	23.5	24.3	25.3	26.2	26.6	27.5

References

1. Anderson, R.; May, R. Population biology of infectious diseases. Part I. *Nature* **1979**, *280*, 361–367. [CrossRef] [PubMed]
2. Coyne, M.; Smith, G.; McAllister, F. Mathematic model for the population biology of rabies in raccoons in the mid-Atlantic states. *Am. J. Vet. Res.* **1989**, *50*, 2148–2154. [PubMed]
3. Kribs-Zaleta, C. Estimating Contact Process Saturation in Sylvatic Transmission of Trypanosoma cruzi in the United States. *PLoS Negl. Trop. Dis.* **2010**, *4*. [CrossRef] [PubMed]
4. Mubayi, A.; Greenwood, P.; Wang, X.; Castillo-Chavez, C.; Gorman, D.; Gruenewald, P.; Saltz, R. Types of Drinkers and Drinking Settings: An Application of a Mathematical Model. *Addiction* **2021**, *106*, 749–758. [CrossRef] [PubMed]
5. Cohen, J. *Schistosomiasis: A Human Host-Parasite System*, 1st ed.; Oxford: Blackwell, UK, 1981; pp. 237–56.
6. Johnson, L. The Interaction between HIV and Other Sexually Transmitted Infections in South Africa: A Model-Based Evaluation. Ph.D. Thesis, Department of Actuarial Science, University of Cape Town, Cape Town, South Africa, 2008.
7. Mubayi, A.; Castillo-Chavez, C.; Chowell, G.; Kribs-Zaleta, C.; Kumar, N.; Siddiqui, N.; Das, P. Transmission dynamics and underreporting of Kala-azar in Indian state of Bihar. *J. Theor. Biol.* **2010**, *262*, 177–185. [CrossRef] [PubMed]
8. Fine, P.; Clarkson, J. Measles in England and Wales I: An Analysis of Factors Underlying Seasonal Patterns. *Int. J. Epidemiol.* **1982**, *11*, 5–14. [CrossRef] [PubMed]
9. Finkenstadt, B.; Grenfell, B. Time series modelling of childhood diseases: a dynamical systems approach. *J. R. Stat. Soc.* **2000**, *49*, 187–205. [CrossRef]
10. Pollicott, M.; Wang, H.; Weiss, H. Extracting the time-dependent transmission rate from infection data via solution of an inverse ODE problem. *J. Biol. Dyn.* **2012**, *6*, 509–523. [CrossRef]
11. Hadeler, K. Parameter Identification in Epidemic Models. *Math. Biosci.* **2011**, *229*, 185–189. [CrossRef] [PubMed]
12. Engs, R.; Hanson, D.; Diebold, B. The Drinking Patterns and Problems of a National Sample of College Students, 1994: Implications for Education. *J. Alcohol Drug Educ.* **1997**, *42*, 54–80.
13. Engs, R.; Hanson, D. Reduction of Consumption Theory: A Test Using the Drinking Patterns and Problems of Collegians in the United States, 1983–1994. *Coll. Stud. J.* **1999**, *33*, 333–345.
14. Kim, J.; Peterson, K.E.; Scanlon, K.S.; Fitzmaurice, G.M.; Must, A.; Oken, E.; Rifas-Shiman, S.L.; Rich-Edwards, J.W.; Gillman, M.W. Trends in overweight from 1980 through 2001 among preschool-aged children enrolled in a health maintenance organization. *Obesity* **2006**, *14*, 1107–1112. [CrossRef]

15. University of California. Office of the President. UC Information Digest. A Reference Guide on Student Access & Performance at the University of California. Available online: http://www.ucop.edu/sas/infodigest03/ (accessed on 4 September 2010).

16. Rahman, M.; Daniel, E. A Reproductive Health Communication Model That Helps Improve Young Women's Reproductive Life and Reduce Population Growth: The Case of PRACHAR from Bihar, India. *Pathfind. Int. Res. Eval. Work. Pap. Ser.* **2010**, *1*, 1–32.

17. Velásquez, G.E.; Aibana, O.; Ling, E.J.; Diakite, I.; Mooring, E.Q.; Murray, M.B. Time from Infection to Disease and Infectiousness for Ebola Virus Disease, a Systematic Review. *Clin. Infect. Dis.* **2015**, *61*, 1135–1140. [CrossRef] [PubMed]

18. Jodara, L.; Santonjaa, F.J.; Gonzalez-Parrab, G. Modeling dynamics of infant obesity in the region of Valencia, Spain. *Comput. Math. Appl.* **2008**, *56*, 679–689. [CrossRef]

19. Mubayi, A.; Kribs-Zaleta, C.; Martcheva, M.; Castillo-Chavez, C. A cost-based comparison of quarantine strategies for new emerging diseases. *Math. Biosci. Eng.* **2010**, *7*, 687. [PubMed]

20. CDC. Ebola Virus Disease-Survivors. 5 November 2019. Available online: https://www.cdc.gov/vhf/ebola/treatment/survivors.html (accessed on 1 January 2021).

21. CDC. 2014 Ebola Outbreak in West Africa Epidemic Curves. 3 April 2019. Available online: https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/cumulative-cases-graphs.html (accessed on 1 January 2021).

22. UN. Department of Economic and Social Affairs Population Dynamics. 2019 Revision of World Population Prospects. 28 August 2019. Available online: https://population.un.org/wpp/ (accessed on 1 January 2021).

23. Riley, S.; Fraser, C.; Donnelly, C.A.; Ghani, A.C.; Abu-Raddad, L.J.; Hedley, A.J.; Leung, G.M.; Ho, L.M.; Lam, T.H.; Thach, T.Q.; et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science* **2003**, *300*, 1961–1966. [CrossRef] [PubMed]

24. Christakis, N.; Fowler, J. The Spread of Obesity in a Large Social Network over 32 Years. *N. Engl. J. Med.* **2007**, *357*, 370–379. [CrossRef] [PubMed]

25. Centers for Disease Control and Prevention. *A Program for Early Release of Selected Estimates from the National Health Interview Survey*; (CDC) National Center for Health Statistics (NCHS): Hyattsville, MD, USA, 2008. Available online: https://www.cdc.gov/nchs/index.htm (accessed on 24 April 2021).