Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Review

Quantifying the transmission potential of pandemic influenza

Gerardo Chowell a,∗, Hiroshi Nishiura b

a School of Human Evolution and Social Change, Arizona State University, Tempe, AZ 85282, USA
b Theoretical Epidemiology, University of Utrecht, Yalelaan 7, 3584 CL Utrecht, The Netherlands

Received 15 November 2007; received in revised form 16 December 2007; accepted 17 December 2007

Available online 10 January 2008

Communicated by J. Fontanari

Abstract

This article reviews quantitative methods to estimate the basic reproduction number of pandemic influenza, a key threshold quantity to help determine the intensity of interventions required to control the disease. Although it is difficult to assess the transmission potential of a probable future pandemic, historical epidemiologic data is readily available from previous pandemics, and as a reference quantity for future pandemic planning, mathematical and statistical analyses of historical data are crucial. In particular, because many historical records tend to document only the temporal distribution of cases or deaths (i.e. epidemic curve), our review focuses on methods to maximize the utility of time-evolution data and to clarify the detailed mechanisms of the spread of influenza.

First, we highlight structured epidemic models and their parameter estimation method which can quantify the detailed disease dynamics including those we cannot observe directly. Duration-structured epidemic systems are subsequently presented, offering firm understanding of the definition of the basic and effective reproduction numbers. When the initial growth phase of an epidemic is investigated, the distribution of the generation time is key statistical information to appropriately estimate the transmission potential using the intrinsic growth rate. Applications of stochastic processes are also highlighted to estimate the transmission potential using similar data. Critically important characteristics of influenza data are subsequently summarized, followed by our conclusions to suggest potential future methodological improvements.

© 2008 Elsevier B.V. All rights reserved.

PACS: 87.19.xd; 87.23.Cc; 87.10.Mn

Keywords: Influenza; Pandemic; Epidemiology; Basic reproduction number; Model

Contents

1. Introduction ................................................................. 51
2. On the definition of the transmission potential ........................................ 52
3. Estimating $R_0$ using a structured epidemic model ................................ 53
   3.1. Parameter estimation .................................................. 56
   3.2. Bootstrap confidence intervals ..................................... 57
4. Primer of mathematics and statistics of $R_0$ and $R(t)$ ............................. 57

∗ Corresponding author. Fax: +505 480 965 7671.
E-mail address: gchowell@asu.edu (G. Chowell).

1571-0645/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.plrev.2007.12.001
1. Introduction

Influenza epidemics are observed around the world during the wintertime and with a strong seasonal component in temperate regions [1,2]. Influenza is a disease caused by the influenza virus, an RNA virus belonging to the Orthomyxoviridae [3]. Many features are common with those of the paramyxovirus infections of the acute upper respiratory tract. Typical symptoms of the disease are characterized by fever, myalgia, severe malaise, non-productive cough, and sore throats. The disease spreads when an infected individual coughs or sneezes and sends the virus into the air, and other susceptible individuals inhale the virus. The virus is also believed to be transmitted when a person touches a surface that is contaminated with the virus (e.g., door knob, etc.) and then touches the nose or eyes. Infected individuals can transmit the virus almost within a day following infection (i.e., latent period). Although it is generally believed that infected individuals can pass the virus for 3–7 days following symptom onset, there is some uncertainty on the duration of the infectious period. The generation time (i.e., sum of latent and infectious periods) for influenza, reported and assumed in the literature, ranges from 3 days [4–6] to about 6 days [7,8].

Individuals that are infected with influenza are believed to become permanently immune against the specific virus strain. Hence, the virus is able to persist in the human population through relatively minor (single point) mutations in the virus composition known as drifts. Influenza (sub)types A/H3N2, A/H1N1 and B are currently co-circulating in the human population [9]. Major changes in the virus composition via recombination or gene reassortment processes (known as genetic shifts) can lead to the emergence of novel influenza viruses with the potential of generating dramatic morbidity and mortality levels around the world [3].

The 1918/19 influenza pandemic known as the Spanish influenza caused by the influenza virus A (H1N1) has been the most devastating in recent history with estimated worldwide mortality ranging from 20 to 100 million deaths [10,11] with a case fatality of 2–6 percent [12,13]. The worldwide 1918 influenza pandemic spread in three waves starting from Midwestern United States in the spring of 1918 [14,15]. The deadly second wave began in late August probably in France while the third wave is generally considered as part of normal more scattered winter outbreaks similar to those observed after the 1889/90 pandemic [14]. Subsequent pandemics during the 20th century are attributed to subtypes A (H2N2) from 1957/58 (Asian influenza) and A (H3N2) in 1968 (Hong Kong influenza) [16].

The ability to quickly detect and institute control efforts at the early stage of an influenza pandemic is directly linked to the final levels of morbidity and mortality in the population [13]. To appropriately assess the disaster size of a probable future pandemic, we have to quantify the transmission potential (and its associated uncertainty). Although it is difficult to directly measure the transmissibility of a future pandemic, historical epidemiologic data is readily available from previous pandemics, and as a reference quantity for future pandemic planning, mathematical and statistical analyses of the historical data can offer various insights. In particular, because many historical records tend to document only the temporal distribution of cases or deaths (i.e., epidemic curve), we modelers have faced with a difficult need to clarify the mechanisms of the spread of influenza using such time-evolution data alone. In this paper, we review a number of mathematical and statistical methods for the estimation of the transmission potential of pandemic influenza, focusing on theoretical techniques to maximize the utility of the temporal distribution of
influenza cases. The methods that have been incorporated in this review include the applications of epidemiologically structured epidemic models, explicitly duration-structured epidemic system, and stochastic processes (i.e. branching and counting processes). Whereas this review does not cover the spread of influenza in space, spatial heterogeneity in transmission and the growing interest in the role of contact networks are briefly discussed as the future challenge.

2. On the definition of the transmission potential

The basic reproduction number, $R_0$ (pronounced as R nought), is a key quantity used to estimate transmissibility of infectious diseases. Theoretically, $R_0$ is defined as the average number of secondary cases generated by a single primary case during its entire period of infectiousness in a completely susceptible population [17]. As the epidemic progresses, the number of susceptible individuals is decreased due to infection, and the reproduction number decays following $R(t) = R_0 S(t)/S(0)$ where $S(t)$ and $S(0)$ are, respectively, the number of susceptible individuals at time $t$ and before the epidemic starts; the latter is equivalent to the total population size $N$ given that all individuals are susceptible before the beginning of an epidemic. Clearly, this definition only applies to (novel) emerging infectious diseases (e.g., the epidemic of severe acute respiratory syndrome (SARS) from 2002–2003) or re-emerging infectious diseases that had not circulated in the population in question for long enough to allow for residual immunity in the population.

The reproduction number is directly related to the type and intensity of interventions necessary to control an epidemic since the objective is to make $R(t) < 1$ as soon as possible. To achieve $R(t) < 1$, one or a combination of control strategies may be implemented. For example, one of the best known uses of $R_0$ is in determining the critical coverage of immunization required to eradicate a disease in a randomly mixing population. That is, when vaccine is available against a disease in question, it is of interest to estimate the critical proportion of the population that needs to be vaccinated (i.e. vaccination coverage) in order to attain $R < 1$ [18,19]. For example, in the U.S. prior to 1963, a vaccine against measles was not available and hence recurrent epidemics of measles were observed with approximately 3–4 million cases per year and a mean of 450 deaths. The introduction of the vaccine in the U.S. reduced the incidence by 98 percent.

The critical vaccination coverage, $p_c$ (in a randomly mixing population) can be estimated from the $R_0$ of the disease in question as follows [20]:

$$p_c > \frac{1}{\epsilon} \left(1 - \frac{1}{R_0}\right),$$  \hspace{1cm} (1)

where $\epsilon$ is the efficacy (i.e. direct effectiveness) of vaccination [21]. $p_c$ given in (1) suggests that the disease could be eradicated even when all susceptible individuals are not vaccinated. The protection conferred to the population by achieving a critical vaccination coverage is known as herd immunity [22,23].

A brief history of the theoretical developments on the basic reproduction number and its analytical computation via epidemic modeling is given elsewhere [23–25]. The mathematical definition and calculation of $R_0$ using next-generation arguments was described initially by Odo Diekmann and his colleagues [17,26], where $R_0$ is the dominant eigenvalue of the resulting next generation matrix. Further elaborations and reviews can be found elsewhere [27–33]. Classically, rather than the threshold phenomena, $R_0$ was used to suggest the severity of an epidemic, because the proportion of those experiencing infection at the end of an epidemic depends only on $R_0$ [34] (see Section 3).

Statistical methods to quantitatively estimate $R_0$ have been reviewed by Klaus Dietz [35]. Depending on the characteristics of data and underlying assumptions of the models, $R_0$ can be estimated using various different approaches [36]. In addition to the final size equation, $R_0$ of an epidemic of newly emerging disease can be estimated from the intrinsic growth rate [4,6,8,18,37,38], which is also referred to as the rate of natural increase, suggesting the natural growth rate of infected individuals in a fully susceptible population (discussed in Section 4). Moreover, for simple epidemic models with relatively few parameters, $R_0$ can be estimated with other unobservable quantities by rigorous curve fitting of model equations to the observed epidemic data (discussed in Section 3) [38–40]. Not only $R_0$ but also $R(t)$ can be estimated from the temporal distribution of infectious diseases, reconstructing the transmission network or inferring the time-inhomogeneous number of secondary transmissions [41–44].

To estimate the basic reproduction number of endemic diseases, different approaches are taken. One would need first to carry out serological surveys to quantify the fraction of the population that is effectively protected against infection (i.e. age- and/or time-specific proportion of those possessing acquired immunity needs to be estimated).
Table 1
Reported estimates of $R_0$ for pandemic influenza during the fall wave (2nd wave) from 1918/19

| Location                     | Serial interval (days) | $R_0$  | Autonomous system fitted with entire epidemic curve | Reference |
|------------------------------|------------------------|--------|---------------------------------------------------|-----------|
| San Francisco, USA           | 6                      | 3.5    | Yes                                               | [39]      |
|                              | 6                      | 2.4    | No                                                | [39]      |
| 45 cities in the USA         | 6$^a$                  | 2.7    | No                                                | [8]       |
| UK (entire England and Wales)$^b$ | 6                      | 1.6    | Yes                                               | [58]      |
| Scandinavian cities          | 6                      | 1.4–1.6| No                                                | [62]      |
| Geneva, Switzerland          | 5.7                    | 3.8    | Yes                                               | [38]      |
| Sao Paulo, Brazil            | 4.6                    | 2.7    | Yes                                               | [59]      |
| Cities in Europe and America | 4.0                    | 1.2–3.0| No                                                | [63]      |
| 83 cities in the UK          | 3.2, 2.6               | 1.7–2.0| No                                                | [5,57]    |
| 45 cities in the USA         | 2.9                    | 1.7    | No                                                | [6]       |
| RAF camp in the UK           | 2.3                    | 2.9    | No                                                | [64]      |
| Featherston Military         | 1.6                    | 3.1    | Yes                                               | [61]      |
| Camp, New Zealand$^c$        | 1.1                    | 1.8    | Yes                                               | [61]      |
|                              | 0.9                    | 1.3    | Yes                                               | [61]      |

$^a$ Sensitivity of the $R$ estimates to different assumptions for the serial interval was examined.
$^b$ Three pandemic waves were simultaneously fitted.
$^c$ The epidemic was observed in a community with closed contact (i.e. military camp).

Through this effort, the force of infection, the rate at which susceptible individuals get infected, is estimated [45]. For example, this is the case for rubella, mumps and measles (that are still circulating in some regions of the world even when high effective vaccination coverage is achieved). Although the estimation of $R_0$ for endemic diseases is out of the scope of this review, methodological details and the applications to estimate the force of infection and $R_0$ can be found elsewhere [46–52].

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. $R$ 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 vaccination campaigns.

With regard to seasonal influenza, the reproduction number ($R$) over the last 3 decades has been estimated at 1.3 (SE 0.05) in the United States, France, and Australia with an overall range of 0.9–2.0 [53]. An $R$ estimate of 1.5 has been reported for a single A/H3N2 season in France [54], and some estimates have been reported in the range 1.4–2.6 for several consecutive influenza seasons in England and Wales [55,56]. A particularly high estimate of $R$ has been suggested for the 1951 influenza epidemic in England and Canada [57].

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 [8,38–40,57–64] 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. Table 1 lists estimates of $R_0$ in recent studies. The variability of $R_0$ estimates suggests that local factors, including geographic and demographic conditions, could play an important role in disease spread [65–67]. In the following sections, we review how these estimates are obtained and how we shall interpret the estimates, starting from a simple structured epidemic model proposed in 1927.

3. Estimating $R_0$ using a structured epidemic model

Mathematical models provide a unique way to analyze the transmission dynamics and study various different scenarios associated to the spread of communicable diseases in population(s) [18,68,69]. The history of the mathe-
Mathematical modeling of infectious diseases greatly remounts to the study of Sir Ronald Ross in 1911 [70] who invented a classic malaria model and also discovered the mosquito-borne transmission mechanisms of malaria. Employing a mass action principle for the spread of malaria, Ross explored the effects of controlling the mosquito population using simple mathematical models [71]. Following his effort, Kermack and McKendrick introduced a classical SIR (susceptible-infectious-removed) epidemic model in 1927, which is most frequently utilized in the present day, given by the following system of nonlinear ordinary differential equations (ODEs) [72]:

\[
\begin{align*}
\frac{dS(t)}{dt} &= -\frac{\beta S(t)I(t)}{N}, \\
\frac{dI(t)}{dt} &= \frac{\beta S(t)I(t)}{N} - \gamma I(t), \\
\frac{dR(t)}{dt} &= \gamma I(t),
\end{align*}
\]

where \( S(t) \) denotes susceptible individuals at time \( t \); \( I(t) \), infected (assumed infectious) individuals at time \( t \); and \( R(t) \), recovered (assumed permanently immune) individuals at time \( t \); \( \beta \) is the transmission rate; \( \gamma \) the recovery rate; and \( N \) is the total population size which is assumed constant for a closed population (i.e. a population without immigration and emigration). Susceptible individuals in contact with the virus enter the infectious class \( I \) at the rate \( \beta I/N \). That is, homogeneous mixing between individuals is assumed.

The basic reproduction number, \( R_0 \), for the epidemic system (2) is given by the product of the transmission rate and the mean infectious period. That is:

\[
R_0 = \frac{\beta}{\gamma}.
\]

Classically, \( R_0 \) has been known as a quantity to suggest severity of an epidemic [34]. Indeed, analytical expression of \( R_0 \) in (3) is derived simply by solving the above system (2). Replacing \( I(t) \) in the right-hand side of \( dS(t)/dt \) by \( (1/\gamma) dR(t)/dt \), we get

\[
\frac{1}{S(t)} \frac{dS(t)}{dt} = -\frac{\beta}{\gamma N} \frac{dR(t)}{dt}.
\]

Integrating both sizes of (4) from 0 to infinity,

\[
\ln \frac{S(\infty)}{S(0)} = -\frac{\beta}{\gamma N} (R(\infty) - R(0)).
\]

Since \( S(\infty) = N - R(\infty) \), and because we assume \( S(0) = N \) and \( R(0) = 0 \), Eq. (5) can be rewritten as

\[
\ln \frac{N - R(\infty)}{N} = -\frac{\beta}{\gamma} \frac{R(\infty)}{N}.
\]

In the above Eq. (6), final size, i.e. the proportion of those experiencing infection among a total number of individuals in a community following a large scale epidemic, is defined as \( p \equiv R(\infty)/N \). That is,

\[
\ln(1 - p) = -R_0 p.
\]

Therefore, the following final size equation of an autonomous SIR (or SEIR) model is obtained:

\[
\hat{R}_0 = -\frac{\ln(1 - p)}{p}.
\]

Eq. (8) can be analytically derived using both deterministic (models governed by ODEs or partial differential equations (PDEs)) [73] and stochastic models [74].

Despite the usefulness of (2), SIR assumptions given by ODEs are not always directly applicable to real data. One of the reasons include that there is no disease where an infected individual can cause secondary transmission immediately after his/her infection.

Accordingly, we have used slightly extended compartmental models in the previous studies to describe the transmission dynamics of the 1918/19 influenza pandemic and estimate the reproduction number [38,39]. We now describe
Fig. 1. Flow chart of the state progression of individuals among the different epidemiological classes as modeled by the complex SEIR model. See Eqs. (10).

two different SEIR (susceptible-exposed-infectious-removed) models that have been used to estimate the reproduction number. The first model is the simple SEIR model, and the second model accounts for asymptomatic and hospitalized individuals.

The simple SEIR model classifies individuals as susceptible ($S$), exposed ($E$), infectious ($I$), recovered ($R$), and dead ($D$) [18]. Susceptible individuals in contact with the virus enter the exposed class at the rate $\beta I(t)/N$, where $\beta$ 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 $\gamma$ and $\delta$, respectively. Recovered individuals are assumed protected for the duration of the outbreak. The mortality rate is given by $\delta = \gamma/\text{CFP}/(1-\text{CFP})$, where CFP is the mean case fatality proportion. The transmission process can be modeled using the system of nonlinear differential equations:

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

(9)

where $C(t)$ is the cumulative number of infectious individuals. The basic reproduction number of the above system (9) is given by the product of the mean transmission rate and the mean infectious period, $R_0 = \beta/(\gamma + \delta)$.

A more complex SEIR model (Fig. 1) classifies individuals as susceptible ($S$), exposed ($E$), clinically ill and infectious ($I$), asymptomatic and partially infectious ($A$), diagnosed and reported ($J$), recovered ($R$), and death ($D$). The birth and natural death rates are assumed to have a common rate $\mu$ (60-year life expectancy as in [38]). The entire population is assumed susceptible at the beginning of the pandemic wave. Susceptible individuals in contact with the virus progress to the latent class at the rate $\beta(I(t) + J(t) + qA(t))/N$ where $\beta$ is the transmission rate, and $0 < q < 1$ is a reduction factor in the transmissibility of the asymptomatic class ($A$). Since there is no explicit evidence
estimating and proving the effectiveness of public health interventions, and because a high burden was placed upon
the sanitary and medical sectors, diagnosed/hospitalized individuals (J) are assumed equally infectious. Although it
is difficult to explicitly evaluate the difference in infectiousness between general community and hospital, we roughly
made this assumption since 78 percent of the nurses of the San Francisco Hospital contracted influenza [75]. A more
rigorous assumption requires either statistical analysis of more detailed time-series data [76] or an epidemiological
comparison of specific groups by contact frequency [77].

The total population size at time t is given by

\[ N = S(t) + E(t) + I(t) + A(t) + J(t) + R(t). \]

We assumed homogeneous mixing of the population and, therefore, the fraction \((I(t) + J(t) + qA(t))/N\) is the probability of a random contact with an infectious individual. A proportion \(0 < \rho < 1\) of latent individuals progress to the clinically infectious class \((I)\) at the rate \(k\) while the remaining \((1 - \rho)\) progress to the asymptomatic partially infectious class \((A)\) at the same rate \(k\) (fixed to 1/1.9 days\(^{-1}\) [8]). Asymptomatic cases progress to the recovered class at the rate \(\gamma_1\). Clinically infectious individuals (class \(I\)) are diagnosed (reported) at the rate \(\alpha\) or recover without being diagnosed (e.g., mild infections, hospital refusals) at the rate \(\gamma_1\). Diagnosed individuals recover at the rate \(\gamma_2 = 1/(1/\gamma_1 - 1/\alpha)\) or die at rate \(\delta\). The mortality rates were adjusted according to the case fatality proportion (CFP), such that \(\delta = \frac{\text{CFP}}{1 - \text{CFP}} \left(\frac{\mu}{\gamma_2}\right)\).

The transmission process can be modeled using the following system of nonlinear differential equations:

\[
\begin{align*}
\frac{dS(t)}{dt} &= \mu N(t) - \frac{\beta S(t)(I(t) + J(t) + q A(t))}{N} - \mu S(t), \\
\frac{dE(t)}{dt} &= \frac{\beta S(t)(I(t) + J(t) + q A(t))}{N} - (k + \mu) E(t), \\
\frac{dA(t)}{dt} &= k(1 - \rho) E - (\gamma_1 + \mu) A(t), \\
\frac{dI(t)}{dt} &= k \rho E(t) - (\alpha + \gamma_1 + \mu) I(t), \\
\frac{dJ(t)}{dt} &= \alpha I(t) - (\gamma_2 + \delta + \mu) J(t), \\
\frac{dR(t)}{dt} &= \gamma_1 (A(t) + I(t)) + \gamma_2 J(t) - \mu R(t), \\
\frac{dD(t)}{dt} &= \delta J(t), \\
\frac{dC(t)}{dt} &= \alpha I(t).
\end{align*}
\]

We assume the cumulative number of influenza notifications, our observed epidemic data, is given by \(C(t)\). Seven
model parameters \((\beta, \gamma_1, \alpha, q, \rho, E(0), I(0))\) are estimated from the epidemic curve by least squares fitting as explained below. The reproduction number for model (10) is given by (see [38]):

\[
\mathcal{R} = \frac{\beta k}{k + \mu} \left\{ \rho \left( \frac{1}{\gamma_1 + \alpha + \mu} + \frac{\alpha}{(\gamma_1 + \alpha + \mu)(\gamma_2 + \delta + \mu)} \right) + (1 - \rho) \left( \frac{q}{\gamma_1 + \mu} \right) \right\}
\]

and the clinical reporting proportion is given by:

\[
O = \frac{\alpha}{\alpha + \gamma_1 + \mu}.
\]

3.1. 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 \(\hat{\Theta}\) 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 \(C(t, \Theta)\). That is,
we minimize:

\[
X(\Theta) = \sum_{i=1}^{n}(y_t - C(t, \Theta))^2.
\]
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 [78]:

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

which can be estimated by

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

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 [79], 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 (Eqs. (9)) is given by:

$$V(\hat{R}_0) \approx \hat{R}_0^2 \left\{ \frac{V(\hat{\beta})}{\hat{\beta}^2} + \frac{V(\hat{\gamma})}{(\hat{\gamma} + \hat{\delta})^2} + \frac{V(\hat{\delta})}{(\hat{\gamma} + \hat{\delta})^2} - \left( \frac{2}{\hat{\beta}(\hat{\gamma} + \hat{\delta})} \right) \left( \text{Cov}(\hat{\beta}, \hat{\gamma}) - \hat{\gamma} \text{Cov}(\hat{\delta}, \hat{\gamma}) + \text{Cov}(\hat{\delta}, \hat{\beta}) \right) \right\}.$$  \hspace{1cm} (16)

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

3.2. Bootstrap confidence intervals

Another method to generate uncertainty bounds on the reproduction number is generating bootstrap confidence intervals by generating realizations of the best-fit curve $C(t)$ [80]. 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. Fig. 2 shows the temporal distributions of the reproduction number and the proportion of the clinical reporting obtained by the bootstrap method after fitting the complex SEIR epidemic model to the initial phase of the Fall influenza wave using 17 epidemic days of the Spanish Flu Pandemic in San Francisco, California.

4. Primer of mathematics and statistics of $R_0$ and $R(t)$

In addition to the estimation of $R_0$, it is of practical importance to evaluate time-dependent variations in the transmission potential. Explanation of the time course of an epidemic can be partly achieved by estimating the effective reproduction number, $R(t)$, defined as the actual average number of secondary cases per primary case at time $t$ (for $t > 0$) [41,42,44,81,82]. Although effective interventions against Spanish influenza may have been limited in the early 20th century, it is plausible that the contact frequency leading to infection varied with time owing to the huge number of deaths and dissemination of information through local media (e.g., newspapers). $R(t)$ shows time-dependent variation with a decline in susceptible individuals (intrinsic factors) and with the implementation of control measures (extrinsic factors). If $R(t) < 1$, it suggests that the epidemic is in decline and may be regarded as being under control at time $t$ (vice versa, if $R(t) > 1$).

---

1 $R(t)$ should not be confused with the number of removed individuals using the same notation. In the following arguments of this paper, $R(t)$ denotes the effective reproduction number.
Fig. 2. Model fits, residuals plots, and the resulting distributions of the reproduction number and the proportion of the clinical reporting obtained after fitting the complex SEIR epidemic model to the initial phase of the Fall influenza wave using 17 epidemic days of the Spanish Flu Pandemic in San Francisco, California. See Eqs. (10) [39]. In the top panel, the epidemic data of the cumulative number of reported influenza cases are the circles, the solid line is the model best fit, and the solid blue lines are 1000 realizations of the model fit to the data obtained through parametric bootstrapping as explained in the main text.

4.1. Estimation of $R_0$ using the intrinsic growth rate

To appropriately understand the theoretical concept of $R(t)$, let us firstly consider an explicitly infection-age structured epidemic model. Whereas Kermack–McKendrick model governed by ODEs (i.e. SIR and SEIR models as discussed above) has been well-known, Kermack and McKendrick had actually proposed an infection-age structured model in their initial publication in 1927 [72], the mathematical importance of which was recognized only after 1970s [83,84]. Let us denote the numbers of susceptible and recovered individuals by $S(t)$ and $U(t)$. Further, let $i(t, \tau)$ be the density of infectious individuals at time $t$ and infection-age $\tau$ (i.e. time since infection). The model is given by

$$\frac{dS(t)}{dt} = -\lambda(t)S(t),$$

$$\left( \frac{\partial}{\partial t} + \frac{\partial}{\partial \tau} \right) i(t, \tau) = -\gamma(\tau)i(t, \tau), \quad i(t, 0) = \lambda(t)S(t),$$

$$\frac{dU(t)}{dt} = \int_{0}^{\infty} \gamma(\tau)i(t, \tau) d\tau,$$  \hspace{1cm} (17)

where $\lambda(t)$ is referred to as the force of infection (foi) (i.e. as discussed in Section 2, foi is defined as the rate at which susceptible individuals get infected) which is given by:

$$\lambda(t) = \int_{0}^{\infty} \beta(\tau)i(t, \tau) d\tau$$  \hspace{1cm} (18)
and \( \gamma(\tau) \) is the rate of recovery at infection-age \( \tau \). It should be noted that the above model has not taken into account the background host demography (i.e. birth and death). In a closed population, the total population size \( N \) is thus given by:

\[
N = S(t) + \int_0^\infty i(t, \tau) \, d\tau + U(t).
\]

(19)

It should also be noted that, although \( i(t, \tau) \) is referred to as *density*, it is not meant to be a normalized density (i.e. integral of \( i(t, \tau) \) over \( t \) and \( \tau \) does not sum up to 1). Rather, we use density to mathematically refer to the absolute frequency in the infection-age space.

The system (17) can be reasonably integrated

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

(20)

where

\[
j(t) = i(t, 0),
\]

\[
\Gamma(\tau) = \exp \left( - \int_0^\tau \gamma(\sigma) \, d\sigma \right)
\]

(21)

and \( j_0(\tau) \) suggests the density of initially infected individuals at the beginning of an epidemic. In the following arguments, we call \( j(t) \) as *incidence of infection* (i.e. new infections at a given point of time \( t \)). It is not difficult to derive

\[
S(t) = S(0) - \int_0^t j(\sigma) \, d\sigma
\]

(22)

from (17)–(21). Thus, the subequation of \( i(t, 0) \) in system (17) is rewritten as

\[
j(t) = \lambda(t) \left[ S(0) - \int_0^t j(\sigma) \, d\sigma \right].
\]

(23)

Taking into account the initial condition in (20), Eq. (23) is rewritten as

\[
j(t) = \left[ S(0) - \int_0^t j(\sigma) \, d\sigma \right] \left[ G(t) + \int_0^t \psi(\tau) j(t - \tau) \, d\tau \right],
\]

(24)

where

\[
\psi(\tau) = \beta(\tau) \Gamma(\tau),
\]

\[
G(t) = \int_0^\infty \beta(\sigma + t) \frac{\Gamma(\sigma + t)}{\Gamma(\sigma)} j_0(\sigma) \, d\sigma.
\]

(25)

Considering the initial invasion phase (i.e. initial growth phase of an epidemic), we get a linearized equation

\[
j(t) = S(0) G(t) + S(0) \int_0^t \psi(\tau) j(t - \tau) \, d\tau.
\]

(26)
Eq. (26) represents Lotka’s integral equation, where the basic reproduction number, $R_0$, is given by

$$R_0 = S(0) \int_0^\infty \psi(\tau) \, d\tau. \quad (27)$$

Thus, the epidemic will grow if $R_0 > 1$ and decline to extinction if $R_0 < 1$. The above model can yield the same final size equation as seen in models governed by ODEs [17].

Assuming that the infection-age distribution is stable, we get a simplified renewal equation

$$j(t) = \int_0^\infty A(\tau) j(t - \tau) \, d\tau, \quad (28)$$

where $A(\tau)$ is the product of $\psi(\tau)$ and $S(0)$. Moreover, assuming that we observe an exponential growth of incidence during the initial phase (i.e. $j(t) = k \exp(rt)$ where $k$ and $r$ are, respectively, a constant ($k > 0$) and the intrinsic growth rate), the following relationship must be met

$$j(t) = j(t - \tau) \exp(r\tau). \quad (29)$$

Replacing $j(t - \tau)$ in the right-hand side of (28) by (29), we get

$$j(t) = \int_0^\infty A(\tau) j(t) \exp(-r\tau) \, d\tau. \quad (30)$$

Removing $j(t)$ from both sides of (30), we get the Lotka–Euler characteristic equation:

$$1 = \int_0^\infty e^{-r\tau} A(\tau) \, d\tau. \quad (31)$$

Further, we consider a probability density of the generation time (i.e. the time from infection of an individual to the infection of a secondary case by that individual [85]), denoted by $w(\tau)$:

$$w(\tau) := A(\tau) \int_0^\infty A(x) \, dx = \frac{A(\tau)}{R_0}. \quad (32)$$

Using (32), Eq. (31) can be replaced by

$$\frac{1}{R_0} = \int_0^\infty \exp(-r\tau) w(\tau) \, d\tau. \quad (33)$$

Eqs. (29)–(33) are what Wallinga and Lipsitch discussed in a recent study [6], reasonably suggesting the relationship between the generation time and $R_0$. Accordingly, the estimator of $R_0$ using the intrinsic growth rate is given by:

$$\hat{R}_0 = \frac{1}{M(-r)}, \quad (34)$$

where $M(-r)$ is the moment generating function of the generation time distribution $w(\tau)$, given the intrinsic growth rate $r$ [6].² Eq. (34) significantly improved the issue of estimating $R_0$ using the intrinsic growth rate alone, because (34) permits validating estimates of $R_0$ by various different distributional assumptions for $w(\tau)$. The issue of assuming

² In the original study of Wallinga and Lipsitch [6], the notation $R_0$ is not used for Eq. (34) and rather document (34) as the estimator of $R$. Most likely, there are two reasons for this. First, we cannot assure if all individuals are susceptible to pandemic influenza before the start of epidemic (as discussed in Section 2). Second, we cannot assume that infection-age distribution is stable during the initial growth phase, which is highlighted in (20). Thus, it should be remembered that the above discussion is mathematically tight in theory, but there are certain number of assumptions to apply the concept to observed data. Since writing $R$ alone is always confusing (as it is unclear if $R$ is concerned with time or immunity status), here we use $R_0$ instead.
explicit distributions for latent and infectious periods has been highlighted in recent studies [86–90] and indeed, this point is in part addressed by (34), because the convolution of latent and infectious periods yields \( w(\tau) \). Moreover, since the assumed lengths of generation time most likely yielded different estimates of \( R_0 \) for Spanish influenza by different studies [60]. Eq. (34) highlights a critical need to clarify the generation time distribution using observed data.

Here we briefly show a numerical example. Fig. 3 shows the daily number of influenza deaths during Spanish influenza pandemic in a suburb of Zurich, 1918 [91]. Since the non-linear phase is difficult to analyze, our interest to estimate \( R_0 \) with this dataset is limited to the initial growth phase only (right panel in Fig. 3). Even though the data represent deaths over time (i.e. not infection events with time), we can directly extract the same intrinsic growth rate as practised with onset data, assuming that death data are a good proxy for morbidity data (see our discussions in Section 6). Assuming exponential growth in deaths as shown in (29), the intrinsic growth rate \( r \) is estimated to be 0.16 per day. Supposing that \( w(\tau) \) is arbitrarily assumed to follow a gamma distribution with mean \( G \) and coefficient of variation, \( CV = \sqrt{\text{Var}(G)}/G \), \( R_0 \) is given by

\[
R_0 = \left(1 + rG(CV)^2\right)^{1/(2CV^2)}.
\]

Although there is no consensus regarding the generation time of Spanish influenza, we assume it ranges from 2–5 days. Assuming further that \( CV = 0.5 \), \( R_0 \) is estimated to range from 1.36 (for \( G = 2 \text{ days} \)) to 2.07 (for \( G = 5 \text{ days} \)).

4.2. The concept of \( R(t) \) and its estimation

In the following, let us consider the non-linear phase of an epidemic. Derivation of \( R_0 \) given by (34) assumes an exponential growth which is applicable only during the very initial phase of an epidemic (or, when the transmission is stationary over time), and thus, it is of practical importance to widen the utility of above-described renewal equations in order to appropriately interpret the time-course of an influenza pandemic. Let us explicitly account for the depletion of susceptible individuals, as we deal with an estimation issue with time-inhomogeneous assumptions (i.e. non-linear phase). Adopting the mass action assumption, we get:

\[
j(t) = S(t) \int_0^\infty \psi(\tau) j(t - \tau) d\tau = \int_0^\infty A(t, \tau) j(t - \tau) d\tau,
\]

where \( A(t, \tau) \) should be interpreted as the reproductive power at time \( t \) and infection-age \( \tau \) at which an infected individual generates secondary cases. We refer to the latter part of Eq. (36) as a non-autonomous renewal equation, where the number of new infection at time \( t \) is proportional to the number of infectious individuals (as assumed in the renewal equation in the initial phase).
Using Eq. (36), the effective reproduction number, $R(t)$ (i.e. instantaneous reproduction number at calendar time $t$) is defined as:

$$R(t) = \int_0^\infty A(t, \tau) d\tau.$$  \hspace{1cm} (37)

Following (37), we can immediately see that $R(t)$ with an autonomous assumption (i.e. where contact and recovery rates do not vary with time) is given by:

$$R(t) = \frac{S(t)}{S(0)} R_0$$  \hspace{1cm} (38)

which is shown in [17]. In practical terms, Eq. (38) suggests that time-varying decrease in transmission potential as well as decline in the epidemic reflects only depletion of susceptible individuals. This corresponds to a classic assumption of the Kermack and McKendrick model.

However, as we discussed in the beginning of this section, we postulate that human contact behaviors (and other extrinsic factors) modify the dynamics of pandemic influenza, assuming that the decline in incidence does reflect not only depletion of susceptibles but also various extrinsic dynamics (e.g., isolation, quarantine and closure of public buildings). Thus, instead of the assumption in (36), we shall assume time-inhomogeneous $\psi(t, \tau)$; i.e.

$$j(t) = \int_0^\infty \psi(t, \tau) j(t - \tau) d\tau = \int_0^\infty A(t, \tau) j(t - \tau) d\tau$$  \hspace{1cm} (39)

to describe $A(t, \tau)$.

To derive simple estimator of $R(t)$, it is convenient to assume separation of variables for $A(t, \tau)$ (implicitly assuming that the relative infectiousness to infection-age is independent of calendar time) [92]. Under this assumption, $A(t, \tau)$ is rewritten as the product of two functions $\phi_1(t)$ and $\phi_2(\tau)$:

$$A(t, \tau) = \phi_1(t) \phi_2(\tau).$$  \hspace{1cm} (40)

Arbitrarily assuming a normalized density for $\phi_2(\tau)$, i.e.

$$\int_0^\infty \phi_2(\tau) d\tau \equiv 1$$  \hspace{1cm} (41)

then, it is easy to find that

$$R(t) = \int_0^\infty A(t, \tau) d\tau = \phi_1(t)$$  \hspace{1cm} (42)

suggesting that the function $\phi_1(t)$ is equivalent to the effective reproduction number $R(t)$. Another function $\phi_2(\tau)$ represents the density of infection events as a function of infection-age $\tau$. Accordingly, we can immediately see that $\phi_2(\tau)$ is exactly the same as $w(\tau)$, the generation time distribution. That is, the above arguments suggest that $A(t, \tau)$ (i.e. the rate at which an infectious individual at calendar time $t$ and infection-age $\tau$ produces secondary transmission) is decomposed as:

$$A(t, \tau) = R(t) w(\tau).$$  \hspace{1cm} (43)

Inserting (43) into (39) yields an estimator of $R(t)$ [92]:

$$\hat{R}(t) = \frac{j(t)}{\int_0^\infty j(t - \tau) w(\tau) d\tau}.$$  \hspace{1cm} (44)

The above Eq. (44) is exactly what was proposed in applications to SARS [41] and foot and mouth disease [93]; i.e. discretizing (44) to apply it to the daily incidence data (i.e. using $j_i$ incident cases infected between time $t_i$ and time
$t_{i+1}$ and discretized generation time distribution $w_i$),
\[
\hat{R}(t_i) = \frac{j_i}{\sum_{j=0}^{n} j_{i-j} w_j}
\]  
(45)

was used as the estimator. However, it should be noted that the study in SARS implicitly assumed that onset data $c(t)$ at time $t$ reflects the above discussed infection event $j(t)$. That is, supposing that we observed $c_i$ onset cases reported between $t_i$ and $t_{i+1}$, $R(t)$ was calculated as
\[
\hat{R}(t_i) = \frac{c_i}{\sum_{j=0}^{n} c_{i-j} s_j},
\]  
(46)

where $s_j$ is the discretized serial interval which is defined as the time from onset of a primary case to onset of the secondary cases [94,95]. The method permits reasonable transformation of an epidemic curve (i.e. temporal distribution of case onset) to the estimates of time-inhomogeneous reproduction number $R(t)$. Employing the relative likelihood of case $k$ infected by case $l$ using the density function of serial interval $s(t)$; i.e.
\[
p(k,l) = \frac{s(t_k - t_l | \theta)}{\sum_{m \neq k} s(t_k - t_m | \theta)}.
\]  
(47)

Using (47), expected value and variance of $R(t_i)$ are given by the following
\[
E(R(t_i)) = \frac{1}{n^2} \sum_{l:t_l=t} \sum_{k=1}^{n-q} P(k,l),
\]  

\[
\text{Var}(R(t_i)) = \frac{1}{n^2} \sum_{k=1}^{n-q} \left( \sum_{l:t_l=t} P(k,l)(1 - P(k,l)) - \sum_{l,m:t_l=t} P(k,l)P(k,m) \right),
\]  
(48)

where $n_i$ is the total number of reported case onsets at time $t$ [96].

In the present day, only by using the above described methods (or similar concepts with similar assumptions), we can transform epidemic curves into $R(t)$ and roughly assess the impact of control measures on an epidemic. However, whereas Eqs. (45) and (46) are similar in theory, we need to explicitly account for the difference between onset and infection event. In fact, when there are many asymptomatic infections and asymptomatic secondary transmissions, serial interval is not equivalent to the generation time, and thus, directly adopting the above methods would be inappropriate. Since this point is particularly important in analyzing influenza data, we discuss this issue in Section 6.

5. Statistical methods to estimate $R_0$

5.1. Branching process

In the previous sections, we discussed several different methods to estimate $R_0$ either by (i) employing detailed curve fitting methods assuming a structured epidemic model or (ii) using the intrinsic growth rate (or doubling time [97,98]). Summarizing the above discussions, we believe that the readers should benefit from memorizing $R_0 = 1/M(-r)$ for the use of the intrinsic growth rate $r$ in estimating $R_0$ [6] and remembering the final size equation $R_0 = -\ln(1-p)/p$, suggesting the severity of an epidemic as the theoretical concept. Indeed, estimators using either the intrinsic growth rate or final size relations have played an important role in discussing $R_0$ of pandemic influenza [63].

However, it should be noted that deterministic models do not permit incorporating stochasticity explicitly (e.g., standard error for $R_0$ is determined by measurement of errors alone), as the models argue only average number of secondary transmissions within the assumed transmission dynamics. That is, our arguments given above explore only the time-evolution of influenza spread in the mean field. To address the issue of variation in secondary transmissions, full stochastic models are called for [99].

From a viewpoint of data science, the discrete-time branching process, which is also referred to as Galton–Watson process, can reasonably assess individual heterogeneity in secondary transmissions [100,101]. As we discussed the initial growth phase of the deterministic model, let us consider the same epidemic phase where we observe a geometric
increase in the number of cases by generation [23]. We denote the initial number of infected individuals by $C_0$ in generation 0. Then, during the first generation, $C_1$ cases are produced by secondary transmissions of $C_0$. Similarly, let $C_n$ be the number of infections in generation $n$. The branching process of this type assumes that every infected individual has an independently and identically distributed stochastic random variable $\rho_i^{(n)}$ representing the number of secondary cases produced by case $i$ in generation $(n)$, and that environmental stochasticity and immigration/emigration can be ignored. Supposing that the pattern of secondary transmission follows a discrete probability distribution $p_k$ with $k$ secondary transmission(s); i.e.

$$p_k = \Pr(\rho_i^{(n)} = k) \quad (k = 0, 1, 2, \ldots)$$

then, the expected number of secondary transmissions and the variance are given by

$$R_0 = E(\rho_i^{(n)}) = \sum_{k=1}^{\infty} kp_k,$$

$$\text{Var}(\rho_i^{(n)}) = \sum_{k=0}^{\infty} (k - R_0)^2 p_k.$$  \hfill (50)

In other words, the concept of probability distribution $p_k$ reflects offspring distribution in population ecology, and this permits explicit modeling of variations in secondary transmissions in infectious diseases [102,103]. This approach is particularly important during the initial phase of an epidemic, because the number of infectious individuals is small in this stage, and thus, it is deemed essential to take into account demographic stochasticity, i.e. variation in the numbers of secondary transmissions by chance. Indeed, the model has been applied to observed outbreak data where we observed the extinction before growing to a major epidemic [104,105].

Let us briefly discuss the variation in secondary transmissions and an estimation method of $R_0$ using the discrete-time branching process, deriving analytical expressions of the expected number of infected individuals in generation $n$, $M_n = E(C_n)$ and the variance $V_n = \text{Var}(C_n)$. It is impossible to avoid using the probability generating function (pgf) to discuss the branching process. The above described $\rho_i^{(n)}$ characterize positive and discrete number of secondary transmissions, and thus, is a non-zero discrete random variable. The pgf of $\rho$, $g_\rho(s)$ is given by

$$g_\rho(s) = E(s^{\rho}) = \sum_{k=0}^{\infty} p_k s^k.$$  \hfill (51)

There are two basic properties concerning $g(s)$ in relation to the epidemic process. First, $R_0$ is by definition the mean value of secondary transmissions (Eq. (50)) and thus given by $g'(1)$. Second, the probability that an infected individual does not cause any secondary transmissions $p_0 = \Pr(\rho = 0)$ is given by $g(0)$, which is useful for discussing threshold phenomena and extinction [101]. If we note that $C_0 = 1$ (i.e. only one index case), the Galton–Watson process has the following pgf identity:

$$g_0(s) = s,$$

$$g_{n+1}(s) = g_n(g(s)) = g(g_n(s)).$$  \hfill (52)

Even when there are $C_0 = a$ cases in generation 0 (where $a > 1$), we just have to assume that there are $a$ different independent infection-trees and thus

$$g_{C_0}(s) = s^a,$$

$$g_{C_n}(s) = (g_n(s))^a.$$  \hfill (53)

From the above discussions, the expected number of cases in generation $n$, $M_n$, and the variance $V_n$ is

$$M_n = R_0^n M_0,$$

$$V_n = \begin{cases} 
  n \text{Var}(\rho) & (R_0 = 1), \\
  R_0^{n-1} \text{Var}(\rho) \frac{R_0^n}{R_0 - 1} & (R_0 \neq 1).
\end{cases}$$  \hfill (54)
The process grows geometrically if $R_0 > 1$, stays constant if $R_0 = 1$, and decays geometrically if $R_0 < 1$. These three cases are referred to as supercritical, critical, and subcritical, respectively. However, unlike the deterministic model, it should be remembered that critical process does not permit continued transmissions, and rather, the process becomes extinct almost surely (i.e. probability of extinction given $R_0 = 1$ is one) [101].

Mathematically, demographic stochasticity in transmission is represented by a Poisson process, which has been practiced in the application of branching processes to epidemics [17]. Assuming that mean value of secondary transmissions is a constant $R_0$, the conditional distribution of observing $C_{n+1}$ cases, given $C_n$ cases, follows a Poisson distribution:

$$C_{n+1} | C_n \sim \text{Poisson}[C_n R_0].$$

Supposing that we analyze influenza data documenting the generations of cases from 0 to $n$ in which we observed geometric growth, the likelihood of estimating $R_0$ is proportional to

$$\prod_{k=1}^{n} (R_0 C_{k-1})^{C_k} \exp(-R_0 C_{k-1}).$$

Here we apply the above model to the Spanish influenza data in Zurich (Fig. 3). Assuming that the generation time of length $\tau$, $w(\tau)$, is given by the following delta function with the mean length 3 days,

$$w(\tau) = \begin{cases} \infty, & \text{for } \tau = 3, \\ 0, & \text{for } \tau \neq 3, \end{cases}$$

then the observed series of data can be grouped by generation $(C_0, C_1, C_2, \ldots)$:

$$1, \ 3, \ 4, \ 7, \ 26, \ 30, \ 37, \ \ldots.$$ (58)

Since we assumed exponential growth during the initial 16 days in the previous section, here we similarly assume a geometric increase up to the 6th generation. Applying (56) to the above data, maximum likelihood estimate of $R_0$ (and the corresponding 95 percent confidence intervals) is 1.51 (1.24, 1.81). The model is simple enough to estimate $R_0$, and indeed, a slight extension of the discrete-time branching process has been employed to estimate $R_0$ as well as the proportion of undiagnosed cases in the analysis of SARS outbreak data [106].

It should be noted that the discrete-time branching process assumes a homogeneous pattern of spread. A technical issue has arisen on this subject during the SARS outbreak. Usually, we observe some cases who produce an extraordinary number of secondary cases compared with other infected individuals, which are referred to as superspreaders. Because of this, observed offspring distributions for directly transmitted diseases tend to be extremely skewed to the right. Empirically, it has been suggested that Poisson offspring distribution is sometimes insufficient to highlight the presence of superspreaders in epidemic modeling [107]. For example, if non-zero discrete distribution of secondary cases follows a geometric distribution with mean $R_0$, the pgf is given by a geometric distribution with mean $R_0$:

$$g(s) = \frac{1}{1 + R_0 (1 - s)}. \tag{59}$$

Moreover, if the offspring distribution follows gamma distribution with mean $R_0$ and dispersion parameter $k$, the pgf $g(s)$ follows negative binomial distribution [108]:

$$g(s) = \left(1 + \frac{R}{k} (1 - s)\right)^{-k}. \tag{60}$$

We still do not know if pandemic influenza is also the case to warrant the skewed offspring distributions. To explicitly test if superspreading events frequently exist in influenza transmission, it is necessary to accumulate contact tracing data for this difficult disease, the cases of which often show flu-like symptoms only (as discussed in Section 1). In addition, it should be noted that we cannot attribute the skewed offspring distribution to the underlying contact network only. To date, there are two major reasons which can generate superspreaders: (i) those who experience very frequent contacts (social superspreader) or (ii) those who are suffering from high pathogen loads or those who can scatter the pathogen through the air such as the use of nebulizer in hospitals (biological superspreader). From the offspring distribution only, we cannot distinguish these two mechanisms.
5.2. Counting process

With regard to the estimation of $R_0$ using final size, we briefly discuss another method based on a stochastic process. As we discussed above, let $S(t)$, $I(t)$ and $U(t)$ be the numbers of susceptible, infectious and recovered individuals at time $t$, respectively. Further, let $\beta$ and $1/\gamma$ denote the transmission rate and the mean duration of the infectious period, respectively. Supposing that $K(t)$, the number of individuals who experienced infection between time 0 and time $t$, is given by $K(t) = S(0) - S(t)$, the two processes $K(t)$ and $U(t)$ are increasing counting processes where the general epidemic is explained by:

$$\Pr(dK(t) = 1, dU(t) = 0 \mid \mathcal{Z}_t) = \beta \bar{S}(t)I(t) dt,$$

$$\Pr(dK(t) = 0, dU(t) = 1 \mid \mathcal{Z}_t) = \gamma I(t) dt,$$

$$\Pr(dK(t) = 0, dU(t) = 0 \mid \mathcal{Z}_t) = 1 - \beta \bar{S}(t)I(t) dt - \gamma I(t) dt, \quad (61)$$

where $\mathcal{Z}_t$ denotes the $\sigma$-algebra generated by the history of the epidemic $\{S(u), I(u); 0 < u < t\}$ and $\bar{S}(t) = S(t)/n$ (where $n$ is the size of the susceptible population at time 0). The latter is equivalent to assuming density-independent transmission (i.e. also referred to as true mass action or frequency dependent assumption [109]). Based on Eq. (61), two zero-mean martingales [110] are defined by:

$$M_1(t) = K(t) - \int_0^t \beta \bar{S}(u)I(u) du,$$

$$M_2(t) = U(t) - \int_0^t \gamma I(u) du. \quad (62)$$

From the martingale theory [111], a zero-mean martingale is given by

$$M(t) = \int_0^t \frac{1}{\bar{S}(u)} dM_1(u) - \frac{\beta}{\gamma} M_2(t) = \frac{n}{S(0)} + \frac{n}{S(0)-1} + \cdots + \frac{n}{S(t)+1} - \frac{\beta}{\gamma} U(t). \quad (63)$$

Thus, the estimator $\hat{\theta} = \hat{\beta}/\hat{\gamma}$ is given by

$$\hat{\theta} = \frac{\left[ \frac{n}{S(0)} + \frac{n}{S(0)-1} + \cdots + \frac{n}{S(T)+1} \right]}{U(T)} = -\ln(1 - \tilde{p}), \quad (64)$$

where $\tilde{p}$ is the observed final size ($= 1 - S(T)/n$) at the end of the epidemic at time $T$. Furthermore, the variance of the zero-martingale is given by

$$\text{Var}(M(t)) = \text{Var}(M_1(t)) + \theta^2 \text{Var}(M_2(t)). \quad (65)$$

From the martingale central limit theorem [112], the estimator $\theta$ is approximately normally distributed in a major outbreak in a large community. The standard error is then consistently estimated by:

$$\text{s.e.}(\hat{\theta}) = \left[ \frac{\left( \frac{n}{S(0)} + \frac{n}{S(0)-1} + \cdots + \frac{n}{S(T)+1} - \theta^2 R(T) \right)}{U(T)} \right]^{1/2}, \quad (66)$$

Consequently, the estimator and standard error of $R_0$ are given by:

$$\hat{R}_0 = n \times \hat{\theta},$$

$$\text{s.e.}(\hat{R}_0) = n \times \text{s.e.}(\hat{\theta}). \quad (67)$$

More detailed mathematical descriptions can be found elsewhere [74,113,114].
Here we show a numerical example. Let us consider a large epidemic of equine influenza (i.e., influenza in horses) as our case study. In 1971, a nationwide epidemic of equine-2 influenza A (H3N8) was observed in Japan [115]. For example, in Niigata Racecourse, 580 influenza cases were diagnosed with influenza among a total of 640 susceptible horses. The final size \( p \) is thus 90.6 percent (95 percent CI: 88.4, 92.9). From this data, we calculate \( R_0 \) and its uncertainty bounds.

Using \( \hat{p} = 0.906 \) and total number of infected \( U(T) = 580 \) in Eq. (64), \( \hat{\theta} \) is estimated as 0.00408. Therefore, the estimate of \( R_0 = 2.60 \) is given by Eq. (67). Moreover, from Eq. (66) where \( S(T) = 639 - 579 = 60 \) and \( S(0) = n = 639 \) (we assume one case was already infected at time \( t = 0 \)), we obtain \( s.e.(\hat{\theta}) = 0.000126 \). Here, \( \hat{\theta} \) is assumed to follow normal distribution. Therefore, the 95 percent confidence interval for \( R_0 \) is given as \( [R_0 - 1.96 \times 640 \times 0.000126, R_0 + 1.96 \times 640 \times 0.000126] = [2.44, 2.76] \).

When applying the final size equation, it should be remembered that (i) we assume all individuals are initially susceptible (in the above described model) and (ii) we assume \( \beta \) and \( \gamma \) are independent of time (i.e., constant), and thus, that any extrinsic factors should not have influenced the course of the observed epidemic.

5.3. Epidemics with two levels of mixing

In the above described models, we always assumed that the pattern of influenza transmission is homogeneous, which is clearly unrealistic to capture the transmission dynamics of influenza. Since the last century, it has already been understood that the transmission dynamics are not sufficiently modeled by assuming homogeneous mixing. However, because more detailed data are lacking (e.g., epidemic records of pandemic influenza with time, age and space), what we could offer has been mainly to extract the intrinsic growth rate from the initial exponential growth, and estimate \( R_0 \) using the estimator based on a model with the homogeneous mixing assumption.

One line of addressing heterogeneous patterns of transmission using the observed data is separating household transmission from community transmission. In other words, it is of practical importance to distinguish between individual and group \( R_0 \) [116]. From the beginning of explicit modeling of influenza [117,118], a method to separately estimate the transmission parameters has been proposed, which has been partly extended in a recent study [92] or applied to further old data of pandemic influenza [119]. Indeed, an important aspect of this issue was highlighted in a recent study which compared estimates of \( R_0 \) between those having casual and close contacts [63]. To estimate key parameters of household and community transmissions of influenza, or to simulate realistic patterns of influenza spread, such a consideration is fruitful.

Mathematically elaborating this concept, there are several publications which proposed the basis of analyzing household transmission data employing stochastic models [120–122]. Moreover, a rigorous study has been made to estimate parameters determining the intrinsic dynamics (e.g., infectious period) using household transmission data with time [123].

Future challenges on the estimation of \( R_0 \) include the application of such theories to the observed data with some extension. For example, as we discussed above, knowing the generation time would be crucial to elucidate a robust estimate of \( R_0 \) [6,89,90,92]. However, we do not know if the generation time varies between close and casual contacts; this should be the case, because, as long as the generation time is given by coevolution of latent and infectious periods, close contact should lead to shorter generation time than casual contact. In future studies, influenza models may better to highlight the increasing importance of considering household transmission to estimate the transmission potential using the temporal distribution of infection events.

6. Characteristics of influenza data

Except for our approach in Section 3, mathematical arguments given in this paper are not particularly special for influenza. In other words, we modelers have employed similarly structured models which describe the population dynamics of other directly transmitted diseases, and such models are applicable not only for influenza but also for many viral diseases including measles, smallpox, chickenpox, rubella and so on [18]. However, influenza has many different epidemiologic characteristics compared to other childhood viral diseases. For instance, following the previous efforts in influenza epidemiology [124–126] and modeling [127,128], we should at least note the following:
1. Detailed mechanisms of immunity have yet to be clarified. Since influenza virus has an wide antigenic diversity (i.e. unlike other childhood viral diseases, antigenic stimulation is not monoclonal), this complicates our understanding in the fraction of immune individuals, cross-protection mechanisms and evolutionary dynamics [129–132].

2. Flu-like symptoms are too common, and thus, we cannot explicitly distinguish influenza from other common viral infections without expensive laboratory tests for each case. Because of this character, it is difficult to effectively implement usual public health measures (e.g., contact tracing and isolation).

3. Although explicit estimates are limited [133,134], a certain fraction of infected individuals does not exhibit any symptoms (following infection). This complicates not only the eradication [135] but also epidemiologic evaluations of vaccines and therapeutics [136].

4. Looking into the details of the intrinsic dynamics [4,123], it appears recently that the generation time and infectious period are much shorter than what were believed previously. Therefore, despite the relatively small $R_0$ estimate, the turn-over of a transmission cycle (i.e. speed of growth) is rather quick. The incubation period of Spanish influenza is as short as 1.5 days [137], complicating the implementation of quarantine measures [138].

Thus, depending on the characteristics of observed data (and the specific purpose of modeling), we have to highlight these factors referring to the best available evidence. This is one of the most challenging issues in designing public health interventions against a potential future pandemic.

6.1. What is reported in the observed data?

In addition to the above described issue, we, of course, must remember what the reported data is. In many studies, the compartment $I(t)$ or relevant class of infectious individuals of the SIR (or SEIR) model was fitted to the observed data. Indeed, in the majority of previous classic studies, $R(t)$ (i.e. removed class; denoted by $U(t)$ in our discussion) of Kermack and McKendrick model was fitted to the data, assuming that the removed class highlights observed data as the reported cases no longer produce secondary cases. However, the observed epidemiologic data is actually neither $I(t)$ nor $U(t)$. Always, what we get as the temporal distribution reflects case onset or deaths with time which is mostly accompanied by some reporting delay.

We believe this is one of the most challenging issues in epidemic modeling. Except for rare examples in sexually transmitted diseases, infection event is not directly observable, and thus, we have to maximize the utility of reported (observed) data, explicitly understanding what the data represents.

In this case, back-calculation of the infection events is called for. Let $c(t)$ denote the number of onsets at time $t$, this should be modeled by using incidence $j(t)$ and the density of the incubation period of length $\tau$, $f(\tau)$:

$$c(t) = \int_0^\infty j(t - \tau) f(\tau) d\tau. \quad (68)$$

Further, supposing that $b(t)$ is the number of reported cases at time $t$ and the density of reporting delay of length $\sigma$ is $h(\sigma)$, observed data is modeled as:

$$b(t) = \int_0^\infty c(t - s) h(s) ds = \int_0^\infty \int_0^\infty j(t - s - \tau) f(\tau) d\tau h(s) ds. \quad (69)$$

That is, only by using the observed data $b(t)$ and known information of the reporting delay $h(s)$ and incubation period distributions $f(\tau)$, we can translate the observed data into infection process $j(t)$.

In some cases, only death data with time, $d(t)$, is available [60]. Similarly, this can be modeled using the back-calculation. Let $q$ denote the case fatality of influenza which is reasonably assumed time-independent, and further let $m(u)$ be the relative frequency of time from onset to death, $d(t)$ is given by:

$$d(t) = q \int_0^\infty c(t - u) m(u) du. \quad (70)$$
Even when using onset data with delay or death data, it should be noted that the intrinsic growth rate is identical to that estimated from the infection event distribution. Assuming that the incidence \( j(t) \) exhibits exponential growth during the initial phase of an epidemic, i.e. \( j(t) = k \exp(rt) \), Eqs. (68) and (70) can be rewritten as

\[
b(t) = \int_0^\infty \int_0^\infty k \exp(r(t-s-\tau)) f(\tau) d\tau h(s) ds = k \exp(rt) \int_0^\infty \int_0^\infty \exp(-r(s+\tau)) f(\tau) d\tau h(s) ds \tag{71}
\]

and

\[
d(t) = q \int_0^\infty \int_0^\infty j(t-u-\tau) f(\tau) d\tau m(u) du = q \int_0^\infty \int_0^\infty k \exp(r(t-u-\tau)) f(\tau) d\tau m(u) du \]

\[
= qk \exp(rt) \int_0^\infty \int_0^\infty \exp(-r(u+\tau)) f(\tau) d\tau m(u) du. \tag{72}
\]

Thus, the growth terms \( \exp(rt) \) (i.e. which depends on time) of \( b(t) \) and \( d(t) \) are still identical to that of incidence \( j(t) \). In other words, mathematically Eqs. (71) and (72) could be a justification to extract an estimate of the intrinsic growth rate from cases with reporting delay or deaths with time. However, we should always remember that the infection-age distribution is not stable during the initial phase, and moreover, this method cannot address individual variation in the secondary transmissions (e.g., superspreaders, as we discussed in Section 5).

### 6.2. What to be learnt from the reported data?

In this way, it is not an easy task to clarify the infection events with time. A similar application of the convolution equation has been intensively studied in modeling HIV/AIDS. Since AIDS has a long incubation period, and because AIDS diagnosis is certainly reported in the surveillance data (at least, in industrialized countries), backcalculation of the number of HIV infections with time using the number of AIDS diagnoses and the incubation period distribution has been an issue to capture the whole epidemiologic picture of HIV/AIDS [139–142]. In the current modeling practice using the temporal distribution of onset events, we are now faced with a need to apply this technique to diseases with much shorter incubation periods.

Now, let us look back at a method to estimate \( R(t) \), which was proposed by Wallinga and Teunis [41]. Whereas the method has a background of mathematical reasoning (as shown in (46), Section 4.2), the estimator was derived implicitly assuming that observed data exactly reflects infection events. If asymptomatic infection and transmission are rare, this assumption might be justifiable as the lengths of the serial interval and generation time become almost identical. However, as long as we cannot ignore asymptomatic transmissions, which is particularly the case for influenza, the assumption \( s(\sigma) = w(\tau) \) might be problematic [85].

Since \( R(t) \) of this method was given by summing up the probability of causing secondary transmissions by an onset case at the onset time of this case \( t \), we should rewrite the assumption using a modified onset-based renewal equation as follows [60]:

\[
c(t) = \int_0^t c(t-\tau)R(t-\tau)s(\tau) d\tau. \tag{73}
\]

For simplicity, we ignore reporting delay in the observed data, roughly assuming that the observed data reflects \( c(t) \). Translating Eq. (73) in words, it is implicitly assumed that secondary transmission happens exactly at the time of onset, and based on this assumption, \( R(t) \) in the right-hand side of (73) can be backcalculated.
To understand the assumptions behind the above equation, let us assume that incidence $j(t)$ is given by

$$j(t) = S(t) \int_0^\infty \beta(\sigma) \Gamma(\sigma) c(t - \sigma) d\sigma,$$

(74)

where $\beta(\sigma)$ is the transmission rate at disease-age $\sigma$ (i.e. the time since onset of infection [143]) and $\Gamma(\sigma)$ is the survivorship of cases following onset. It should be noted that Eq. (74) ignores secondary transmissions before onset of illness. As we discussed above, $c(t)$ is given by $j(t)$ and the incubation period distribution $f(\tau)$,

$$c(t) = \int_0^\infty j(t - \tau) f(\tau) d\tau.$$  

(75)

Replacing $c(t)$ in the right-hand side of (74) by (75), we get

$$j(t) = S(t) \int_0^\infty j(t - s) \phi(s) ds,$$

(76)

where $s$ represents infection-age (i.e. time since infection), and $\phi(s)$ is given by

$$\phi(s) = \int_0^s \beta(a) \Gamma(a) f(s - a) da$$

(77)

which represents generation time distribution. From Eqs. (76) and (77), we can find that $R(t)$ is given by

$$R(t) = S(t) \int_0^\infty \phi(s) ds = S(t) \int_0^\infty \beta(\sigma) \Gamma(\sigma) d\sigma \int_0^\infty f(\tau) d\tau.$$  

(78)

Eq. (78) can be further reduced to $R(t) = R_0 S(t)/S(0)$ which represents Kermack and McKendrick’s assumption. Replacing $j(t)$ in the right-hand side of (75) by (74), we get

$$c(t) = \int_0^\infty \psi(t, \sigma) c(t - \sigma) d\sigma,$$

(79)

where $\psi(t, \sigma)$ denotes the serial interval distribution of calender time $t$ and disease-age $\sigma$:

$$\psi(t, \sigma) = \int_0^\sigma \beta(\sigma - \tau) \Gamma(\sigma - \tau) f(\tau) S(t - \tau) d\tau.$$  

(80)

Eq. (80) is difficult to solve as it includes $S(t - \tau)$ in the right-hand side. However, in the special case, e.g., let us say when we can assume $\beta(\tau) \Gamma(\tau) = k \delta(\tau)$ (where $k$ is constant and $\delta(\cdot)$ is delta function),

$$\psi(t, \sigma) = kf(\sigma) S(t - \sigma).$$  

(81)

Inserting (81) back to (79),

$$c(t) = \int_0^\infty kf(\sigma) S(t - \sigma) c(t - \sigma) d\sigma = \int_0^\infty R(t - \sigma) c(t - \sigma) \frac{f(\sigma)}{\int_0^\infty f(\tau) d\tau} d\sigma$$

(82)

which is onset-based renewal equation which was presented in (73). What to be learnt from (82) is, the assumption that secondary transmission happens immediately after onset suggests that the incubation period distribution is identical to the serial interval distribution as shown above, which is a bit funny conclusion. Maximizing the utility of observed data has still remained an open question.
In addition to modeling the temporal distribution, explicit modeling of asymptomatic infection is also called for [135]. Provided that there are so many asymptomatic transmissions which are not in the negligible order, we need to shift our concept of transmissibility; e.g., rather than \( R_0 \), a threshold quantity of symptomatic infection is required. In such a case, application of type-reproduction number \( T \) might be useful [144,145], and it has already been put into practice [146].

7. What to be clarified further?

In this review, we focused on the use of the temporal distribution of influenza to estimate \( R_0 \) (or \( R(t) \)) and the relevant key parameters. It must be remembered that our arguments, almost necessarily, employed the homogeneous mixing assumption, as we cannot extract information on heterogeneous patterns of infection from a single stream of temporal data alone. Presently, more information (e.g., at least, spatio-temporal distribution) is becoming available for influenza. In this section, we briefly sketch what can be (and should be) done in the future to quantify the transmission dynamics of pandemic influenza.

7.1. What \( R_0 \) really means?

It is not a new issue that heterogeneous patterns of transmission could even destroy the mean field theory in infectious diseases. For example, in a pioneering study of gonorrhea transmission dynamics by Hethcote and Yorke [147], an importance of contact heterogeneity was sufficiently highlighted. Since a simple model assuming homogeneous mixing did not reflect the patterns of gonorrhea transmission in the United States, Hethcote and Yorke divided the population in question into two; those who are sexually very active and not, the former of which was referred to as core group. Compared with the temporal distribution of infection given by a model with homogeneous mixing assumption, the simple heterogeneous model with a core group revealed much quicker increase in epidemic size, showing rather different trajectory of an epidemic. Given that the variance of sexual partnership is extremely large (i.e. if the distribution of the frequency of sexual intercourse is extremely skewed to the right with a very long right tail), the estimate of \( R_0 \) is shown to become considerably high. The finding supports a vulnerability of our society to the invasion of sexually transmitted diseases. Following this pioneering study, considerable efforts have been made to approximately model the heterogeneous patterns of transmission using extended mean field equations [18,26,148,149]. In addition to such an approximation of heterogeneous transmission, recent progress in epidemic modeling with explicit contact network structures suggests that variance of the contact frequency plays a key role in determining the threshold quantity, and in some special cases, the concept of threshold phenomena could be confused [150–153].

In Section 4, we defined the force of infection as

\[
\lambda(t) = \int_0^\infty \beta(\tau) i(t, \tau) d\tau.
\]  

(83)

In deterministic models given by simple ODEs (which ignores infection-age), \( \lambda(t) \) is equivalent to \( \beta I(t) \). These are what classical mean field models suggest.

Let us account for an epidemic on networks, whose node-connectivity distribution (i.e. the distribution of probabilities that nodes have exactly \( k \) neighbors) follows some explicit distribution \( P(k) \). The force of infection \( \lambda_c \), which yields \( R_0 = 1 \), in a static contact network is given by

\[
\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}.
\]

(84)

Here \( \langle k \rangle \) denotes the average connectivity of the nodes. Assuming that \( P(k) \) follows a power law of the form \( P(k) = c k^{-v} \) (where \( c \) is constant),

\[
\langle k^2 \rangle = c \sum_k k^{2-v}.
\]

(85)

Given that \( v \leq 3 \), \( \lambda_c = 0 \), and in such a case, \( R_0 \) even becomes infinite. This implies that the disease spread will continue for any mean estimate of \( R_0 \). Such a network structure is referred to as scale free [154], complicating disease
control efforts in public health [155,156]. The importance of the network structure would also be highlighted for \( v > 3 \).

For sexually transmitted infections, contact frequency is countable (unlike airborne infection or transmission through droplets), and \( v \) is estimated to be around 3 or a little larger [157]. Following such a finding, many non-sexual directly transmitted diseases are also modeled in the present day assuming the scale-free network [150]. However, it should be noted that the pattern of contact does not necessarily follow scale-free for all directly transmitted diseases. Indeed, there is no empirical evidence which suggests that the contact structure of any droplet infections follows the power law (i.e. we do not know if the above described contact heterogeneity is the case for diseases except for sexually transmitted diseases). A typical example of confusion is seen in the superspreading events during the 2002/03 SARS epidemics [158], where we cannot explicitly attribute the phenomena either to contact network or biological factors (as long as contact and infection event are not directly observable). We still do not know how we should account for the distribution \( P(k) \) for influenza and other viral respiratory diseases (i.e. power law or not) which remains to be clarified for each disease in future research.

7.2. New theory to replace mass action principle

Methodological developments have been made to account for the network heterogeneity with data [159]. An approximate approach to address this issue is highlighted particularly in spatio-temporal modeling, an excellent account of which is reviewed by Matt Keeling [160].

Even though it is difficult to quantify the transmission dynamics with an explicit contact network with time, there are useful analytical approximations to capture the dynamics of influenza (and other respiratory transmitted viral diseases) and estimate the transmission potential. For example, the force of infection with a power law approximation is reasonably given by:

\[
\lambda(t) = \beta I(t)^{1+\alpha} S(t)^{1+\Psi}.
\]

In (86), \( \alpha \) and \( \Psi \) characterize the epidemic dynamics; e.g., initial growth (i.e. if \( \alpha \) is less than 0, the modified form acts to dampen the exponential growth of incidence) and endemic equilibrium (i.e. when \( \Psi \) is greater than 0, density-dependent damping is increased). A model of this type was actually validated with measles data in England and Wales, comparing the prediction with that of employing the mass action principle [161].

Another approximation might be a pair-wise model [162], which can explicitly account for the correlation between connected pairs. The model reasonably permits deriving the force of infection \( \lambda \) using the number of various connected pairs, which implies wide applicability to the epidemiologic data of sexually transmitted infections. Incorporating spatial heterogeneity in an approximate manner would shed light on further quantifications [163,164], and thus, simple and reasonably tractable models which permit spatio-temporal modeling of influenza are expected (e.g., [165]).

7.3. Which kind of data do we have to explore?

Summarizing the above discussions, we have presented modeling approaches that can quantify the transmission potential of pandemic influenza. As we have shown, temporal case distributions have been analyzed in many instances, and previous efforts have come close to maximize the utility of temporal distributions (i.e. epidemic curve). However, at the same time, we have also learned that we can extract almost the intrinsic growth rate alone from a single time-evolution data. Accordingly, we are now faced with a need to clarify heterogeneous patterns of transmission and more detailed intrinsic dynamics of influenza [166–168]. With regard to the latter, primitive epidemiologic questions (e.g., probability of clinical attack given infection) remain to be answered for Spanish, Asian and Hong Kong influenza. Let us summarize what we need to clarify theoretically about pandemic influenza in list:

1. Acquired immunity.
2. Evolutionary dynamics.
3. Multi-host species transmission.
4. Asymptomatic transmission.
5. Attack rate (i.e. \( \text{Pr(onset}\mid\text{infection}) \)).
6. Case fatality (i.e. \( \text{Pr(death}\mid\text{onset}) \)).
7. Generation time and serial interval.
8. Latent, incubation, infectious and symptomatic periods with further data.
9. Transmission potential with time, space and antigenic types.
10. Transmission potential with time and age.

These issues highlight the importance to quantify the transmission of influenza using not only cases (i.e. those followed onset of symptoms) but also some hint suggesting the infection event. For example, the majority of the above listed issues could be reasonably addressed by implementing serological surveys (e.g., antibody titers of individuals and, preferably, time-delay distribution from infection to seroconversion). Since the proportion of those who do not experience symptomatic infection (i.e. probability of asymptomatic infection) is not small for influenza [64,146], case records can tell us little to address the above mentioned issues, and thus, historical data of Spanish influenza may hardly offer further information. By maximizing the utility of observed data, we have to clarify the dynamics of influenza further, and identify key information which characterize the specific mechanisms of spread.

References

[1] Simonsen L. The global impact of influenza on morbidity and mortality. Vaccine 1999;17:S3–10.
[2] Park AW, Glass K. Dynamic patterns of avian and human influenza in east and southeast Asia. Lancet Infectious Diseases 2007;7:543–8.
[3] Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of influenza A viruses. Microbiological Reviews 1992;56:152–79.
[4] Ferguson NM, Cummings DAT, Cauchemez S, Fraser C, Riley S, Meeyai A, et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature 2005;437:209–14.
[5] Ferguson NM, Cummings DA, Fraser C, Cajka JC, Cooley PC, Burke DS. Strategies for mitigating an influenza pandemic. Nature 2006;442:448–52.
[6] Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proceedings B 2007;274:599–604.
[7] Longini IM, Halloran ME, Nizam A, Yang Y. Containing pandemic influenza with antiviral agents. American Journal of Epidemiology 2004;159:623–33.
[8] Mills CE, Robins JM, Lipsitch M. Transmissibility of 1918 pandemic influenza. Nature 2004;432:904–6.
[9] Nicholson K, Hay A. Textbook of influenza. Oxford: Blackwell; 1998.
[10] Cunha BA. Pandemic influenza: a review. Population and Development Review 2007;33:429–51.
[11] Murray CJ, Lopez AD, Chin B, Feehan D, Hill KH. Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: a quantitative analysis. Lancet 2006;368:2211–8.
[12] Sydenstricker E. Variations in case fatality during the influenza epidemic of 1918. Public Health Reports 1921;36:2201–11.
[13] Markel H, Lipman HB, Navarro JA, Stoan A, Michalsen JR, Stern AM, et al. Nonpharmaceutical interventions implemented by US cities during the 1918–1919 influenza pandemic. JAMA 2007;298:644–54.
[14] Patterson KD, Pyle GF. The geography and mortality of the 1918 influenza pandemic. Bulletin of the History of Medicine 1991;65:4–21.
[15] Johnson NP, Mueller J. Updating the accounts: global mortality of the 1918–1920 “Spanish” influenza pandemic. Bulletin of the History of Medicine 2002;76:105–15.
[16] MacKellar L. Pandemic influenza: a review. Population and Development Review 2007;33:429–51.
[17] Diekmann O, Heesterbeek JAP. Mathematical epidemiology of infectious diseases: model building, analysis and interpretation. New York: John Wiley and Sons; 2000.
[18] Anderson RM, May RM. Infectious diseases of humans: dynamics and control. Oxford: Oxford University Press; 1991.
[19] Anderson RM, May RM. Directly transmitted infectious diseases: control by vaccination. Science 1982;215:1053–60.
[20] Smith CE. Factors in the transmission of virus infections from animal to man. The Scientific Basis of Medicine Annual Reviews 1964:125–50.
[21] Halloran ME, Haber M, Longini IM, Struchiner CJ. Direct and indirect effects in vaccine efficacy and effectiveness. American Journal of Epidemiology 1991;133:323–31.
[22] Fine PE. Herd immunity: history, theory, practice. Epidemiologic Reviews 1993;15:265–302.
[23] Nishiura H, Dietz K, Eichner M. The earliest notes on the reproduction number in relation to herd immunity: Theophil Lotz and smallpox vaccination. Journal of Theoretical Biology 2006;241:964–7.
[24] Heesterbeek JAP. A brief history of $R_0$ and a recipe for its calculation. Acta Biotheoretica 2002;50:189–204.
[25] Nishiura H, Inaba H. Discussion: Emergence of the concept of the basic reproduction number from mathematical demography. Journal of Theoretical Biology 2007;244:357–64.
[26] Diekmann O, Heesterbeek JAP, Metz JAJ. On the definition and the computation of the basic reproductive ratio $R_0$ in models for infectious diseases. Journal of Mathematical Biology 1990;35:503–22.
[27] van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Mathematical Biosciences 2002;180:29–48.
[28] Castillo-Chavez C, Feng Z, Huang W. On the computation of $R_0$ and its role on global stability. In: Mathematical approaches for emerging and reemerging infectious diseases: an introduction. IMA, vol. 125. Berlin: Springer-Verlag; 2002. p. 229–50.
[64] Mathews JD, McCaw CT, McVernon J, McBayde ES, McCaw JM. A biological model for influenza transmission: pandemic planning implications of asymptomatic infection and immunity. PLoS ONE 2007;2:e1220.

[65] Chowell G, Bettencourt LMA, Johnson N, Alonso WJ, Viboud C. The 1918–1919 influenza pandemic in England and Wales: Spatial patterns in transmissibility and mortality impact. Proceedings of the Royal Society B 2008, in press, doi:10.1098/rspb.2007.1477.

[66] Sattenspiel L, Herring DA. Structured epidemic models and the spread of influenza in the central Canadian subarctic. Human Biology 1998;70:91–115.

[67] Sattenspiel L, Herring DA. Simulating the effect of quarantine on the spread of the 1918–19 flu in central Canada. Bulletin of Mathematical Biology 2003;65:1–26.

[68] Brauer F, Castillo-Chavez C. Mathematical models in population biology and epidemiology. New York: Springer-Verlag; 2000.

[69] Hethcote HW. The mathematics of infectious diseases. SIAM Review 2000;42:599–653.

[70] Ross R. The prevention of malaria. London: John Murray; 1911.

[71] Dietz K. Mathematical models for transmission and control of malaria. In: Wernsdorfer WH, McGregor I, editors. Malaria, principles and practice of malariology. Edinburgh: Churchill Livingstone; 1988. p. 1091–133.

[72] Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics. I. Proceedings of the Royal Society Series A 1927;115:700–21; reprinted in Bulletin of Mathematical Biology 1991;53:33–55.

[73] Ma J, Earn DJ. Generality of the final size formula for an epidemic of a newly invading infectious disease. Bulletin of Mathematical Biology 2006;68:679–702.

[74] Becker NG. Analysis of infectious disease data. New York: Chapman and Hall; 1989.

[75] Hrenoff AK. The influenza epidemic of 1918–1919 in San Francisco. The Military Surgeon 1941;89:805–11.

[76] Cooper B, Lipsitch M. The analysis of hospital infection data using hidden Markov models. Biostatistics 2004;5:223–37.

[77] Nishiura H, Kuratsuji T, Quy T, Phi NC, Van Ban V, Ha LE, et al. Rapid awareness and transmission of severe acute respiratory syndrome in Hanoi French Hospital, Vietnam. American Journal of Tropical Medicine and Hygiene 2005;73:17–25.

[78] Davidian M, Giltinan DM. Nonlinear models for repeated measurement data. Monographs on statistics and applied probability, vol. 62. New York: Chapman and Hall; 1995.

[79] Bickel P, Doksum KA. Mathematical statistics. Oakland, California: Holden-Day; 1977.

[80] Efрон B, Tibshirani RJ. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Statistical Science 1986;1:54–75.

[81] Haydon DT, Chase-Topping M, Shaw DJ, Matthews L, Friis SM, Wilesmith J, et al. The construction and analysis of epidemic trees with reference to the 2001 UK foot-and-mouth outbreak. Proceedings of the Royal Society of London Series B 2003;270:121–7.

[82] Cauchemez S, Boelle PY, Donnelly CA, Ferguson NM, Thomas G, Leung GM, et al. Real-time estimates in early detection of SARS. Emerging Infectious Diseases 2006;12:110–3.

[83] Diekmann O. Limiting behaviour in an epidemic model. Nonlinear Analysis, Theory, Methods and Applications 1977;1:459–70.

[84] Efron B, Tibshirani RJ. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Statistical Science 1986;1:54–75.

[85] Sato HJ, Rohani P, Keeling MJ. Appropriate models for the management of infectious diseases. PLoS Medicine 2005;2:e174.

[86] Lloyd AL. Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. Proceedings of the Royal Society of London Series B 2001;268:985–93.

[87] Lloyd AL. Realistic distributions of infectious periods in epidemic models: Changing patterns of persistence and dynamics. Theoretical Population Biology 2001;60:59–71.

[88] Roberts MG, Heesterbeek JA. Model-consistent estimation of the basic reproduction number from the incidence of an emerging infection. Journal of Mathematical Biology 2007;55:803–16.

[89] Yan P. 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. Journal of Theoretical Biology 2008, in press.

[90] Imahorn A. Epidemiologische Beobachtungen ueber die Grippe-Epidemie 1918 im Oberwallis. Inaugural-Dissertation zur Erlangung der Doktorwuerde der Medizinischen Fakultaet der Universitaet Zuerich. Zurich: Universitaet Zuerich; 2000 [in German].

[91] Fraser C. Estimating individual and household reproduction numbers in an emerging epidemic. PLoS ONE 2007;2:e758.

[92] Ferguson NM, Donnelly CA, Anderson RM. Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. Nature 2001;413:542–8.

[93] Fine PE. The interval between successive cases of an infectious disease. American Journal of Epidemiology 2003;158:1039–47.

[94] Nishiura H. Epidemiology of a primary plague pandemic in Kantoshu, Manchuria, from 1910 to 1911: statistical analysis of individual records collected by the Japanese Empire. International Journal of Epidemiology 2006;35:1059–65.

[95] Cowling BJ, Ho LM, Leung GM. Effectiveness of control measures during the SARS epidemic in Beijing: a comparison of the Rt curve and the epidemic curve. Epidemiology and Infection 2007, in press, doi:10.1017/S0950268807008722.

[96] Marques CA, Forattini OP, Massad E. The basic reproduction number for dengue fever in Sao Paulo state, Brazil: 1990–1991 epidemic. Transactions of the Royal Society of Tropical Medicine and Hygiene 1994;88:58–9.

[97] Galvani AP, Lei X, Jewell NP. Severe acute respiratory syndrome: temporal stability and geographic variation in case-fatality rates and doubling times. Emerging Infectious Diseases 2003;9:991–4.

[98] Lessler J, Cummings DA, Fishman S, Vora A, Burke DS. Transmissibility of swine flu at Fort Dix, 1976. Journal of the Royal Society Interface 2007;4:755–62.

[99] Haccou P, Jagers P, Vatutin VA. Branching processes: variation, growth and extinction of populations. Cambridge: Cambridge University Press; 2005.
[140] Jewell NP, Dietz K, Farewell VT, editors. AIDS epidemiology: methodological issues. Boston: Birkhauser; 1992.

[141] Colton T, Johnson T, Machin D, editors. Proceedings of the conference on quantitative methods for studying AIDS. Blaubeuren, Germany, June 14–18, 1993. Statistics in Medicine 1994;13:1899–2188.

[142] Nishiura H. Lessons from previous predictions of HIV/AIDS in the United States and Japan: epidemiologic models and policy formulation. Epidemiologic Perspectives and Innovations 2007;4:3.

[143] Nishiura H, Eichner M. Infectiousness of smallpox relative to disease age: estimates based on transmission network and incubation period. Epidemiology and Infection 2007;135:1145–50.

[144] Heesterbeek JA, Roberts MG. The type-reproduction number T in models for infectious disease control. Mathematical Biosciences 2007;206:3–10.

[145] Roberts MG, Heesterbeek JA. A new method for estimating the effort required to control an infectious disease. Proceedings of the Royal Society of London Series B 2003;270:1359–64.

[146] Inaba H, Nishiura H. On the dynamical system formulation of the type reproduction number for infectious diseases and its application to the asymptomatic transmission model. Mathematical Biosciences 2007, submitted for publication.

[147] Hethcote HW, Yorke JY. Gonorrhea transmission dynamics and control. In: Levin S, editor. Lecture notes in biomathematics, vol. 56. Springer: Berlin; 1980.

[148] Woolhouse ME, Dye C, Ettari JD, Smith T, Charlwood JD, Garnett GP, et al. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. Proceedings of the National Academy of Science USA 1997:94:338–42.

[149] Blower SM, McLean AR. Mixing ecology and epidemiology. Proceedings of the Royal Society of London Series B 1991;245:187–92.

[150] Newman MEJ. Spread of epidemic disease on networks. Physical Review E 2002;66:016128.

[151] May RM, Lloyd AL. Infection dynamics on scale-free networks. Physical Review E 2001;64:066112.

[152] Miller JC. Epidemic size and probability in populations with heterogeneous infectivity and susceptibility. Physical Review E 2007;76:010101.

[153] Duerr HP, Scheweh M, Leary CC, De Vlas SJ, Eichner M. Epidemic size and probability in populations with heterogeneous infectivity and susceptibility. Epidemiology and Infection 2007;135:1124–32.

[154] Albert R, Barabasi AL. Statistical mechanics of complex network. Review of Modern Physics 2002;74:47–97.

[155] Kiss IZ, Green DM, Kao RR. Disease contact tracing in random and clustered networks. Proceedings B 2005;272:1407–14.

[156] Kiss IZ, Green DM, Kao RR. Infectious disease control using contact tracing in random and scale-free networks. Journal of the Royal Society Interface 2006;3:55–62.

[157] Liljeros F, Edling CR, Amaral LAN, Stanley HE, Aberg Y. The web of human sexual contacts. Nature 2001;411:907–8.

[158] Masuda N, Konno N, Aihara K. Transmission of severe acute respiratory syndrome in dynamical small-world networks. Physical Review E 2004;69:031917.

[159] Keeling MJ, Eames KT. Networks and epidemic models. Journal of the Royal Society Interface 2005;2:295–307.

[160] Keeling MJ. Extensions to mass-action. In: Cuddington K, Besiner BE, editors. Ecological paradigms lost. Routes of theory change. Amsterdam: Elsevier; 2005. p. 107–42.

[161] Bjornstad ON, Finkenstadt BF, Grenfell BT. Dynamics of measles epidemics: Estimating scaling of transmission rates using a time series SIR model. Ecological Monographs 2002;72:169–84.

[162] Keeling MJ. The effects of local spatial structure on epidemiological invasions. Proceedings of the Royal Society of London Series B 1999;266:859–69.

[163] Riley S. Large-scale spatial transmission models of infectious disease. Science 2007;316:1298–301.

[164] Keeling MJ, Brooks SP, Gilligan CA. Using conservation of pattern to estimate spatial parameters from a single snapshot. Proceedings of the National Academy of Science USA 2004;101:9155–60.

[165] Viboud C, Bjornstad ON, Smith DL, Simonsen L, Miller MA, Grenfell BT. Synchrony, waves, and spatial hierarchies in the spread of influenza. Science 2006;312:447–51.

[166] Nishiura H. Smallpox during pregnancy and maternal outcomes. Emerging Infectious Diseases 2006;12:1119–21.

[167] Duerr HP, Brockmann SO, Pielotowski I, Scheweh M, Eichner M. Influenza pandemic intervention planning using InfluSim: pharmaceutical and non-pharmaceutical interventions. BMC Infectious Diseases 2007;7:76.

[168] Eichner M, Scheweh M, Duerr HP, Brockmann SO. The influenza pandemic preparedness planning tool InfluSim. BMC Infectious Diseases 2007;7:17.