MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **7**, Number **4**, October **2010**

doi:10.3934/mbe.2010.7.851

pp. **851–869**

TIME VARIATIONS IN THE GENERATION TIME OF AN INFECTIOUS DISEASE: IMPLICATIONS FOR SAMPLING TO APPROPRIATELY QUANTIFY TRANSMISSION POTENTIAL

Hiroshi Nishiura

[1] PRESTO, Japan Science and Technology Agency (JST)
4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan
[2] Theoretical Epidemiology, University of Utrecht
Yalelaan 7, 3584 CL, Utrecht, The Netherlands

(Communicated by Jia Li)

Abstract. Although the generation time of an infectious disease plays a key role in estimating its transmission potential, the impact of the sampling time of generation times on the estimation procedure has yet to be clarified. The present study defines the period and cohort generation times, both of which are time-inhomogeneous, as a function of the infection time of secondary and primary cases, respectively. By means of analytical and numerical approaches, it is shown that the period generation time increases with calendar time, whereas the cohort generation time decreases as the incidence increases. The initial growth phase of an epidemic of Asian influenza A (H2N2) in the Netherlands in 1957 was reanalyzed, and estimates of the basic reproduction number, R_0, from the Lotka-Euler equation were examined. It was found that the sampling time of generation time during the course of the epidemic introduced a time-effect to the estimate of R_0. Other historical data of a primary pneumonic plague in Manchuria in 1911 were also examined to help illustrate the empirical evidence of the period generation time. If the serial intervals, which eventually determine the generation times, are sampled during the course of an epidemic, direct application of the sampled generation-time distribution to the Lotka-Euler equation leads to a biased estimate of R_0. An appropriate quantification of the transmission potential requires the estimation of the cohort generation time during the initial growth phase of an epidemic or adjustment of the time-effect (e.g., adjustment of the growth rate of the epidemic during the sampling time) on the period generation time. A similar issue also applies to the estimation of the effective reproduction number as a function of calendar time. Mathematical properties of the generation time distribution in a heterogeneously mixing population need to be clarified further.

1. **Introduction.** Understanding the time intervals between successive generations of infected individuals is crucial to appropriately quantify the transmission dynamics of infectious diseases. The generation time is defined as the time interval between infection of a primary case and infection of a secondary case caused by the primary case [40]. Similarly, the defined interval of the onset event, the serial interval, is the time interval between onset of a primary case and onset of a secondary case generated by the primary case [12]. Although it is difficult to directly observe the actual

2000 *Mathematics Subject Classification.* Primary: 00A71, 37N25; Secondary: 92C60, 92D30.
Key words and phrases. Transmission, epidemiology, model, influenza, plague.
The work of H Nishiura was supported by the JST PRESTO program.

time of infection of a non-sexual directly transmitted disease, the serial interval can be partly observed in practice (e.g., based on contact tracing which indicates who acquired infection from whom with calendar times of onset among traced cases [25, 36] or on the time intervals between the onset of the first and of the subsequent cases in households [3]), which eventually elucidates the generation time [40].

The generation-time distribution is known to play a key role in estimating the transmission potential of a disease [29, 38, 45] which is measured by the basic reproduction number, R_0, defined as the average number of secondary cases generated by a single primary case in a fully susceptible population [8, 9, 10]. From the initial growth phase of an epidemic, the intrinsic growth rate, r_0, i.e., the natural rate of increase in infected individuals [11], is estimated, and R_0 is subsequently estimated using the Lotka-Euler equation:

$$\frac{1}{R_0} = \int_0^\infty \exp(-r_0\sigma) g_0(\sigma)\, d\sigma \tag{1}$$

where $g_0(\sigma)$ is the probability density of the generation time of length σ. In many instances (e.g., see [7, 25, 27, 44]), R_0 is inferred by using the estimate of r_0 and by assuming that the generation-time distribution is known, and in addition, by implicitly ignoring the time-dependency of the generation time [35]. Nevertheless, the estimation methods of the generation time and its sampling scheme have yet to be firmly developed, and hence, advantages and disadvantages of the sampling time of generation time have to be clarified. A similar issue applies to the estimation of the effective reproduction number as a function of calendar time t, $R(t)$, defined as the average number of secondary cases per primary case at time t [4, 17]. To date, several different algorithms have been proposed to transform the epidemic curve into $R(t)$ using the distribution of either generation time or serial interval [5, 6, 13, 15, 34, 46, 47], but only a few studies explicitly accounted for the time-inhomogeneity of the generation time or serial interval [4, 6, 21], and moreover, the impact of the time-inhomogeneity on the estimation procedure of $R(t)$ has been only partly clarified [4, 21].

To understand the concept and role of the generation time as a function of calendar time, the present study aimed to comprehensively discuss the time inhomogeneous generation time and to characterize the impact of the sampling time on the estimation of R_0. In the next section, epidemiological definitions of the generation time and the serial interval are discussed in light of the historical development of their concepts. The subsequent three sections are devoted to defining two different generation times as a function of calendar time, offering practical interpretations with respect to their increase and decrease with calendar time. A numerical illustration follows in section 6. A reanalysis of historical datasets of Asian influenza and pneumonic plague in section 7, suggest epidemiologic consequences of the time-inhomogeneity of the generation time on the estimation of R_0, and partly demonstrate empirical evidence of the time variation. A discussion and future implications follow in section 8.

2. Generation time and serial interval in epidemiology.

As briefly mentioned in the Introduction, the generation time and the serial interval are distinguished by their definitions; although both are concerned with the time-interval between a primary case and a secondary case directly infected from the primary case, the generation time represents the interval between their times of infection and the serial interval refers to the interval between their times of illness onset

[40]. To the best of the author's knowledge, this distinction was not explicitly made when Pickles [37] initially referred to the serial interval as the **transmission interval** with reference to empirical observations of a hepatitis epidemic in the United Kingdom. The followers of Pickles, most notably Hope Simpson [18] and Bailey [3], used the term serial interval and clearly defined it as the interval between successive illness onsets. The distinction became much clearer only recently, when Fine [12] reviewed that (i) the transmission interval (i.e., the interval between successive infections) and (ii) the clinical onset serial interval (i.e., the interval between successive clinical cases) are different both conceptually and quantitatively. In the present day, the former is referred to as the generation time, and the latter as the serial interval.

It is worth mentioning another issue of the epidemiological definition of the generation time in relation to model building. Anderson and May [2] considered the generation time as a sum of the latent period (i.e., non-infectious period) and the infectious period [7]. According to strict arguments by Fine [12], the sum of the latent and infectious periods is concerned with *the course of a single infection* (in a primary case), and thus, is different from the *interval between successive infections*. The sum of latent and infectious periods may be identical to the generation time if the contact frequency and infectiousness (e.g., conditional probability of transmission given a contact) are independent of the time since infection in a primary case, and this is particularly the case for many models written by ordinary differential equations [2]. Nevertheless, the infectiousness profile tends to vary during the course of infection in a primary case (e.g., HIV infection), and the generation time which strictly reflects the interval between successive infections may well be shorter than the sum of latent and infectious periods [48]. The generation time in the present study is derived from a simple class-age structured model, but refers to the interval between successive infections.

3. Infection-age structured Kermack and McKendrick model.

We first present the system of equations with which we define the time-inhomogeneous generation times. To appropriately capture the characteristics of time-inhomogeneity, and also, to reflect the interval between successive infections in the model, we used an infection-age structured Kermack and McKendrick model [22]. Hereafter, we refer to the time since infection of an infected individual as the **infection-age**. Although a model of Kermack and McKendrick, governed by ordinary differential equations, may be more widely known than that given by the McKendrick-von-Foerster equations (i.e., partial differential equations), it should be noted that the infection-age structured model was actually proposed and investigated in an initial publication in 1927 [22]. Indeed, the infection-age structured model has been shown to better capture the detailed intrinsic dynamics [14].

The following assumptions are made throughout this article: (i) the population is homogeneously mixing, (ii) the epidemic occurs in a fully susceptible population, (iii) the time scale of the epidemic is assumed to be sufficiently shorter than the average life expectancy at birth of the host, and the background demographic dynamics (i.e., natural birth and death) are ignored, (iv) the epidemic occurs in a closed constant population without immigration and emigration, (v) once an infected individual recovers, he/she becomes completely immune against further infections, and (vi) the initial attack size is extremely small as compared with the total population size N, so that the number of susceptible individuals at the beginning of an

epidemic ($t = 0$), $S(0)$ is sufficiently approximated to N.

Let us denote the numbers of susceptible and recovered individuals at calendar time t by $S(t)$ and $U(t)$ (we use the notation $U(t)$ for recovered individuals to avoid any confusion with the effective reproduction number at calendar time t, $R(t)$). Furthermore, let $i(t,\tau)$ be the density of infectious individuals at calendar time t and infection-age τ. In reality, these numbers have to be discrete (i.e., a positive integer) as is practiced using the stochastic Markov jump process. Here a deterministic approximation is made with constant population size $N = S(t)+\int_0^\infty (t,\tau)d\tau+U(t)$, and thus, for example, the approximated proportion susceptible is given by $S(t)/N$. The SIR (susceptible-infected-recovered) model is given by

$$\begin{aligned} \frac{dS(t)}{dt} &= -\lambda(t)S(t), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \tau}\right) i(t,\tau) &= -\gamma(\tau)i(t,\tau),\ i(t,0)=\lambda(t)S(t), \\ \frac{dU(t)}{dt} &= \int_0^\infty \gamma(\tau)i(t,\tau)d\tau, \end{aligned} \tag{2}$$

where $\lambda(t)$ is the force of infection at calendar time t which is given by

$$\lambda(t) = \int_0^\infty \beta(\tau)i(t,\tau)d\tau, \tag{3}$$

and $\gamma(\tau)$ is the recovery rate at infection-age τ and $\beta(\tau)$ is the effective transmission rate at infection-age τ. $\gamma(\tau)$ is intended to represent infection-age dependent rate of ending infectious period, while $\beta(\tau)$ reflects infection-age dependent variations in contact frequency and transmission probability per contact. In an SIR model written by ordinary differential equations, both are dealt as constant γ and β. To ease our interpretation, we integrate (2) as

$$i(t,\tau) = \begin{cases} \Gamma(\tau)j(t-\tau), & \text{for } t-\tau>0 \\ \frac{\Gamma(\tau)}{\Gamma(\tau-t)}j_0(\tau-t), & \text{for } \tau-t>0 \end{cases} \tag{4}$$

where

$$\begin{aligned} j(t) &= i(t,0), \\ \Gamma(\tau) &= \exp\left(-\int_0^\tau \gamma(\sigma)d\sigma\right), \end{aligned}$$

and $j_0(\tau)$ is the density of initially infected individuals at the beginning of an epidemic. By $j_0(\tau)$, we intend to represent imported infected individuals (e.g., from abroad) to the previously fully susceptible population. In the following discussion, we assume that the contribution of the infection-age of initially infected individuals to the epidemic is negligible. It should be noted that $j(t)$ is often referred to as incidence (i.e., the number of newly infected individuals at calendar time t).

We simplify our discussion using a renewal equation of $j(t)$ using the rate of secondary transmissions at time t and infection-age τ, i.e.,

$$j(t) = \int_0^\infty A(t,\tau)j(t-\tau)d\tau \tag{5}$$

where $A(t,\tau)$ is interpreted as the rate of secondary transmissions (or the transient number of secondary transmissions) per single primary case, at calendar time t, whose infection-age is τ. The classic mass action principle adopted by Kermack and McKendrick assumes that the non-linearity of an epidemic is characterized by

the depletion of susceptible individuals alone (i.e., contact and recovery rates are independent of calendar time). The assumption indicates that $A(t,\tau)$ is decomposed as

$$A(t,\tau) = S(t)\beta(\tau)\Gamma(\tau), \tag{6}$$

for the autonomous system (2). The importance of this decomposition was recently emphasized in relation to the generation time [48]. Since R_0 is the average number of secondary cases caused by a single primary case throughout his/her course of infection in a fully susceptible population ($t \to 0$), we get

$$R_0 = S(0)\int_0^\infty \beta(\tau)\Gamma(\tau)d\tau. \tag{7}$$

There are two different effective reproduction numbers as a function of calendar time t, the notations for which we followed Fraser [13] and Grassly and Fraser [16]. The first is the instantaneous reproduction number, expressed as the average number of secondary transmissions occurring at calendar time t, i.e.,

$$R(t) = \int_0^\infty A(t,\tau)d\tau. \tag{8}$$

The second is the cohort reproduction number, $R_c(t)$, representing the average number of secondary transmissions caused by those in an infection cohort having experienced infection at calendar time t.

$$R_c(t) = \int_0^\infty A(t+\tau,\tau)d\tau. \tag{9}$$

Preceding these definitions in infectious disease epidemiology, both $R(t)$ and $R_c(t)$ have been explicitly defined as the period and cohort total fertility rates, respectively, in mathematical demography [1]. The difference is highlighted when a specific event at calendar time t occurs (e.g., a public health intervention starts at calendar time t). Then, $R(t)$ abruptly changes (e.g., declines) with calendar time t, but $R_c(t)$ smoothly changes, because $R_c(t)$ smoothes out the timing (i.e., infection-age) of secondary transmissions among a cohort who experienced infection at calendar time t [33].

4. **Period and cohort generation times.** As mentioned above, the generation time distribution in the present study is considered using a model which rests on the homogeneously mixing assumption. We now follow with the definition of two generation times, both of which are a function of calendar time t. As we illustrate in Fig. 1, it is possible to define two generation times at calendar time t during the course of an epidemic. If we regard calendar time t as the infection time of secondary cases, the generation time is calculated backward from time t, which we refer to as the period generation time of length τ, $g_p(t,\tau)$. The period generation time is regulated by the incidence prior to calendar time t, because the frequency distribution of primary cases (i.e., incidence) varies with calendar time. On the other hand, if the calendar time t is regarded as the infection time of primary cases (i.e., a cohort of infected individuals born at calendar time t), the forward calculation applies to the generation time, which we refer to as the cohort generation time of length σ, $g_c(t,\sigma)$. The number of susceptible individuals declines with calendar time, which influences the length of the cohort generation time.

Accordingly, the period generation-time distribution reflects the relative frequency of secondary transmissions occurring at calendar time t by the primary

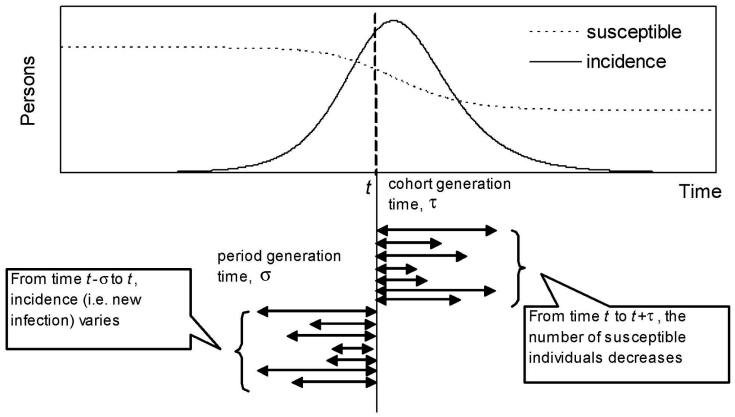

FIGURE 1. **The concept of the time-inhomogeneous generation time.** Two different generation times are illustrated with respect to calendar time t. The forward calculation of the generation time τ, which starts with the infection time of primary cases at calendar time t, is referred to as the cohort generation time. Future decline in susceptible individuals with calendar time influences the length of τ. The backward calculation of the generation time σ, which considers the infection time of secondary cases at calendar time t, is referred to as the period generation time. Variation in incidence with calendar time characterizes the frequency of primary cases at calendar time $t-\sigma$ and thus, influences the length of σ.

cases at infection-age τ. Since incidence $j(t)$ varies with calendar time, the instantaneous measure of the generation time does not only depend on the rate of secondary transmission at calendar time t and infection-age τ (i.e., $A(t,\tau)$) but also incidence $j(t-\tau)$, i.e.,

$$g_p(t,\tau) = \frac{A(t,\tau)j(t-\tau)}{\int_0^\infty A(t,\sigma)j(t-\sigma)d\sigma}. \tag{10}$$

On the other hand, the distribution of cohort generation time, $g_c(t,\tau)$, represents the relative frequency of secondary transmissions caused by those who were infected at calendar time t and are at infection-age τ, i.e.,

$$g_c(t,\tau) = \frac{A(t+\tau,\tau)}{\int_0^\infty A(t+\sigma,\sigma)d\sigma}, \tag{11}$$

which was computed in a different way in a recent study, for the reconstruction of a transmission network from an epidemic curve using a stochastic model [21]. Of course, the mean period and cohort generation times at calendar time t, $T_p(t)$ and $T_c(t)$, follow equations (10) and (11):

$$\begin{aligned} T_p(t) &= \int_0^\infty \tau g_p(t,\tau)d\tau, \\ T_c(t) &= \int_0^\infty \tau g_c(t,\tau)d\tau. \end{aligned} \tag{12}$$

Equations (10) and (11) also directly offer their interpretations. Equation (10) is rearranged as

$$A(t,\tau)j(t-\tau) = g_p(t,\tau)j(t). \tag{13}$$

Both sides indicate the transient number of secondary transmissions occurring at calendar time t caused by those at infection-age τ. Similarly, equation (11) is rearranged as

$$R_c(t)g_c(t,\tau) = A(t+\tau,\tau). \tag{14}$$

Both sides indicate the absolute number of secondary transmissions per single primary case at infection-age τ who belongs to a cohort having experienced infection at calendar time t.

Accordingly, employing a classic assumption of Kermack and McKendrick (i.e., equation (6)), the mean period generation time at calendar time t is

$$T_p(t) = \frac{\int_0^\infty \tau\beta(\tau)\Gamma(\tau)j(t-\tau)d\tau}{\int_0^\infty \beta(\sigma)\Gamma(\sigma)j(t-\sigma)d\sigma}, \tag{15}$$

indicating that the time-effect is caused by an increase (or decrease) in incidence prior to calendar time t. Similarly, the mean cohort generation time at calendar time t is,

$$T_c(t) = \frac{\int_0^\infty \tau S(t+\tau)\beta(\tau)\Gamma(\tau)d\tau}{\int_0^\infty S(t+\sigma)\beta(\sigma)\Gamma(\sigma)d\sigma}, \tag{16}$$

suggesting that the rate of depletion of susceptible individuals generates the time-effect. These mean estimates will be used in Sections 6 and 7.

5. **Comparison of generation times.** Using equation (6), the previously and frequently used generation-time distribution $g_0(\tau)$ in the Lotka-Euler equation (1) (e.g., see [45]), which was implicitly assumed to be independent of calendar time, is expressed as

$$g_0(\tau) = \frac{\beta(\tau)\Gamma(\tau)}{\int_0^\infty \beta(\sigma)\Gamma(\sigma)d\sigma}. \tag{17}$$

As an important premise for further discussions, it must be noted that we assume that the normalization of the rate of infection $A(\tau)$ (or alternatively, $A(t+\tau,\tau)$ or $A(t,\tau)j(t-\tau)$ in the case of time-inhomogeneous generation times in equations (10) and (11)) results in the probability density function of the generation time; the assumption implies that the infection process $A(\tau)$ has a Markov property. Most likely, this is not the case if the rate $A(\tau)$ is decomposed to the product of $\beta(\tau)$ and $\Gamma(\tau)$, resulting in a need to account for dependency between R_0 and $g_0(\tau)$ [48]. In the present study, we ignore this issue for simplicity (see Section 8).

Classically, $g_0(\tau)$ was referred to as the cohort generation time in mathematical demography [23]. However, it is inappropriate to deem $g_0(\tau)$ as the cohort generation time in infectious disease epidemiology, since a depletion of susceptible individuals and an increase (and decrease) in infected individuals generate the nonlinear dynamics (i.e., dependent happening). Thus, for clarity, $g_0(\tau)$ is hereafter referred to as the basic generation time.

The basic generation time, $g_0(\tau)$, is the special case of both $g_p(t,\tau)$ and $g_c(t,\tau)$. As for $g_p(t,\tau)$, assuming that the incidence $j(t)$ grows exponentially with an intrinsic growth rate r_0, i.e., $j(t) = k\exp(r_0 t)$ where k is a constant (so that we have the

density $i(t,\tau) = B\exp[r_0(t-\tau)]\Gamma(\tau)$ where B is a scaling factor), the equation (10) is rearranged as

$$g_p(t,\tau) = \frac{\exp(-r_0\tau)\beta(\tau)\Gamma(\tau)}{\int_0^\infty \exp(-r_0\sigma)\beta(\sigma)\Gamma(\sigma)d\sigma}. \tag{18}$$

If the SIR model (2) with demographic dynamics (i.e., an SIR system with birth and death rates of the host) reaches a stationary state (i.e., $r = 0$), $g_p(t,\tau)$ is identical to $g_0(\tau)$. In addition, from the Lotka-Euler equation (1) and R_0 in equation (7), the denominator of equation (18) is $1/S(0)$. Thus, $T_p(t)$ with stable growth of infected individuals is the monotonically decreasing function with calendar time t for $r_0 > 0$, and thus is shorter than the mean basic generation time. Because of the reflection of the frequency of mothers' age distribution at delivery (which is equivalent to infection-age distribution of secondary transmission among primary cases) at calendar time t, $T_p(t)$ has been referred to as **the mean age of childbearing in the stable population** in mathematical demography [23].

The relationship between $g_0(\tau)$ and $g_p(t,\tau)$ can be further simplified by applying our concept to an SIR model given by ordinary differential equations where the infectious period is exponentially distributed with mean $1/\gamma$. In the SIR model, $\beta(\tau)$ is constant β and $\Gamma(\tau) = \exp(-\gamma\tau)$. Thus, $g_0(\tau)$ in equation (17) should read

$$g_0(\tau) = \frac{\beta\exp(-\gamma\tau)}{\int_0^\infty \beta\exp(-\gamma\sigma)d\sigma} = \gamma\exp(-\gamma\tau). \tag{19}$$

Thus, the mean basic generation time (or, equivalently, the mean period generation time with **zero** intrinsic growth rate), T_0, is $1/\gamma$. For T_p, the mean period generation time with an exponential growth in incidence, we consider the linearized version in equation (18), i.e.,

$$\begin{aligned} T_p &= \beta S(0)\int_0^\infty \tau\exp(-r_0\tau)\exp(-\gamma\tau)d\tau \\ &= \frac{\beta S(0)}{(r_0+\gamma)^2} \end{aligned} \tag{20}$$

From the Lotka-Euler equation (1) and R_0 in equation (7),

$$\int_0^\infty \exp(-r_0\tau)\beta\exp(-\gamma\tau)d\tau = \frac{\beta}{r_0+\gamma} = \frac{1}{S(0)}. \tag{21}$$

Thus, we can simplify equation (20) to

$$T_p = \frac{1}{r_0+\gamma} \tag{22}$$

which reasonably confirms that $T_p = T_0$ for $r_0 = 0$ and $T_p < T_0$ for $r_0 > 0$.

The cohort generation time, $g_c(t,\tau)$, can be related to $g_0(\tau)$ when we consider the initial growth phase of an epidemic. Using equation (6), $g_c(t,\tau)$ is rewritten as

$$g_c(t,\tau) = \frac{S(t+\tau)\beta(\tau)\Gamma(\tau)}{\int_0^\infty S(t+\sigma)\beta(\sigma)\Gamma(\sigma)d\sigma}. \tag{23}$$

As mentioned above, we ignore the contribution of the small number of initially infected individuals j_0 to the infection-age distribution of all infected individuals for simplicity. Assuming that the length of generation time is sufficiently shorter than the calendar timescale of an epidemic (so that multiple generations of cases exist during the early exponential growth phase), and moreover, assuming that the depletion of susceptible individuals is negligible during the initial growth phase, we

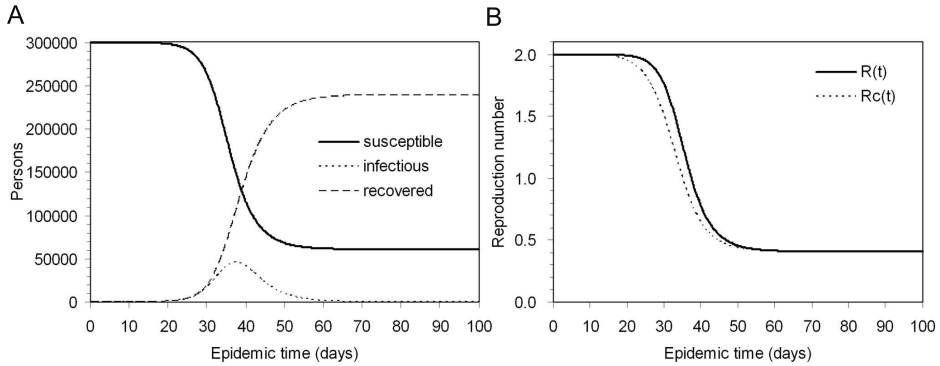

FIGURE 2. **Time course of an epidemic in a homogeneously mixing population.** (A) Epidemic curve of susceptible, infectious and recovered individuals in a population of 300,000. R_0 is assumed as 2. (B) Instantaneous and cohort reproduction numbers estimated from the epidemic curve in A. In both panels, the infectious period (i.e., the basic generation time) is assumed to follow an exponential distribution with a mean of 3 days.

approximate $S(t+\tau) \to S(0)$ for $t \to 0$. This leads equation (23) to (17), and thus, the mean basic generation time, T_0, might also be deemed the mean *initial* cohort generation time. Except for the initial growth phase, it is difficult to make an explicit analytical comparison, but equations (14) and (17) yield the relationship between $g_c(t,\tau)$ and $g_0(\tau)$, i.e.,

$$R_c(t)g_c(t,\tau) = R_0\frac{S(t+\tau)}{S(0)}g_0(\tau), \tag{24}$$

under homogeneously mixing assumption. That is, adopting the classic assumption of Kermack and McKendrick (i.e., equation (6)), equation (24) suggests that the absolute number of secondary transmissions per single primary case at infection-age τ who is in a cohort having experienced infection at calendar time t, is expressed as the absolute number of secondary transmissions per single primary case at infection-age τ during the initial phase of an epidemic, weighted by the density of susceptible individuals at calendar time $t+\tau$. This argument shares a problem in common with parameterization of the survivorship function of infected individuals with HIV [20, 24]; i.e., the survivorship data collected with calendar time are right-truncated.

6. **Numerical illustrations.** Here we numerically illustrate the time inhomogeneous generation times using the simplest SIR model (i.e., an SIR model for a homogeneously mixing population without a latent period and with an exponentially distributed infectious period) given by

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\beta S(t)I(t), \\
\frac{dI(t)}{dt} &= \beta S(t)I(t) - \gamma I(t), \\
\frac{dU(t)}{dt} &= \gamma I(t),
\end{aligned} \tag{25}$$

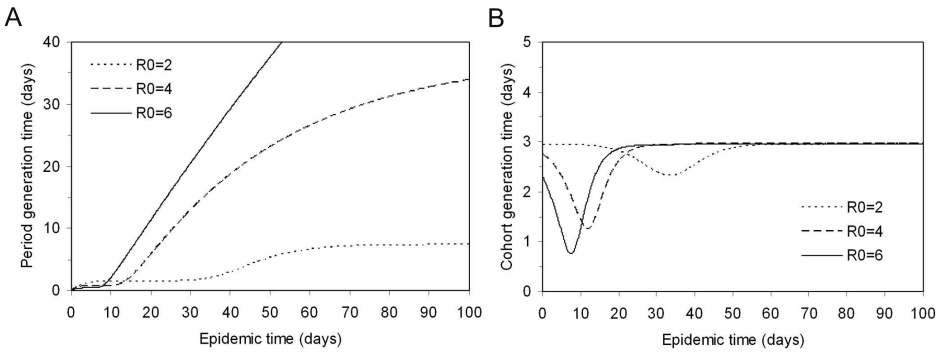

FIGURE 3. **Time course of the mean period and cohort generation times.** Mean period (A) and cohort generation times (B) as a function of calendar time are illustrated, assuming R_0 as 2, 4 and 6. In both panels, the infectious period (i.e., the basic generation time) is assumed to follow an exponential distribution with a mean of 3 days.

where $I(t)$ is the number of infectious individuals at calendar time t. Again it should be noted that we use $U(t)$ for recovered individuals to avoid any confusion with the instantaneous reproduction number, $R(t)$. We assumed that one index case successfully invaded into a new community with a total size N of 300,000 individuals. Except for the index case, all are assumed susceptible to a disease at calendar time 0 and experience homogeneous mixing. The mean infectious period, $1/\gamma$, is 3 days, which is equivalent to the mean basic generation time. The transmission rate, β, was calculated as $\beta = R_0\gamma/N$ where we assumed three different values for R_0 of 2, 4 and 6. The instantaneous and cohort reproduction numbers, $R(t)$ and $R_c(t)$, are computed as

$$R(t) = \int_0^\infty \beta S(t)\exp(-\gamma\tau)d\tau = \frac{\beta S(t)}{\gamma} = \frac{S(t)}{S(0)}R_0, \tag{26}$$

$$R_c(t) = \beta\int_0^\infty S(t+\tau)\exp(-\gamma\tau)d\tau. \tag{27}$$

Fig. 2A shows the temporal dynamics of an epidemic ($S(t), I(t)$ and $U(t)$) assuming $R_0 = 2$. Corresponding to the epidemic curve, Fig. 2B illustrates qualitative patterns of $R(t)$ and $R_c(t)$, confirming that $R_c(t)$ declines earlier than $R(t)$ reflecting future decline in susceptible individuals at calendar time t (i.e., reflecting $S(t+\tau)$ rather than $S(t)$).

We further computed the mean period and cohort generation times at calendar time t, $T_p(t)$ and $T_c(t)$ by

$$T_p(t) = \frac{\int_0^\infty \tau\exp(-\gamma\tau)j(t-\tau)d\tau}{\int_0^\infty \exp(-\gamma\sigma)j(t-\sigma)d\sigma} \tag{28}$$

and

$$T_c(t) = \frac{\int_0^\infty \tau S(t+\tau)\exp(-\gamma\tau)d\tau}{\int_0^\infty S(t+\sigma)\exp(-\gamma\sigma)d\sigma}. \tag{29}$$

$T_p(t)$ and $T_c(t)$ with different R_0 are comparatively shown in Figs. 3A and 3B. Whereas $T_p(t)$ is an increasing function with calendar time t, $T_c(t)$ is decreased once during the epidemic and recovers close to the mean estimate of 3 days. The greater the R_0, the longer would be the mean period generation time during the late phase of an epidemic. $T_p(t)$ exceeds T_0 shortly before observing the peak incidence. On the other hand, the mean cohort generation time is always shorter than the mean basic generation time, and the magnitude of the decline during the course of an epidemic corresponds to the value of R_0.

Fig. 4A further illustrates the mean period generation time as a function of the period reproduction number. Here, $T_p(t)$ and $R(t)$ are, respectively, monotonically increasing and decreasing functions with calendar time t. The non-linear relationship in Fig. 4A is seen, because $T_p(t)$ is an inverse function of the growth rate of an epidemic at calendar time t (i.e., equation (22)) and $R(t)$ indirectly reflects the growth rate of infection at calendar time t which is a monotonically decreasing function. It should be noted that the monotonically increasing $T_p(t)$ is consistent with numerical illustrations given by Burr and Chowell [4]. Fig. 4B shows contraction of the mean cohort generation time which coincides with the increase in incidence, $j(t)$. Although the prevalence, $I(t)$, was previously suggested to coincide with $T_c(t)$ [21], it is worth noting that the incidence more precisely captures the decline in $T_c(t)$, because the decrease in future susceptible individuals at calendar time $t+\tau$ (or the rate of the decline in susceptible individuals) depends on new infection (i.e., incidence) at calendar time $t+\tau$. Therefore, the magnitude of contraction in $T_c(t)$ reflects the absolute incidence, because there is a competition among an infection cohort (who were infected at calendar time t, $j(t)$) in finding the susceptible individuals at calendar time $t+\tau$, which becomes harder in the population when $j(t)$ is greater (and thus, the cohort generation time is shortened by an increase in $j(t)$).

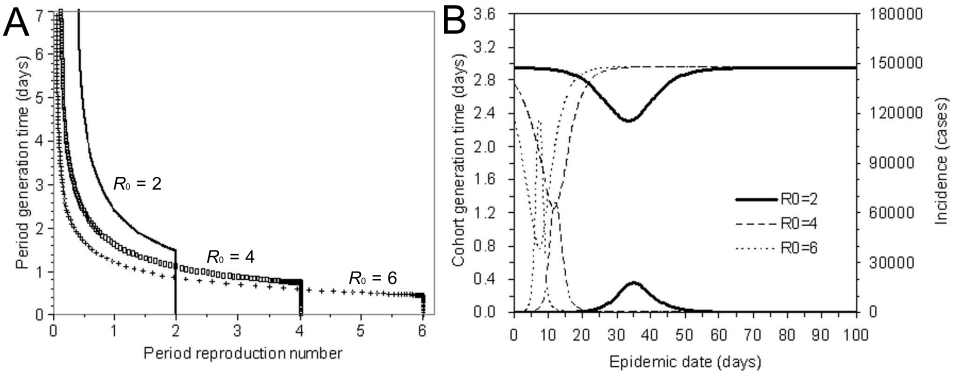

FIGURE 4. **Properties of the mean period and cohort generation times.** (A) Relationship between the mean period generation time and the period reproduction number. R_0 was assumed as 2, 4 or 6. The infectious period is assumed to follow an exponential distribution with a mean of 3 days. (B) Relationship between the mean cohort generation time and incidence (i.e., the number of new infections) at calendar time t. The parameter values are equivalent to those in A.

7. **Estimation of R_0.** Here we examine two different historical datasets, i.e., epidemics of influenza and pneumonic plague. The influenza dataset is the daily number of deaths during an epidemic of Asian influenza A (H2N2) in the Netherlands in 1957 [28]. From 1 September to 30 November, 1957, a total of 1,230 influenza deaths were reported (Fig. 5A). Examining the initial growth phase in detail, we assess the impact of the sampling time of generation time on the estimate of R_0 using the Lotka-Euler equation (1). Fig. 5B compares the observed and predicted values of the daily number of deaths during the first 18 days. Assuming Poisson-distributed errors, the intrinsic growth rate, r_0, is estimated as 0.131 (95% confidence interval: 0.113, 0.147) per day.

Using the estimate of r_0 and assuming that the distribution of the basic generation-time, $g_0(\tau)$, is known, the Lotka-Euler equation can yield an estimate of R_0. However, as we discussed in the Introduction, known generation time is in practice often informed by the serial interval from the contact tracing procedure, which is performed during the course of an epidemic. Although it is difficult to directly observe the generation time, G, the serial interval, S, is thus partly observed from the contact tracing, and moreover, using the incubation periods of primary cases, F_p, and secondary cases, F_s, the generation time is given by [40]

$$G = S + F_p - F_s \tag{30}$$

as long as two conditions are met; i.e., (i) there is no asymptomatic transmission [19] and (ii) there is no dependency between the incubation period and the generation time [12, 40]. Although both may not be the case for various directly transmitted viral diseases, we assume at least for now that the equation (30) is justified. We now consider the impact of mistakenly regarding $g_p(t,\tau)$ or $g_c(t,\tau)$ as $g_0(\tau)$ in the estimation of R_0 using the Lotka-Euler equation (1). Although unrealistic, here we assume the simplest case where the Asian influenza epidemic in Fig. 5A was fully described by the SIR model (25) with 300,000 initially fully susceptible individuals among whom homogeneous mixing takes place and with an exponentially distributed infectious period (i.e., the basic generation time) with mean $1/\gamma = 3$ days. Again, we ignore the underlying demographic dynamics. From the Lotka-Euler equation (1), the unbiased estimate of R_0 is thus $1 + r_0/\gamma = 1.39$. If the autonomous system (25) fully captured the dynamics, the final size would have been 50.6% (i.e., 151,867 cases). Furthermore, assuming that 70% of infection resulted in symptomatic disease [26, 42], this would yield a (perhaps overestimated) case fatality of 1.16%.

We assume that the generation times are sampled during the middle phase of an epidemic (i.e., Days 60, 80, 100 and 120), because the prevalence is high in that stage and contact tracing likely happens then (Fig. 5C). The contact tracing during the midst of an epidemic tends to fail to identify all the contacts, but we ignore the issue of incomplete observation and any other relevant epidemiological problems for simplicity. Days 60 and 80 are before the peak incidence, whereas Days 100 and 120 are shortly after the peak of the epidemic curve. We assume that the serial intervals (which will determine the generation times) are sampled at these points of time; although in reality the sampled serial intervals may reflect observation at a certain interval $[t_1, t_2]$ (and thus, reflects the generation time in this period weighted by the sampling frequency with calendar time), we ignore this issue for simplicity. Fig. 5D shows the consequence of the sampling time on the estimate of R_0 using the Lotka-Euler equation. Application of the period generation-time distribution before

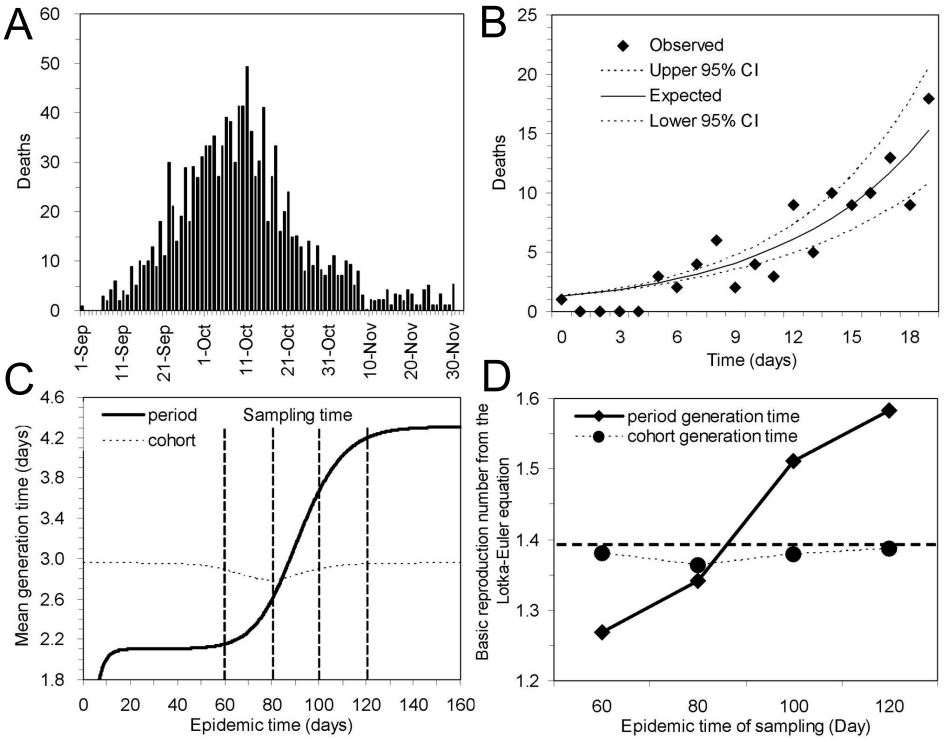

FIGURE 5. **Analysis of the Asian influenza epidemic in the Netherlands, 1957.** (A) The reported number of deaths as a function of calendar time. In total, 1230 deaths were reported from 1 September to 30 November [28]. (B) Observed and expected values of the daily number of deaths during the first 18 days. Assuming Poisson-distributed errors, the intrinsic growth rate is estimated as 0.131 (95% confidence interval: 0.113, 0.147) per day. (C) Mean period and cohort generation times, based on an SIR model, are illustrated as a function of calendar time t. The infectious period is assumed to follow an exponential distribution with a mean of $1/\gamma = 3$ days. Using the intrinsic growth rate $r_0 = 0.131$ per day, the unbiased $R_0 = 1 + r_0/\gamma = 1.394$. We assume that the generation times are sampled at a point in time, on Days 60, 80, 100 or 120. (D) Estimates of R_0, if the generation time is sampled at a point in time during the course of an epidemic, and if the obtained generation-time distribution at calendar time t is mistakenly applied to the Lotka-Euler equation as if the observation were the basic generation time. The dashed horizontal line is the reference line of the unbiased estimate (i.e., $R_0 = 1.394$).

observing the peak incidence (i.e., Days 60 and 80) leads to underestimation of R_0, whereas the use of the distribution after the peak (i.e., Days 100 and 120) results in the overestimation. On the other hand, if the cohort generation-time distribution

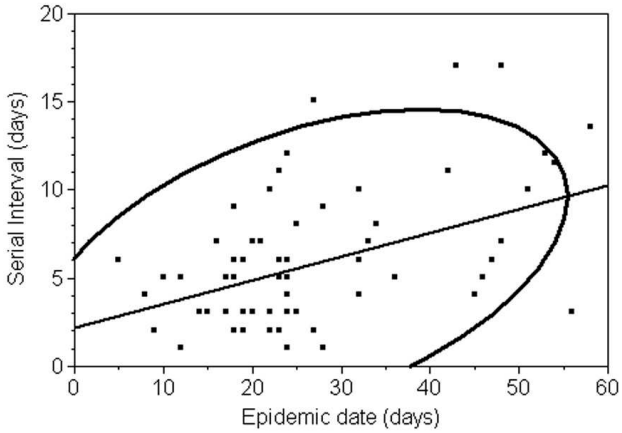

FIGURE 6. **Serial intervals of primary pneumonic plague in Manchuria, 1911, as a function of calendar time ($n = 88$).** Serial intervals are plotted according to the calendar time of onset of secondary cases. Among a total of 228 cases that were reported in the politically directly-controlled area of the Japanese Empire, 88 serial intervals were identified by means of contact tracing. The straight line represents the linear predictor. The density ellipsoid was computed from the bivariate normal distribution fit to the serial interval and calendar time. The dataset was extracted from historical publications [31, 41].

is mistakenly regarded as the basic generation-time distribution, R_0 will always be underestimated. Although the extent of this underestimation is small in Fig. 5D, this would be more obvious if the incidence is greater (or if the unbiased R_0 is greater than 1.39; see Fig. 4B).

As another example, we sought to validate our findings empirically. Fig. 6 shows the serial interval estimates of individual pairs of cases of primary pneumonic plague in Manchuria, China in 1910-11 [31, 41]. In total, 88 serial intervals were observed based on contact tracing among a total of 228 cases developing the disease in an area under direct political control of the Japanese Empire. In the original study, the date of illness onset for each confirmed case (i.e., secondary case) was recorded, and the contact tracing practice (e.g., interviewing any potential contacts within 1 week preceding the onset of illness through household visits) has indicated some sources of infection in a backward fashion based on household sharing or clearly identified contact in the community [31]. Originally, the contact tracing practice was conducted to estimate the incubation period, because classically the serial intervals were mistakenly regarded as equivalent to the incubation periods [30]. Subsequently, the serial intervals were plotted as a function of the onset time of secondary cases (Fig. 6), which reflects the period generation time if we apply equation (30). The overall mean serial interval was estimated as 5.66 days (SD = 3.65). The observed serial interval increased with calendar time t. A linear predictor suggests that the mean serial interval at calendar time t is $0.135t + 2.185$

days (coefficient of determination $R^2 = 0.20$). A density ellipsoid (showing $p = 0.95$) suggests a positive correlation between serial interval and epidemic date (correlation $= 0.450$, $p < 0.01$). Therefore, assuming that the historical data correctly recorded serial intervals, and assuming further that the known samples of serial interval represent the population estimate, the mean estimate of the serial interval as a function of onset time of secondary cases is certainly increased with calendar time. This supports our finding in the mean period generation time which is increased with calendar time t.

Similarly, it should be noted that the serial interval as a function of the onset time of the index case, which roughly captures the cohort generation time, was previously studied in an epidemic of severe acute respiratory syndrome (SARS) in Singapore (i.e., Fig. 1F in Lipsitch et al. [25]), showing a comparable qualitative pattern to our Fig. 4B. The above mentioned two different examples (i.e., plague and SARS) highlight the importance of emphasizing sampling method and its timing in estimating the generation time.

8. **Discussion.** The present study examined two different generation times which were time-inhomogeneous. Unlike the incubation period (i.e., the time from infection to symptom onset), the generation time is concerned with transmission process (i.e., is an interval from infection of the primary case to infection of the secondary cases generated by the primary case), and thus, the interval is not independent of calendar time and is greatly influenced by underlying transmission dynamics. The period generation time takes the calendar time t as the infection time of secondary cases. As the incidence varies prior to calendar time t, and as the growth rate of an epidemic accordingly changes (declines) with calendar time, the period generation time is increased with calendar time. On the other hand, the cohort generation time reflects the frequency of secondary transmissions caused by an infection cohort having experienced infection at calendar time t, which was shown to be decreased as the incidence increased. The reduction of the cohort generation time τ is caused by greater difficulty in finding a susceptible individual at calendar time $t + \tau$, among the infection cohort, born at calendar time t. The practical consequence of the sampling time of the generation times during the course of an epidemic was illustrated using the historical data of Asian influenza, where both period and cohort generation times would yield a biased estimate of R_0. In addition, our finding of the period generation time was supported by empirical evidence of the serial interval of pneumonic plague as a function of onset time of secondary cases. The increase in the period generation time was also noted in a stochastic simulation [4], and the contraction of the cohort generation time was indicated in an empirical observation of SARS [25] and survival analysis [21]. The present study contributed to an explicit definition of these two different generation times and comprehensively discussed their properties, adopting an important assumption of homogeneous mixing.

Two important public health implications are drawn from our exercise, which are relevant to the estimation of the basic and effective reproduction numbers from real-time epidemic growth data. First, the generation time is shown to vary greatly with calendar time which has to be taken into account to improve the sampling scheme of generation time in epidemiologic observations. The distribution of the serial interval can be extracted from contact tracing data, but, except for transmission experiments, it is very rare that contact tracing takes place throughout the duration of an epidemic. Rather, such data are extracted from the middle or late

stage of the epidemic. If the period generation time is sampled in the late stage of an epidemic (e.g., after observing the peak incidence), its application to the Lotka-Euler equation (1) would result in overestimation of R_0. Similarly, employing the cohort generation time would always underestimate R_0. Thus, to appropriately estimate R_0 using the Lotka-Euler equation, we have to **(i)** estimate the basic generation time (as the cohort generation time) during the initial growth phase or **(ii)** make an adjustment of the growth rate of an epidemic, which varies with calendar time t, for the period generation time. The latter adjustment of the period to the basic generation times might be achieved by using equation (18) (e.g., linear approximation of an epidemic curve during the period of sampling and an adjustment of period generation time, using the approximate growth rate, might allow estimation of the basic generation time). Another possible strategy is to estimate the generation-time distribution without contact tracing data; if R_0 is estimated from the final size, the Lotka-Euler equation (1) with known intrinsic growth rate can yield $g_0(\sigma)$.

Second, our exercise indicates that the statistical framework of real-time estimation of $R(t)$ or $R_c(t)$ has to be carefully re-assessed, if it is assumed that there is a known (basic) generation time [32]. For instance, if the distribution of generation-time is employed to reconstruct the transmission network as performed in SARS [46], the estimation of the $R_c(t)$ has to account for time-variations in $T_c(t)$ in the estimator which was successfully addressed using the hazard-based approach in a recent study [21]. That is, if the estimator of $R(t)$ or $R_c(t)$ (equations (8) and (9)) is based on appropriate understanding of the explicitly structured system (which captures the reality), the effect of time-variation will be readily included in the estimator and, thus, the time-inhomogeneous generation time may not be an issue. However, if the distribution of the basic generation-time is assumed to be known (and independent of time), and if the estimation procedure is developed without reference to the explicitly structured model, the present issue would be of critical importance. In addition, as an important assumption in the present study, it should be emphasized that we normalized the rate of infection $A(t+\tau,\tau)$ or $A(t,\tau)j(t-\tau)$ to yield the probability density function of the time-inhomogeneous generation time. The normalization may yield the density function, as long as the infection process has a Markov property. If not, the dependency of the generation time on R_0 has to be explicitly modeled [48], and the correct density function for the time-to-infection calls for a truncated distribution. The resulting qualitative patterns with respect to the time-inhomogeneity using a conditional generation-time distribution (given R_0) may not be too different from what were discussed in the present study, but the need to account for dependency between R_0 and the generation time distribution indicates critical importance to reconsider the relevant estimation frameworks for R_0, $R(t)$ and $R_c(t)$. That is, naive statistical inference of the basic generation-time distribution may well be unsuccessful based only on a sample distribution of the serial intervals, as indicated by both the present study on the time-inhomogeneity and Yan's study [48] indicating that the generation-time distribution is not independently and identically distributed.

A realistic issue that we did not address in this paper is concerned with the extrinsic dynamics (e.g., public health interventions). Since our illustrations and estimations were solely based on an assumption of Kermack and McKendrick (equation (6)), our exercise of the time-inhomogeneity highlighted the depletion of susceptible individuals alone (i.e., only the intrinsic dynamics). If the effective contact tracing,

quarantine and isolation are implemented during the course of an epidemic, the equation (6) may better read

$$A(t,\tau) = S(t)\beta(t,\tau)\Gamma(t,\tau) \tag{31}$$

where the reduction in contact frequency with calendar time t, $\beta(t,\tau)$, and early removal of infectious individuals at calendar time t, $\Gamma(t,\tau)$, may be caused by these interventions. Earlier secondary transmissions of a primary case as a result of very stringent contact tracing and isolation may well transiently shorten the period generation time at calendar time t. Similarly, the cohort generation time may become much shorter than those we showed in Fig. 4B (especially during the late stage of epidemic), which is intuitively seen from SARS data at the very end of an outbreak (i.e., cohort serial intervals in April in Fig. 1F in Lipsitch et al. [25]). The impact of the extrinsic dynamics on the generation-time distribution could be partly addressed as more detailed data become available (e.g., for this purpose, we have to analyze not only the incidence but also the number of susceptible individuals with calendar time).

Since the clarification of the generation time of infectious diseases is intended to improve the quantification of the transmission potential and objective interpretation of the time course of an epidemic, we should accentuate future implications for improving the relevant estimation frameworks. To appropriately understand the secondary transmission phenomena, the generation time has to account for various heterogeneous patterns in transmission. To allow analytical interpretation, all the arguments in the present study relied on a homogeneously mixing assumption. Although we focused only on the time-inhomogeneity of the generation time in a homogeneously mixing population, the generation time which addresses the heterogeneity with respect to age, space and social structure would be useful to offer more appropriate estimates of R_0, $R(t)$ and $R_c(t)$ using similar (but type-structured) real-time growth data. In particular, the generation time in a multi-layered population (e.g., separating household transmissions from those in the community) has yet to be formally defined. Also, an age-effect on the length of the serial interval of tuberculosis was previously discussed as a practical matter for making statistical inference [43] (i.e., tuberculosis infection among children leads to longer serial intervals than those in adulthood). Although the theory to support the relevant estimation framework has progressed greatly, we should note that the sampling scheme of the generation time (e.g., using equation (30)) has not yet been fully developed. Thus, sampling methods as well as estimation procedures, which can account for not only time-inhomogeneity, but also heterogeneous mixing, are warranted. Despite these proposed tasks, our study offers some technical measures for sampling the generation times, and should enhance discussion on the estimation of the transmission potential from real-time growth data.

During the final stages of revision, it came to our attention that a statistical study on generation time with similar scopes, using a stochastic modeling approach, has been published online [39].

REFERENCES

[1] J. M. Alho and B. D. Spencer, "Statistical Demography and Forecasting," Springer, New York, 2005.

[2] R. M. Anderson and R. M. May, "Infectious Diseases of Humans: Dynamics and Control," Oxford University Press, Oxford, 1991.

[3] N. T. J. Bailey, "The Mathematical Theory of Infectious Diseases and its Applications," 2^{nd} edition, Charles Griffin and Company Ltd., London, 1975.
[4] T. Burr and G. Chowell, *The reproduction number $R(t)$ in structured and nonstructured populations*, Math. Biosci. Eng., **6** (2009), 239–259.
[5] S. Cauchemez, P. Y. Boelle, C. A. Donnelly, N. M. Ferguson, G. Thomas, G. M. Leung, A. J. Hedley, R. M. Anderson and A. J. Valleron, *Real-time estimates in early detection of SARS*, Emerg. Infect. Dis., **12** (2006), 110–113.
[6] S. Cauchemez, P. Y. Boelle, G. Thomas and A. J. Valleron, *Estimating in real time the efficacy of measures to control emerging communicable diseases*, Am. J. Epidemiol., **164** (2006), 591–597.
[7] G. Chowell, H. Nishiura and L. M. Bettencourt, *Comparative estimation of the reproduction number for pandemic influenza from daily case notification data*, J. R. Soc. Interface., **4** (2007), 155–166.
[8] O. Diekmann and J. A. P. Heesterbeek, "Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation," Wiley, New York, 2000.
[9] O. Diekmann, J. A. P. Heesterbeek and J. A. Metz, *On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations*, J. Math. Biol., **28** (1990), 365–382.
[10] K. Dietz, *The estimation of the basic reproduction number for infectious diseases*, Stat. Methods. Med. Res., **2** (1993), 23–41.
[11] L. I. Dublin and A. J. Lotka, *On the true rate of natural increase*, J. Am. Stat. Assoc., **151** (1925), 305–339.
[12] P. E. M. Fine, *The interval between successive cases of an infectious disease*, Am. J. Epidemiol., **158** (2003), 755–760.
[13] C. Fraser, *Estimating individual and household reproduction numbers in an emerging epidemic*, PLoS One, **2** (2007), e758.
[14] C. Fraser, S. Riley, R. M. Anderson and N. M. Ferguson, *Factors that make an infectious disease outbreak controllable*, Proc. Natl. Acad. Sci. U.S.A., **101** (2004), 6146–6151.
[15] T. Garske, P. Clarke and A. C. Ghani, *The transmissibility of highly pathogenic avian influenza in commercial poultry in industrialised countries*, PLoS One, **2** (2007), e349.
[16] N. C. Grassly and C. Fraser, *Mathematical models of infectious disease transmission*, Nat. Rev. Microbiol., **6** (2008), 477–487.
[17] D. T. Haydon, M. Chase-Topping, D. J. Shaw, L. Matthews, J. K. Friar, J. Wilesmith and M. E. Woolhouse, *The construction and analysis of epidemic trees with reference to the 2001 UK foot-and-mouth outbreak*, Proc. R. Soc. Lond. Ser. B., **270** (2003), 121–127.
[18] R. E. Hope Simpson, *The period of transmission in certain epidemic diseases: An observational method for its discovery*, Lancet, **2** (1948), 755–760.
[19] H. Inaba and H. Nishiura, *The state-reproduction number for a multistate class age structured epidemic system and its application to the asymptomatic transmission model*, Math. Biosci., **216** (2008), 77–89.
[20] J. D. Kalbfleisch and J. F. Lawless, *Inference based on retrospective ascertainment: An analysis of the data on transfusion-related AIDS*, J. Am. Stat. Assoc., **84** (1989), 360–372.
[21] E. Kenah, M. Lipsitch and J. M. Robins, *Generation interval contraction and epidemic data analysis*, Math. Biosci., **213** (2008), 71–79.
[22] W. O. Kermack and A. G. McKendrick, *Contributions to the mathematical theory of epidemics. I.*, Proc. R. Soc. Ser. A., **115** (1927), 700–721, reprinted in Bull. Math. Biol. **53** (1991), 33–55.
[23] N. Keyfitz, "Applied Mathematical Demography," John Wiley and Sons, New York, 1977.
[24] S. W. Lakagos, L. M. Barraj and V. De Gruttola, *Nonparametric analysis of truncated survival data, with application to AIDS*, Biometrika, **75** (1988), 515–523.
[25] M. Lipsitch, T. Cohen, B. Cooper, J. M. Robins, S. Ma, L. James, G. Gopalakrishna, S. K. Chew, C. C. Tan, M. H. Samore, D. Fisman and M. Murray, *Transmission dynamics and control of severe acute respiratory syndrome*, Science, **38** (2003), 115–126.
[26] N. Masurel and W. M. Marine, *Recycling of Asian and Hong Kong influenza A virus hemagglutinins in man*, Am. J. Epidemiol., **97** (1973), 44–49.
[27] C. E. Mills, J. M. Robins and M. Lipsitch, *Transmissibility of 1918 pandemic influenza*, Nature, **432** (2004), 904–906.
[28] J. Mulder, N. Masurel and J. F. P. Hers, *De Aziatische-influenza-pandemie van 1957* (Dutch), Ned. Tijdschr. Geneeskd., **102** (1958), 1992–1999.

[29] H. Nishiura, *Correcting the actual reproduction number: A simple method to estimate R_0 from early epidemic growth data*, Int. J. Environ. Res. Public. Health., **7** (2010), 291–302.
[30] H. Nishiura, *Early efforts in modeling the incubation period of infectious diseases with an acute course of illness*, Emerg. Themes. Epidemiol., **4** (2007), 2.
[31] H. Nishiura, *Epidemiology of a primary pneumonic plague in Kantoshu, Manchuria, from 1910 to 1911: Statistical analysis of individual records collected by the Japanese Empire*, Int. J. Epidemiol., **35** (2006), 1059–1065.
[32] H. Nishiura, *Time variations in the transmissibility of pandemic influenza in Prussia, Germany, from 1918-19*, Theor. Biol. Med. Model., **4** (2007), 20.
[33] H. Nishiura and G. Chowell, *The effective reproduction number as a prelude to statistical estimtion of time-dependent epidemic trends*, in "Mathematical and Statistical Estimation Approaches in Epidemiology" (eds. G. Chowell, J. M. Hyman, L. M. A. Bettencourt and C. Castillo-Chavez), Springer, New York, (2009), 103–121.
[34] H. Nishiura, G. Chowell, H. Heesterbeek and J. Wallinga, *The ideal reporting interval for an epidemic to objectively interpret the epidemiological time course*, J. R. Soc. Interface., **7** (2010), 297–307.
[35] H. Nishiura, G. Chowell, M. Safan and C. Castillo-Chavez, *Pros and cons of estimating the reproduction number from early epidemic growth rate of influenza A (H1N1) 2009*, Theor. Biol. Med. Model., **7** (2010), 1.
[36] H. Nishiura, M. Schwehm, M. Kakehashi and M. Eichner, *Transmission potential of primary pneumonic plague: Time inhomogeneous evaluation based on historical documents of the transmission network*, J. Epidemiol. Community. Health., **60** (2006), 640–645.
[37] W. Pickles, "Epidemiology in Country Practice," John Wright and Sons, Bristol, 1939.
[38] M. G. Roberts and J. A. Heesterbeek, *Model-consistent estimation of the basic reproduction number from the incidence of an emerging infection*, J. Math. Biol., **55** (2007), 803–816.
[39] G. Scalia Tomba, A. Svensson, T. Asikainen and J. Giesecke, *Some model based considerations on observing generation times for communicable diseases*, Math. Biosci., **223** (2010), 24–31.
[40] A. Svensson, *A note on generation times in epidemic models*, Math. Biosci., **208** (2007), 300–311.
[41] Temporary Quarantine Section, K. Totokufu, "Epidemic Record of Plague During 1910-1911" (Meiji 43,4-nen 'Pest' Ryuko-shi), Manchurian Daily Press, Dalian, 1912.
[42] E. Vynnycky and W. J. Edmunds, *Analyses of the 1957 (Asian) influenza pandemic in the United Kingdom and the impact of school closures*, Epidemiol. Infect., **136** (2008), 166–179.
[43] E. Vynnycky and P. E. Fine, *Lifetime risks, incubation period, and serial interval of tuberculosis*, Am. J. Epidemiol., **152** (2000), 247–263.
[44] E. Vynnycky, A. Trindall and P. Mangtani, *Estimates of the reproduction numbers of Spanish influenza using morbidity data*, Int. J. Epidemiol., **36** (2007), 881–889.
[45] J. Wallinga and M. Lipsitch, *How generation intervals shape the relationship between growth rates and reproductive numbers*, Proc. R. Soc. Lond. Ser. B., **274** (2007), 599–604.
[46] J. Wallinga and P. Teunis, *Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures*, Am. J. Epidemiol., **160** (2004), 509–516.
[47] L. F. White and M. Pagano, *A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic*, Stat. Med., **27** (2007), 2999–3016.
[48] P. Yan, *Separate roles of the latent and infectious periods in shaping the relation between the basic reproduction number and the intrinsic growth rate of infectious disease outbreaks*, J. Theor. Biol., **251** (2008), 238–252.

Received November 3, 2009; Accepted June 7, 2010.

E-mail address: h.nishiura@uu.nl