On epidemic growth rates and the estimation of the basic reproduction number

Fred Brauer[1], Gerardo Chowell[2,3] *;

[1] Department of Mathematics, The University of British Columbia

Vancouver, B.C., Canada V6T 1Z2

[2] Mathematical, Computational, Modeling Sciences Center,

School of Human Evolution and Social Change, Arizona State University, Tempe, AZ 85287, USA;

[3] Division of Epidemiology and Population Studies, Fogarty International Center,

National Institutes of Health, Bethesda, MD, USA;

*Corresponding author. Email: gchowell@asu.edu Fax: (505) 480-965-7671

Keywords: Influenza; pandemic; epidemiology; basic reproduction number; model.

1 Initial exponential growth rates in epidemic models

Early estimation of the basic reproduction number $\mathcal{R}_0$ as early as possible in the face of an epidemic is crucial for public health decision making [2, 4]. The reason for this is that knowledge of the reproduction number gives valuable information for the estimation of the transmission potential of the epidemic and for the comparison of possible epidemic management strategies. In an epidemic, there is an initial stochastic phase followed by exponential growth continuing until the number of members of the population who are no longer susceptible to infection becomes a significant fraction of the population. It may be possible to obtain real time data sufficient to estimate this initial exponential growth rate, and this leads to the task of using these data to estimate $\mathcal{R}_0$.

However, the estimate of $\mathcal{R}_0$ obtained depends on the specific model being used for the epidemic. Such an estimate has been suggested in [3] and [8]. Here, we give a full mathematical derivation of this estimation rule, with some examples.

2 The age of infection epidemic model

The general Kermack-McKendrick age of infection model [6] for a population of constant total population size N is

$$\begin{aligned} S' &= -a\frac{S}{N}\varphi \qquad (1)\\ \varphi(t) &= \varphi_0(t) + \int_0^t a\frac{S(t-\tau)}{N}\varphi(t-\tau)A(\tau)d\tau \\ &= \varphi_0(t) + \int_0^t [-S'(t-\tau)]A(\tau)d\tau. \end{aligned}$$

Here $\varphi(t)$ represents the total infectivity of all infectious individuals at time $t, \varphi_0(t)$ represents the total infectivity at time t of all individuals who were already infected at time $t = 0$, and $A(\tau)$ represents the mean infectivity of all individuals who had been infected τ time units previously, including those who are no longer infectious. The function $A(\tau)$ is the product of the function representing the fraction of infected members still infected at infection age τ and the relative infectivity $\pi(\tau)$ at infection age τ. For this model, it is known [1] that

$$\mathcal{R}_0 = a\int_0^\infty A(\tau)d\tau, \qquad (2)$$

and the final size relation between the reproduction number and the number S_∞ of individuals who go through the epidemic without being infected is

$$ln\frac{S(0)}{S_\infty} = \mathcal{R}_0\left[1 - \frac{S_\infty}{N}\right]. \qquad (3)$$

In (3) it is assumed that all individuals infected initially have infection age zero. If there are individuals with positive infection age, there is a positive correction term that must be subtracted from the right side of (3).

In the simplest SIR model in a population of constant total size N [6]

$$\begin{aligned} S' &= -a\frac{S}{N}I \\ I' &= a\frac{S}{N}I - \alpha I \end{aligned} \tag{4}$$

it is clear that one should define an epidemic as a case in which $I'(0) > 0$. For the general age of infection model (1), this is not possible. In fact, for a disease with a very long exposed period in a population into which a small number of infectives is introduced the number of infectives could decrease initially before starting to grow as exposed individuals become infective. Thus we must give a different definition of an epidemic for more general models such as (1). The definition that we propose is the following:

In a disease transmission model with no demographic effects, there is no epidemic if the equilibrium with all members of the population susceptible is (locally) asymptotically stable, and there is an epidemic if this equilibrium is unstable, in each case considering only perturbations of the equilibrium with positive infected initial states.

In order to validate this definition it is necessary to develop the analysis of equilibria of (1). The first step is to replace (1) by the *limit equation*

$$\begin{aligned} S' &= -a\frac{S}{N}\varphi \\ \varphi(t) &= \int_0^\infty a\frac{S(t-\tau)}{N}\varphi(t-\tau)A(\tau)d\tau. \end{aligned} \tag{5}$$

This limit equation is just the model (1) with a particular choice of initial function φ_0. According to the asymptotic theory of integral equations [7], the asymptotic behavior of (1) is the same as that of the limit equation (5) for every initial function with

$\lim_{t\to\infty}\varphi_0(t)=0$. To analyze (5), we would ordinarily linearize about an equilibrium, but this approach is not applicable since there is a line of equilibria $\varphi=0$. In order to avoid this problem, we replace the model (5) by the model

$$\begin{aligned} S' &= \mu N - a\frac{S}{N}\varphi - \mu S \\ \varphi(t) &= \int_0^\infty a\frac{S(t-\tau)}{N}\varphi(t-\tau)e^{-\mu\tau}A(\tau)d\tau. \end{aligned} \tag{6}$$

The model (6) is obtained from the model (5) by including a birth rate μN of susceptibles and a proportional death rate μ in each class. In fact, the age of infection epidemic model neglects demographic processes, arguing that these operate on a much slower time scale than the epidemiological process. In the model (6) we are restoring the demographic process; we will assume that μ is small and will ultimately return to (5) and (1) by letting $\mu\to 0$.

The model (6) has a disease - free equilibrium $S=N,\varphi=0$, and the linearization at this equilibrium is

$$\begin{aligned} u'(t) &= -\mu u + av(t) \\ v(t) &= a\int_0^\infty v(t-\tau)e^{-\mu\tau}A(\tau d\tau. \end{aligned}$$

Note that if $v(0)=0$, the linearization has solution

$$u(t)=u(0)e^{-\mu t},\quad v(t)\equiv 0.$$

In our definition of an epidemic we have ruled out such initial states. The characteristic equation is the condition on λ that the linearization have a solution $u=u_0e^{\lambda t}, v=v_0e^{\lambda t}$,

and this is

$$det\begin{bmatrix} -(\mu+\lambda) & a \\ 0 & a\int_0^\infty e^{-(\lambda+\mu)\tau}A(\tau)d\tau - 1 \end{bmatrix} = 0.$$

There are two roots of the characteristic equation, namely $\lambda = -\mu$ and the solution of

$$a\int_0^\infty e^{-(\lambda+\mu)\tau}A(\tau)d\tau = 1. \tag{7}$$

Because this is true for every value of $\mu > 0$, we may take the limit as $\mu \to 0$, to see that the characteristic equation of the linearization of (5) at the disease - free equilibrium $S = N, \varphi = 0$ is the solution of

$$a\int_0^\infty e^{-\lambda\tau}A(\tau)d\tau = 1. \tag{8}$$

The integral in (8) is a decreasing function $F(\lambda)$ of λ and

$$F(0) = \mathcal{R}_0.$$

Thus the solution λ of (8) is positive if $\mathcal{R}_0 > 1$ and negative if $\mathcal{R}_0 < 1$. We have thus established the following result.

There is an epidemic for the model (1) if and only if $\mathcal{R}_0 > 1$. If there is an epidemic, there are solutions with exponential growth rate given by the solution of (8).

According to [8], the initial exponential growth rate in an epidemic is the solution λ of (8). The supporting argument given in [8] is not mathematically complete as it assumes an exponentially growing population and a stable age of infection distribution. These restrictive assumptions are not necessary, and the analysis given here justifies this equation for the initial exponential growth rate.

3 Example: The general $SEIR$ model

We consider an $SEIR$ model with general distributions of stay in both the exposed and infectious period. Suppose the fraction of exposed individuals who are still in the exposed class s time units after being exposed is $Q(s)$ and the fraction of individuals who are still in the infectious class s time units after entering the infectious class is $P(s)$, with $Q(s), P(s)$ non-negative, non-increasing functions such that

$$Q(0) = P(0) = 1, \quad \int_0^\infty Q(s)ds < \infty, \quad \int_0^\infty Q(s)ds < \infty.$$

Then Q and P represent survival probabilities in the classes E and I respectively. The probability density function for E, which will appear in the model, is

$$q(\tau) = -Q'(\tau).$$

We assume that E_0 newly exposed members enter the exposed class at time $t = 0$. Then

$$\begin{aligned} S' &= -a\frac{S}{N}I \\ E(t) &= E_0Q(t) + \int_0^t [-S'(\tau)]Q(t-\tau)d\tau. \end{aligned}$$

Differentiation of the equation for $E(t)$ gives

$$E'(t) = E_0Q'(t) - S'(t) + \int_0^t [-S'(\tau)]Q'(t-\tau)d\tau,$$

and this shows that the input to the infectious stage at time t is

$$-E_0Q'(t) - \int_0^t [-S'(\tau)]Q'(t-\tau)d\tau.$$

and

$$I(t) = -E_0 \int_0^t Q'(u)P(t-u)du - \int_0^t [-S'(s)]Q'(u-s)dsP(t-u)du.$$

The first term in this expression may be written as $I_0(t)$, and the second term may be simplified, using interchange of the order of integration in the iterated integral, to yield

$$\int_0^t \int_0^u [-S'(s)]q(u-s)dsP(t-u)du = \int_0^t \int_s^t q(u-s)duP(t-u)[-S'(s)]ds.$$

If we define

$$T(t-s) = \int_s^t q(u-s)P(t-u)du = \int_0^{t-s} q(t-s-v)P(v)dv,$$

or

$$T(\tau) = \int_0^\tau q(\tau - v)P(v)dv, \tag{9}$$

we obtain

$$I(t) = I_0(t) + \int_0^t [-S'(s)]T(t-s)ds.$$

Then the model is

$$\begin{aligned} S' &= -a\frac{S}{N}I \qquad (10)\\ E(t) &= E_0Q(t) + \int_0^t [-S'(s)]Q(t-s)ds \\ I(t) &= I_0(t) + \int_0^t [-S'(s)]T(t-s)ds, \end{aligned}$$

which is in age of infection form with $\varphi = I$ and $A(\tau) = T(\tau)$, and we have an explicit expression for $T(\tau)$. Then

$$\begin{aligned}
\mathcal{R}_0 &= a\int_0^\infty A(\tau)d\tau \\
&= a\int_0^\infty \int_0^\tau [-Q'(\tau-u)]P(u)dud\tau \\
&= a\int_0^\infty \int_u^\infty [-Q'(\tau-u)]d\tau P(u)du \\
&= a\int_0^\infty \int_0^\infty [-Q'(v)]dvP(u)du \\
&= a\int_0^\infty P(u)du.
\end{aligned} \tag{11}$$

The initial exponential growth rate satisfies

$$a\int_0^\infty e^{-\lambda\tau}\int_0^\tau [-Q'(\tau-u)]P(u)dud\tau = 1,$$

which reduces to

$$\begin{aligned}
1 &= a\int_0^\infty q(v)e^{-\lambda v}dv\int_0^\infty e^{-\lambda u}P(u)du \\
&= a\left[1-\lambda\int_0^\infty e^{-\lambda v}Q(v)dv\right]\int_0^\infty e^{-\lambda u}P(u)du,
\end{aligned} \tag{12}$$

with the aid of integration by parts.

Suppose we let

$$F(\lambda) = a\left[1-\lambda\int_0^\infty e^{-\lambda v}Q(v)dv\right]\int_0^\infty e^{-\lambda u}P(u)du,$$

and choose $\lambda = \lambda_0$ with $F(\lambda_0) = 1$.

Suppose we lengthen the exposed period by increasing the function Q, without altering any other aspects of the model. This decreases the value of $F(\lambda_0)$ and therefore would produce a decrease in the initial exponential growth rate. The limiting case as

$Q \to 0$ is the SIR model, whose exponential initial growth rate is the solution λ of

$$a \int_0^\infty e^{-\lambda u} P(u) du = 1.$$

4 Example: A general quarantine/isolation epidemic model

To cope with a disease outbreak for which there is no (as yet) known treatment, the only methods available are isolation of diagnosed infective and quarantine of suspected exposed members of the population. This approach was used during the $SARS$ epidemic of 2003, and modelled in [5]. An $SEIR$ model with general exposed and infective periods as well as quarantine and isolation has been anlyzed in [9]. While the model in [9] was deterministic, the analysis was from a probabilistic point of view. Here we add quarantine and isolation to the model (13) and derive some of the results of [9] from a compartmental approach.

We move members from the exposed class to a quarantine class Q at rate ψ and from the infective class to an isolated class H at rate φ. For the moment, we assume that both quarantine and isolation are perfectly effective, so that no infections are transmitted from either quarantined or isolated members. Then we need not include Q or H in the model unless we wish to track the number of individuals quarantined and isolated. We need only adjust the model (13) to include the removals. With quarantine but not yet isolation, the new equation for E is

$$E(t) = E_0 e^{-\psi t} Q(t) + \int_0^t [-S'(s)] e^{-\psi(t-s)} Q(t-s) ds.$$

The input to I at time u becomes

$$E_0 q(u) + \int_0^u [-S'(\tau)] e^{-\psi(u-\tau} q(u-\tau) d\tau.$$

Now, $I(t)$ is given by

$$I(t) = E_0 \int_0^t q(u) e^{-\psi(t-u)} P(t-u) du + E_0 \int_0^t [-S'(s)] q(u-s) E^{-\psi(u-s)} ds P(t-u) du.$$

The first term in this expression may be written as $I_0(t)$, and the second term may be simplified, using interchange of the order of integration in the iterated integral much as in the previous section, to yield

$$\int_0^t [-S'(s)] T(t-s) ds$$

with

$$T(v) = \int_0^v q(y) e^{-\psi y} P(v-y) dy.$$

Then the model is

$$\begin{aligned} S' &= -a \frac{S}{N} I \\ E(t) &= E_0 e^{-\psi t} Q(t) + \int_0^t [-S'(s)] e^{-\psi(t-s)} Q(t-s) ds. \\ I(t) &= I_0(t) + \int_0^t [-S'(s)] T(t-s) ds. \end{aligned} \tag{13}$$

If we add isolation at a rate φ of infectives, we obtain

$$I(t) = e^{-\varphi t} I_0(t) + \int_0^t [-S'(s)] e^{-\varphi(t-s)} T(t-s) ds,$$

and the quarantine/isolation model is

$$\begin{aligned} S' &= -a \frac{S}{N} I \\ E(t) &= E_0 e^{-\psi t} Q(t) + \int_0^t [-S'(s)] e^{-\psi(t-s)} Q(t-s) ds. \\ I(t) &= E_0 e^{-\varphi t} I_0(t) + \int_0^t [-S'(s)] e^{-\varphi(t-s)} T(t-s) ds. \end{aligned} \tag{14}$$

The reproduction number is now a control reproduction number $\mathcal{R}_c(\psi, \varphi)$, depending on the control parameters ψ and φ,

$$\begin{aligned}
\mathcal{R}_c(\psi,\varphi) &= a\int_0^\infty e^{-\varphi\tau}T(\tau)d\tau \\
&= a\int_0^\infty e^{-\varphi\tau}E_0\int_0^\tau q(y)e^{-\psi y}P(\tau-y)dyd\tau \\
&= a\int_0^\infty e^{-\psi y}q(y)\left[\int_y^\infty e^{-\varphi(\tau-y)}P(\tau-y)d\tau\right]dy \\
&= a\int_0^\infty e^{-\psi y}q(y)dy\int_0^\infty e^{-\varphi u}P(u)du.
\end{aligned}$$

If we know the functions P and q we may calculate the sensitivity of $\mathcal{R}_)$ to the quarantine and isolation rates and thus compare quarantine and isolation as management strategies. In [9], this expression is written in terms of the Laplace transforms of q and P,

$$\mathcal{R}_c(\psi,\varphi) = aE_0\mathcal{L}_q(\psi)\mathcal{L}_P(\varphi).$$

A naive view of epidemic models suggests that if we know the basic reproduction number, we can determine the size of an epidemic.There are at least two quite different problems with this view. Estimation of the basic reproduction number from early observed data depends on the structure of the model being assumed. For example, if there is an exposed period, the initial exponential growth rate is less than if infected individuals become infectious immediately but this does not affect the basic reproduction number. This implies that an estimate of the basic reproduction number from the initial exponential growth rate will be too small. A second problem is that in an epidemic early data may be incomplete and inaccurate, and its use may lead to a poor estimate of the reproduction number.

5 Estimating R_0 using a compartmental epidemic model

In practice, the reproduction number denoted simply by R and defined as the number of secondary cases generated by a primary infectious cases in a partially protected population might be useful [10, 11]. R can also be estimated from the initial growth phase of an epidemic in such a partially immunized population. In a randomly mixing population, the relationship between the basic reproduction number (R_0) and the reproduction number (R) is given by $R = (1 - p)R_0$ where p is the proportion of the population that is effectively protected against infection (in the beginning of an epidemic). Besides, for many recurrent infectious diseases including seasonal influenza, estimating the background immunity p in the population is extremely difficult due to cross-immunity of antigenically-related influenza strains and prior vaccination campaigns.

Statistical methods to quantitatively estimate R_0 have been reviewed by Klaus Dietz [12]. Depending on the characteristics of data and underlying assumptions of the models, R_0 can be estimated using various different approaches [13]. Here we focus on the estimation of R_0 from an inverse problem perspective using compartmental epidemic models based on systems of ordinary differential equations. A recent review on methods for the estimation of the basic reproduction number in the context of the 1918-19 influenza pandemic has been given by Chowell and Nishiura (2008) [14].

The simple SEIR model classifies individuals as susceptible (S), exposed (E), infec-

tious (I), recovered (R), and dead (D) [15]. Susceptible individuals in contact with the virus enter the exposed class at the rate $\beta I(t)/N$, where β is the transmission rate, $I(t)$ is the number of infectious individuals at time t and $N = S(t) + E(t) + I(t) + R(t)$ is the total population for any t. The entire population is assumed to be susceptible at the beginning of the epidemic. Individuals in latent period (E) progress to the infectious class at the rate k (where $1/k$ suggests the mean latent period). We assume homogeneous mixing (*i.e.* random mixing) between individuals and, therefore, the fraction $I(t)/N$ is the probability of a random contact with an infectious individual in a population of size N. Since we assume that the time-scale of the epidemic is much faster than characteristic times for demographic processes (natural birth and death), background demographic processes are not included in the model. Infectious individuals either recover or die from influenza at the mean rates γ and δ, respectively. Recovered individuals are assumed protected for the duration of the outbreak. The mortality rate is given by $\delta = \gamma$ [CFP/(1-CFP)], where CFP is the mean case fatality proportion. The transmission process can be modeled using the system of nonlinear differential equations:

$$\begin{cases} \frac{dS(t)}{dt} &= -\frac{\beta S(t) I(t)}{N} \\ \frac{dE(t)}{dt} &= \frac{\beta S(t) I(t)}{N(t)} - kE(t) \\ \frac{dI(t)}{dt} &= kE(t) - (\gamma + \delta) I(t) \\ \frac{dR(t)}{dt} &= \gamma I(t) \\ \frac{dD(t)}{dt} &= \delta I(t) \\ \frac{dC(t)}{dt} &= kE(t) \end{cases} \tag{15}$$

where $C(t)$ is the cumulative number of infectious individuals. The basic reproduction number of the above system (15) is given by the product of the mean transmission rate and the mean infectious period, $R_0 = \beta/(\gamma + \delta)$.

5.1 Parameter estimation

In the simplest manner, model parameters can be estimated via least-square fitting of the model solution to the observed data. That is, one looks for the set of parameters $\hat{\mathbf{\Theta}}$ whose model solution best fits the epidemic data by minimizing the sum of the squared differences between the observed data y_t and the model solution $C(t, \mathbf{\Theta})$. That is, we minimize:

$$X(\mathbf{\Theta}) = \sum_{\mathbf{t=1}}^{\mathbf{n}} (\mathbf{y_t} - \mathbf{C}(\mathbf{t}, \mathbf{\Theta}))^{\mathbf{2}} \tag{16}$$

The standard deviation of the parameters can be estimated by computing the asymptotic variance-covariance $AV(\hat{\Theta})$ matrix of the least-squares estimate by [16]:

$$\mathbf{AV}(\hat{\mathbf{\Theta}}) = \sigma^{\mathbf{2}} (\nabla_{\mathbf{\Theta}} \mathbf{C}(\mathbf{\Theta_0})\ \nabla_{\mathbf{\Theta}} \mathbf{C}(\mathbf{\Theta_0})^{\mathbf{T}})^{-\mathbf{1}} \tag{17}$$

which can be estimated by

$$\hat{\sigma}^2 (\hat{\nabla} \Theta \mathbf{C}(\hat{\mathbf{\Theta}}) \hat{\nabla} \mathbf{\Theta} \mathbf{C}(\hat{\mathbf{\Theta}})^{\mathbf{T}})^{-\mathbf{1}} \tag{18}$$

where n is the total number of observations, $\hat{\sigma}^2$ is the estimated variance, and $\hat{\nabla} C$ are numerical derivatives of C. Estimates of $\hat{R_0}$ can be obtained by substituting the corresponding individual parameter estimates into an analytical formula of R_0. Further, using the delta method [17], we can derive an expression for the variance of the

estimated basic reproduction number $\hat{R_0}$. An expression for the variance of R_0 for the simple SEIR model (Equations 15) is given by:

$$V(\hat{R_0}) \approx \hat{R_0}^2 \{\frac{V(\hat{\beta})}{\hat{\beta}^2} + \frac{V(\hat{\gamma})}{(\hat{\gamma}+\hat{\delta})^2} + \frac{V(\hat{\delta})}{(\hat{\gamma}+\hat{\delta})^2} - (\frac{2}{\hat{\beta}(\hat{\gamma}+\hat{\delta})})(Cov(\hat{\gamma},\hat{\beta}) - \frac{\hat{\beta}Cov(\hat{\delta},\hat{\gamma})}{\hat{\gamma}+\hat{\delta}} + Cov(\hat{\delta},\hat{\beta}))\}. \quad (19)$$

This expression depends on the variance (denoted by V) of the individual parameter estimates as well as their covariance (denoted by Cov).

5.2 Bootstrap confidence intervals

Another method to generate uncertainty bounds on the reproduction number is generating bootstrap confidence intervals by generating sets of realizations of the best-fit curve $C(t)$ [18]. Each realization of the cumulative number of case notifications $C_i(t)$ ($i = 1, 2, \ldots, m$) is generated as follows: for each observation $C(t)$ for $t = 2, 3, \ldots, n$ days generate a new observation $C_i'(t)$ for $t \geq 2$ ($C_i'(1) = C(1)$) that is sampled from a *Poisson* distribution with mean: $C(t) - C(t-1)$ (the daily increment in $C(t)$ from day $t-1$ to day t). The corresponding realization of the cumulative number of influenza notifications is given by $C_i(t) = \sum_{j=1}^{t} C_i'(t)$ where $t = 1, 2, 3, \ldots, n$. The reproduction number was then estimated from each of 1000 simulated epidemic curves to generate a distribution of R estimates from which simple statistics can be computed including 95% confidence intervals. These statistics need to be interpreted with caution. For example, 95% confidence intervals for R derived from our bootstrap sample of R should be interpreted as containing 95% of future estimates when the same assumptions are

made and the only noise source is observation error. It is tempting but incorrect to interpret these confidence intervals as containing the *true* parameters with probability 0.95.

5.3 Example: The transmissibility of the 1918 influenza pandemic in Winnipeg, Canada

The 1918-19 influenza pandemic known as the Spanish influenza has been the most devastating in recent history with estimated worldwide mortality ranging from 20 to 100 million deaths [19][36] with a case fatality of 2-6 percent [21][22]. The first pandemic wave arrived to Winnipeg at the end of September 1918 probably brought by returning soldiers at the end of war (Figure 1). The pandemic appears to have moved from the south of the city into the north (from the wealthy to the poor populations) [23]. The influenza mortality rate of influenza was 90 deaths per thousand in the north end, and 46 per thousand in the south.

Because influenza pandemics such as the Spanish flu from 1918-19 are associated to the emergence of novel influenza strains to which most of the population is susceptible, it might be reasonable to assume that the reproduction number $R \approx R_0$. Previous studies have estimated that R_0 of the 1918-19 influenza pandemic ranged between 1.5 and 5.4 [35][25][28][26][42][**?**][34][39][41][24][43] depending on the specific location and pandemic wave considered, type of data, estimation method, and level of spatial aggregation, which has ranged from small towns to entire nations with several million inhabitants. The variability of R_0 estimates suggests that local factors, including ge-

ographic and demographic conditions, could play an important role in disease spread [30][40].

We estimated the reproduction number of the 1918 influenza pandemic in Winnipeg, Canada by fitting the simple SEIR model (15) to the initial phase of the cumulative number of reported cases. Figures 2 and 3 show the model fit to the epidemic data and the corresponding distributions of the reproduction number obtained from parametric bootstrap of the model best fit using 14, 21 and 28 epidemic days of data and a generation interval of 3 and 6 days, respectively. Following a generation interval of 3 days [**?**, 45], the reproduction number was estimated to be ~ 2 (SD 0.1) using the first 14 days and ~ 1.6 (SD 0.03) using the first 21 epidemic days.

References

[1] F. Brauer *Age of infection models and the final size relation*, Math. Biosc. & Eng. **5** (2008), 681-690.

[2] B. Davoudi, B. Pourbohloul, J.C. Miller, R. Meza, L.A. Meyers, and D.J.D. Earn *Early real-time estimation of infectious disease reproduction number* arXiv 0905 0728 (2009).

[3] O. Diekmann and J. A. P. Heesterbeek *Mathematical Epidemiology of Infectious Diseases*, Wiley, Chichester (2000).

[4] C. Fraser, C.A. Donnelly, S. Cauchemez, W.P. Hanage,M.D. Van Kerkhove, T.D. Hollingsworth, J. Griff, R.F. Baggaley, H.E. Jenkins, E.J. Lyons, T. Jom-

bart, W.R. Hinsley, N.C. Grassly, F. Balloux, A.C. Ghani, and N.M. Ferguson Pandemic potential od a strain of influenza A (H1N1): Early findings Science **3**24 (2009), 1557–1561.

[5] A. Gumel, S. Ruan, T. Day, J. Watmough, P. van den Driessche, F. Brauer, D. Gabrielson, C. Bowman, M.E. Alexander, S. Ardal, J. Wu and B.M. Sahai *Modeling strategies for controlling SARS outbreaks based on Toronto, Hong Kong, Singapore and Beijing experience*, Proc. Roy. Soc. London B **2**71 (2004), 2223-2232.

[6] W.O. Kermack and A.G McKendrick *A contribution to the mathematical theory of epidemics*, Proc. Royal Soc. London **115** (1927), 700–721.

[7] J.J. Levin and D.F, Shea *On the asymptotic behavior of the bounded solutions of some integral equations, I, II, III*, J. Math. Anal. & Appl. **37** (1972), 42–82, 288–326, 537–575

[8] J. Wallinga and M. Lipsitch *How generation intervals shape the relationship between growth rates and reproductive numbers*, Proc. Royal Soc. B **274** (2007), 599–604.

[9] P. Yan and Z. Feng *Variability order of the latent and the infectious periods in a deterministic $SEIR$ epidemic model and evaluation of control effectiveness* Math. Biosc. **224** (2010), 43–52.

[10] G. Chowell, F. Brauer. The basic reproduction number of infectious diseases: Computation and estimation using compartmental epidemic models. G. Chowell,

J.M. Hyman, L.M.A. Bettencourt, C. Castillo-Chavez (Eds.) Mathematical and Statistical Estimation Approaches in Epidemiology. Springer. 2009 (pdf)

[11] H. Nishiura, G. Chowell.The effective reproduction number as a prelude to statistical estimation of time-dependent epidemic trends. G. Chowell, J.M. Hyman, L.M.A. Bettencourt, C. Castillo-Chavez (Eds.) Mathematical and Statistical Estimation Approaches in Epidemiology. Springer. 2009 (pdf)

[12] Dietz K (1993) The estimation of the basic reproduction number for infectious diseases. *Statistical Methods in Medical Research* 2:23-41.

[13] De Jong MC, Diekmann O and Heesterbeek JA (1994) The computation of R_0 for discrete-time epidemic models with dynamic heterogeneity. *Mathematical Biosciences* 119:97-114.

[14] G. Chowell, H. Nishiura (2008) Quantifying the transmission potential of pandemic influenza. *Physics of Life Reviews* 5, 50-77.

[15] Anderson RM and May RM (1991) Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford

[16] Davidian M and Giltinan DM (1995) Nonlinear Models for Repeated Measurement data. Monographs on Statistics and Applied Probability 62. Chapman and Hall, New York.

[17] Bickel P and Doksum KA (1977) Mathematical Statistics. Holden-Day, Oakland, California

[18] Efron B and Tibshirani RJ (1986) Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1:54-75.

[19] Cunha BA (2004) Influenza: historical aspects of epidemics and pandemics. *Infectious Disease Clinics of North America* 18:141-155.

[20] Ma J. and Earn DJ (2006) Generality of the final size formula for an epidemic of a newly invading infectious disease. *Bulletin of Mathematical Biology* 68:679-702.

[21] Sydenstricker E (1921) Variations in case fatality during the influenza epidemic of 1918. *Public Health Reports* 36:2201-2211.

[22] Markel H, Lipman HB, Navarro JA, Sloan A, Michalsen JR, Stern AM and Cetron MS (2007) Nonpharmaceutical interventions implemented by US cities during the 1918-1919 influenza pandemic. *JAMA*298:644-654.

[23] EW (2005) Co-operarion in All Human Endeavour: Quarantine and Immigrant Disease Vectors in the 1918-1919 Influenza Pandemic in Winnipeg, Canadian Bulletin of Medical History 22:57-82.

[24] Andreasen V, Viboud C and Simonsen L (2008) Epidemiologic characterization of the summer wave of the 1918 influenza pandemic in Copenhagen: Implications for pandemic control strategies. *Journal of Infectious Diseases* **197**:270-8.

[25] Chowell G, Ammon CE, Hengartner NW and Hyman JM (2006) Transmission Dynamics of the Great Influenza Pandemic of 1918 in Geneva, Switzerland: As-

sessing the Effects of Hypothetical Interventions. *Journal of Theoretical Biology* 241:193-204.

[26] Chowell G, Nishiura H and Bettencourt LM (2007) Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. *Journal of the Royal Society Interface* 4:155-166.

[27] Chowell G, Miller MA and Viboud C (2007) Seasonal influenza in the United States, France, and Australia: transmission and prospects for control. *Epidemiology and Infection* in press.

[28] Chowell G, Ammon CE, Hengartner NW and Hyman JM (2007) Estimating the reproduction number from the initial phase of the Spanish flu pandemic waves in Geneva, Switzerland. *Mathematical Biosciences and Engineering* 4:457-470.

[29] Chowell G, Bettencourt LMA, Johnson N, Alonso WJ and Viboud C (2008) The 1918-1919 influenza pandemic in England and Wales: Spatial patterns in transmissibility and mortality impact *Proceedings of the Royal Society B* 275:501-509.

[30] Chowell G, Bettencourt LMA, Johnson NPAS, Alonso WJ and Viboud C (2008) The 1918-19 influenza pandemic in England and Wales: Spatial patterns in transmissibility and mortality impact submitted

[31] Diekmann O and Heesterbeek JAP (2000) Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation. John Wiley and Sons, New York.

[32] Heesterbeek JAP (2002) A brief history of R_0 and a recipe for its calculation. *Acta Biotheoretica* 50:189-204.

[33] Heffernan JM, Smith RJ and Wahl LM (2005) Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface* 2:281-293.

[34] Massad E, Burattini MN, Coutinho FA and Lopez LF (2007) The 1918 influenza A epidemic in the city of Sao Paulo, Brazil. *Medical Hypotheses* 68:442-445.

[35] Mills CE, Robins JM and Lipsitch M (2004) Transmissibility of 1918 pandemic influenza. *Nature* 432:904-906.

[36] Murray CJ, Lopez AD, Chin B, Feehan D and Hill KH (2006) Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918-20 pandemic: a quantitative analysis. *Lancet* 368:2211-2218.

[37] H. Nishiura, Mathematical and statistical analyses of the spread of dengue. *Dengue Bulletin* **30** (2006) 51-67.

[38] Nishiura H and Inaba H (2007) Discussion: Emergence of the concept of the basic reproduction number from mathematical demography. *Journal of Theoretical Biology* 244:357-364.

[39] Nishiura H (2007) Time variations in the transmissibility of pandemic influenza in Prussia, Germany, from 1918-19. *Theoretical Biology and Medical Modelling* 4:20.

[40] Sattenspiel L and Herring DA (2003) Simulating the effect of quarantine on the spread of the 1918-19 flu in central Canada. *Bulletin of Mathematical Biology* 65:1-26.

[41] Sertsou G, Wilson N, Baker M, Nelson P and Roberts MG (2006) Key transmission parameters of an institutional outbreak during the 1918 influenza pandemic estimated by mathematical modelling. *Theoretical Biology and Medical Modelling* 3:38.

[42] Viboud C, Tam T, Fleming D, Handel A, Miller MA and Simonsen L (2006) Transmissibility and mortality impact of epidemic and pandemic influenza, with emphasis on the unusually deadly 1951 epidemic. *Vaccine* 24:6701-6707.

[43] Vynnycky E, Trindall A and Mangtani P (2007) Estimates of the reproduction numbers of Spanish influenza using morbidity data. *International Journal of Epidemiology* 36:881-889.

[44] Wallinga J, Lipsitch M (2007) How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings B* 274:599-604 .

[45] White LC, Pagano MA (2007) likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. Stat Med; in press (doi: 10.1002/sim.3136).

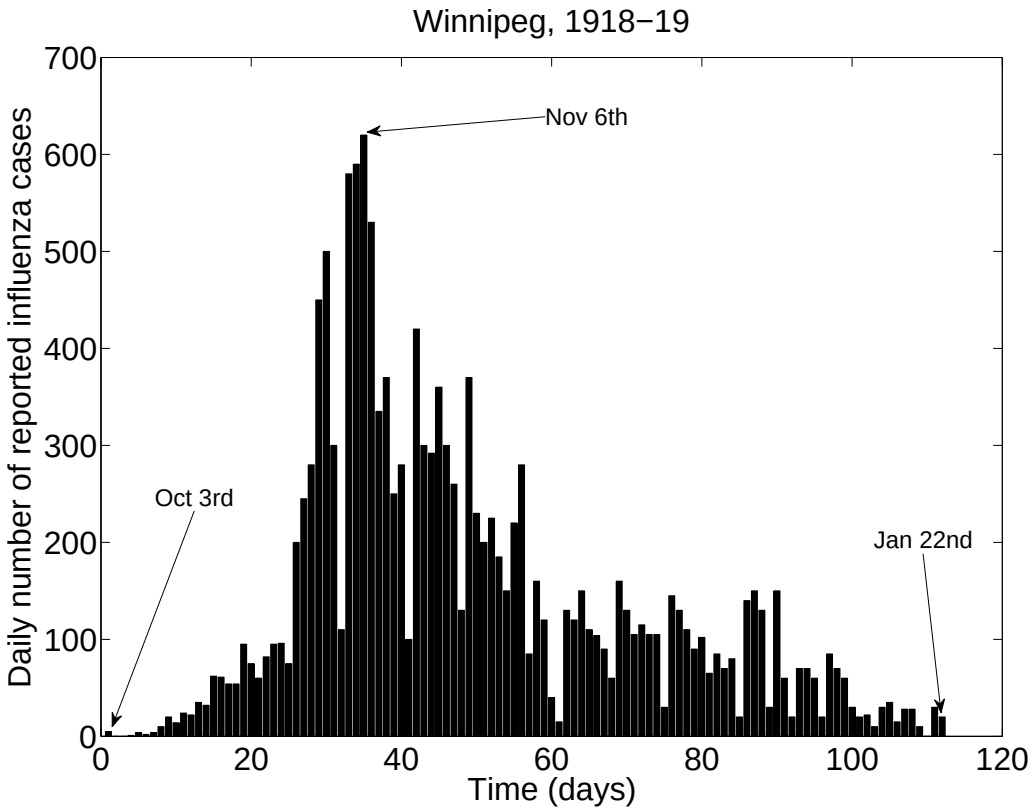

Figure 1: Temporal distribution of Spanish influenza in Winnipeg, Canada in 1918. A total of 14868 cases were reported from October 3rd to January 22nd. Data source: [**?**]

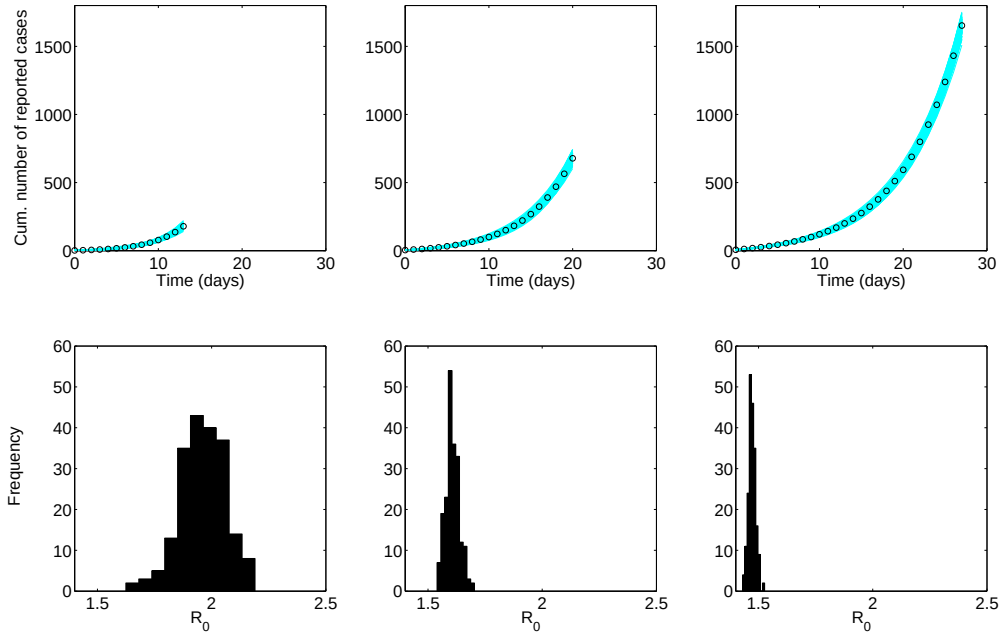

Figure 2: Model fits (top panels) and the resulting distributions of the reproduction number (bottom panels) obtained assuming a generation interval of 3 days after fitting the simple SEIR epidemic model to the initial phase of the Fall influenza wave using 14, 21 and 28 epidemic days of the Spanish Flu Pandemic in Winnipeg, Canada. In the top panel, the epidemic data of the cumulative number of reported influenza cases are the circles and the solid blue lines are 200 realizations of the model fit to the data obtained through parametric bootstrapping as explained in the text.

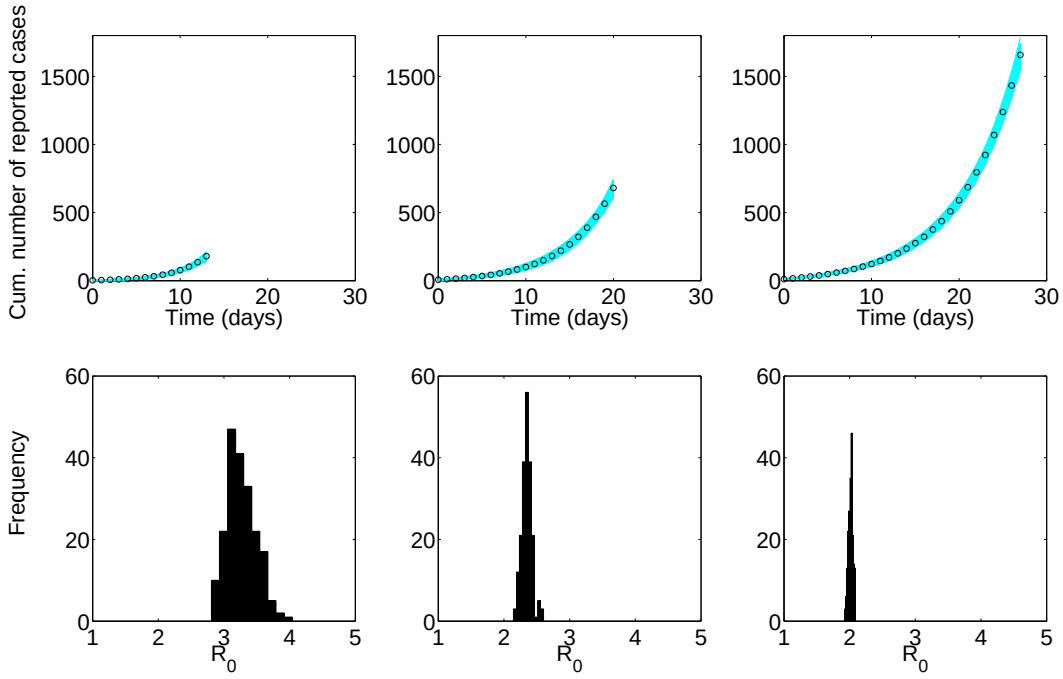

Figure 3: Model fits (top panels) and the resulting distributions of the reproduction number (bottom panels) obtained assuming a generation interval of 6 days after fitting the simple SEIR epidemic model to the initial phase of the Fall influenza wave using 14, 21 and 28 epidemic days of the Spanish Flu Pandemic in Winnipeg, Canada. In the top panel, the epidemic data of the cumulative number of reported influenza cases are the circles and the solid blue lines are 200 realizations of the model fit to the data obtained through parametric bootstrapping as explained in the text.