MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **4**, Number **3**, July **2007**

http://www.mbejournal.org/

pp. **457–470**

ESTIMATING THE REPRODUCTION NUMBER FROM THE INITIAL PHASE OF THE SPANISH FLU PANDEMIC WAVES IN GENEVA, SWITZERLAND

Gerardo Chowell

Center for Nonlinear Studies (MS B284), Los Alamos National Laboratory
Los Alamos, NM 87545, USA

Catherine E. Ammon

Institute of Social and Preventive Medicine, Faculty of Medicine, CMU
PoBox 1211, Geneva 4, Switzerland

Nicolas W. Hengartner

Discrete Simulation Sciences (CCS-5), Los Alamos National Laboratory
Los Alamos, NM 87545, USA

James M. Hyman

Center for Nonlinear Studies (MS B284), Los Alamos National Laboratory
Los Alamos, NM 87545, USA

(Communicated by Abba Gumel)

Abstract. At the outset of an influenza pandemic, early estimates of the number of secondary cases generated by a primary influenza case (reproduction number, R) and its associated uncertainty can help determine the intensity of interventions necessary for control. Using a compartmental model and hospital notification data of the first two waves of the Spanish flu pandemic in Geneva, Switzerland in 1918, we estimate the reproduction number from the early phase of the pandemic waves. For the spring and fall pandemic waves, we estimate reproduction numbers of 1.57 (95% CI: 1.45, 1.70) and 3.10 (2.81, 3.39), respectively, from the initial epidemic phase comprising the first 10 epidemic days of the corresponding wave. Estimates of the variance of our point estimates of R were computed via a parametric bootstrap. We compare these estimates with others obtained using different observation windows to provide insight into how early into an epidemic the reproduction number can be estimated.

1. **Introduction.** The presence of the highly pathogenic A/(H5N1) influenza virus in avian populations in several regions of the world could lead to the next influenza pandemic. Pandemics are worldwide epidemics that affect a significant fraction of the population in terms of morbidity and mortality rates. The 1918 influenza pandemic (Spanish flu) was the most devastating of recent history with an estimated death toll of 20 million to 50 million deaths worldwide [1].

2000 *Mathematics Subject Classification.* 92D30.

Key words and phrases. spanish flu; 1918 influenza pandemic; reproduction number; bootstrapping; nonlinear curve fitting; uncertainty quantification; Geneva; Switzerland.

In an influenza pandemic, public health officials must balance the costs associated with interventions, while containing the outbreak. The average number of secondary cases generated by a primary infectious case (reproduction number, R) provides a measure of the transmissibility of the infectious agent that can be used to help set appropriate priorities for the interventions. When the population is entirely susceptible at the beginning of an epidemic, the basic reproduction number (R_0) [2, 3] provides a measure of the transmissibility of an infectious disease. An epidemic is expected to occur whenever $R_0 > 1$, while an outbreak is unlikely to occur with $R_0 < 1$. For influenza and other recurrent infectious diseases, fraction p of the population is often protected against the infectious agent because of previous exposures to influenza and vaccination campaigns. In this context, a measure of transmissibility is referred to as the reproduction number (R) and under the assumption of a well-mixed population is related to R_0 by $R = (1 - p)R_0$.

As the epidemic progresses, the effective susceptible population is depleted and the reproduction number decays. Significant reductions in the reproduction number can result from the use of public health interventions that can reduce the transmission rate over time (e.g., quarantine of suspected cases, isolation of infectious individuals) and the effective susceptible population (e.g., vaccination, prophylaxis with antiviral medications). The goal of public health interventions is to reduce the reproduction number quickly, at minimum cost.

Estimating the reproduction number directly from contact-tracing data (e.g., [4]) is not always feasible, especially for rapidly disseminating diseases such as influenza. Estimation of the reproduction number by fitting epidemic models to the cumulative tally of cases is probably the most widespread approach (examples of recent work include [5, 6]). A common approach to estimating the reproduction number is first to fit the initial exponential growth rate characteristic of many rapidly disseminating infectious diseases in human populations. This estimate of the initial exponential growth rate can be used to approximate the reproduction number through a formula derived from the linearization of the deterministic epidemic model (e.g., [2, 7, 8, 9]). Other recent methods for estimating the reproduction number include a removal method based on the chain binomial model [10] and approaches that estimate a time-varying reproduction number as the epidemic develops [11, 12, 20]. A recent review on the estimation of the reproduction number from epidemiological data is given by Heffernan et al. (2005) [13].

At the beginning of a developing epidemic, public health officials must make decisions about the types and intensity of interventions necessary for epidemic control. Such decisions can be guided by estimates of the reproduction number and the variance (uncertainty) of these estimates. A relevant question is how sensitive these estimates are to the number of days of epidemic curve data used in the estimation process, both in terms of bias and standard errors. For example, certain model assumptions may hold during the initial epidemic phase but not for the whole epidemic period. Therefore, sensitivity analyses to assess the robustness of estimates of the reproduction number should be part of any contingency plan. This paper is an extension of previous work [6] to address these questions in the context of the Spanish flu pandemic in Geneva, Switzerland using a compartmental epidemic model (Figure 1).

2. MATERIALS AND METHODS.

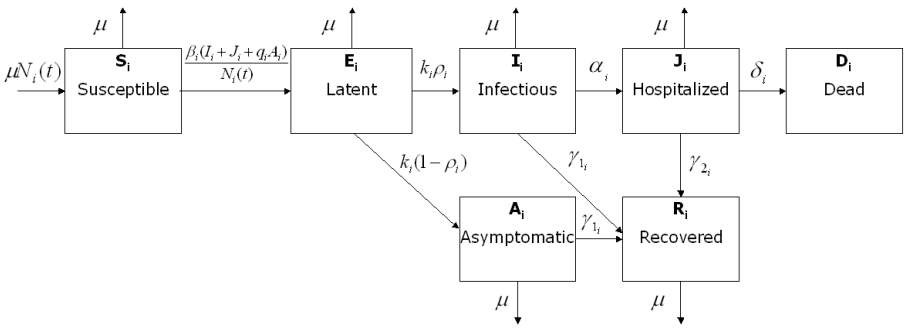

FIGURE 1. Flow chart of the state progression of individuals among the different epidemiological classes as modeled by Equations (1).

2.1. **Historical background.** The 1918-1919 influenza pandemic known as the "Spanish flu" was caused by the influenza virus A (H1N1). In the Canton of Geneva, Switzerland, the Spanish flu affected more than 50% of the population, and the mortality rate was highest in males in the age group 21-40 years [14]. The Canton of Geneva is located in the western side of Switzerland, surrounded in its majority by France, and covers an area of 282 km^2. In Geneva, the spring wave occurred from July through September 1918 while the fall wave spanned October through December 1918 and was deadlier than the spring wave. Control measures included school and church closures, prohibition of public events and visits to hospitals, and mandatory spraying of disinfectants on the streets. There is no evidence of the effectiveness of the control measures, because disruptions in the sanitary, medical, private, and public sectors were common [14, 15]. This was reflected in the social climate of insecurity, fear, and doubts among the population about the effectiveness of the control strategies. Other modern interventions, including influenza vaccines and antiviral medications were not yet available.

2.2. **Epidemic model.** We use a compartmental epidemic model (Figure 1) previously developed for studying the transmissibility and the effect of hypothetical interventions for the 1918 influenza pandemic in Geneva, Switzerland [6]. Individuals are classified as susceptible (S_i), exposed (E_i), clinically ill and infectious (I_i), asymptomatic and partially infectious (A_i), hospitalized and reported (J_i), recovered (R_i), and dead (D_i) where $i = 1{,}2$ indexes the spring and fall pandemic waves, respectively. The population is considered to be completely susceptible to the spring wave of infection. Individuals that recover during the spring wave are assumed protected to the fall wave [16], and the numbers of susceptible, recovered, and dead individuals at the end of the spring wave are set to be the corresponding initial conditions to model the fall wave. Susceptible individuals in contact with the virus progress to the latent class at the rate $\beta_i(I_i(t) + J_i(t) + q_i A_i(t))/N_i(t)$, where β_i is the transmission rate for wave i, and $0 < q_i < 1$ is a reduction factor in the transmissibility of the asymptomatic class (A_i).

Because there is no evidence of the effectiveness of interventions and because disruptions in the sanitary and medical sectors were common [15], hospitalized

individuals (J_i) are assumed infectious. Overworked doctors had to impose restrictions on reception hours, and many of them died from influenza [15]. As a result, announcements were issued in newspapers to call for volunteers to help in overburdened hospitals [15].

The total population size at time t for wave i is given by $N_i(t) = S_i(t) + E_i(t) + I_i(t) + A_i(t) + J_i(t) + R_i(t)$. This formulation allows for the total population size to change over time as a result of influenza-related deaths and demographic changes.

We assume the migration rate of people in and out of the population, including the birth and death rates, to have common rate μ. Assuming homogeneous mixing of the population, the fraction $(I_i(t) + J_i(t) + q_i A_i(t))/N_i(t)$ is the probability that a random contact would be with an infectious individual. A proportion $0 < \rho_i < 1$ of latent individuals progresses to the clinically infectious class (I_i) at the rate k_i while the rest $(1 - \rho_i)$ progress to the asymptomatic partially infectious class (A_i) at the same rate k_i. Asymptomatic cases progress to the recovered class at the rate γ_{1_i}. Clinically infectious individuals (class I_i) are hospitalized (reported) at the rate α_i or recover without being diagnosed (e.g., mild infections, hospital refusals [15]) at the rate γ_{1_i}. Hospitalized individuals (reported) recover at the same rate as individuals in the I_i class and hence $\gamma_{2_i} = 1/(1/\gamma_{1_i} - 1/\alpha_i)$ to account for the residence time in the J_i class or die at rate δ_i. The mortality rates were adjusted according to the case fatality proportion (CFP) such that $\delta_i = \frac{CFP}{1-CFP}(\mu + \gamma_{2_i})$ (see Figure 1).

The transmission process (for each influenza wave) can be modeled using the system of nonlinear differential equations:

$$\begin{cases} \dot{S}_i(t) &= \mu N_i(t) - \beta_i S_i(t)(I_i(t) + J_i(t) + q_i A_i(t))/N_i(t) - \mu S_i(t) \\ \dot{E}_i(t) &= \beta_i S_i(t)(I_i(t) + J_i(t) + q_i A_i(t))/N_i(t) - (k_i + \mu)E_i(t) \\ \dot{A}_i(t) &= k_i(1 - \rho_i)E_i - (\gamma_{1_i} + \mu)A_i(t) \\ \dot{I}_i(t) &= k_i \rho_i E_i(t) - (\alpha_i + \gamma_{1_i} + \mu)I_i(t) \\ \dot{J}_i(t) &= \alpha_i I_i(t) - (\gamma_{2_i} + \delta_i + \mu)J_i(t) \\ \dot{R}_i(t) &= \gamma_{1_i}(A_i(t) + I_i(t)) + \gamma_{2_i} J_i(t) - \mu R_i(t) \\ \dot{D}_i(t) &= \delta_i J_i(t) \\ \dot{C}_i(t) &= \alpha_i I_i(t) \end{cases} \tag{1}$$

where the index $i = 1, 2$ denotes the spring and fall waves of infection, respectively. The dot denotes the time derivatives. The cumulative number of hospital notifications, our observed data, is given by $C_i(t)$.

2.3. **The reproduction number and clinical reporting proportion.** An expression of the reproduction number for the above epidemic model can be obtained using standard methods in mathematical epidemiology [6]. The reproduction number is given by:

$$R_i = \frac{\beta_i k_i}{k_i + \mu}\left\{\rho_i\left(\frac{1}{\gamma_{1_i} + \alpha_i + \mu} + \frac{\alpha_i}{(\gamma_{1_i} + \alpha_i + \mu)(\gamma_{2_i} + \delta_i + \mu)}\right) + (1-\rho_i)\left(\frac{q_i}{\gamma_{1_i} + \mu}\right)\right\} \tag{2}$$

where $i = 1, 2$ indicates the reproduction numbers for the spring and fall influenza waves of infection, respectively. Notice that R_1 can be considered to be a basic reproduction number because the population was mostly naive to the influenza virus at the beginning of the pandemic. Also, it can be seen from the above expression that the reproduction number is the sum of the contributions to infection from three

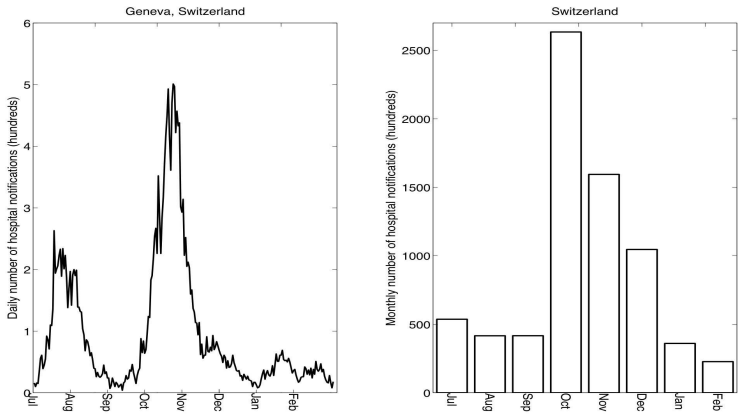

FIGURE 2. Daily number of hospital notifications in Geneva, Switzerland during the 1918-1919 influenza pandemic [14] and the monthly numbers of notifications for all of Switzerland for the same period [28, 29].

different classes of infectious individuals: infectious but not hospitalized, infectious and hospitalized, and those asymptomatic but partially infectious. Estimates of R_i are obtained by substituting the model parameter estimates into (2).

The clinical reporting proportion is the fraction of clinical cases, O_i, that are reported in hospitals over the total number of clinical cases which can be estimated by:

$$O_i = \frac{\alpha_i}{\alpha_i + \gamma_{1_i} + \mu} \tag{3}$$

2.4. **Data sources.** In 1917, life expectancy in the Canton of Geneva was about 60 years [17] with a population size of 174 673, which is about 42% of today's population (Dubois J. E-mail communication. Office Cantonal de la Statistique-Genève. July 11, 2005). Daily epidemic data for the Canton of Geneva was obtained from the mandatory notifications registry in Switzerland during the period 01 July 1918 to 28 February 1919 (Figure 2). While misdiagnosis may be common during seasonal epidemics of influenza because of the limited reliability of clinical diagnosis (nonspecific symptoms), clinical diagnosis for the case of pandemic influenza should have been more reliable because of the severity of symptoms, particularly during the fall wave [15]. We make the simplistic assumption that errors in the data occur at random and are as likely to overstate as to understate the true number of hospitalized flu cases.

The overall case fatality of the Spanish flu in Geneva was 4.2% [14]. It is well documented that the case fatality of the fall wave was much higher than that of the spring wave [14]. Since we do not have the separate case fatality proportions for the spring and fall waves of infection of the 1918 influenza pandemic in Geneva, we used the case fatality for the spring (0.7%) and fall (3.25%) waves of the 1918 influenza pandemic in UK as reported by Gani et al. [18].

2.5. **Parameter estimation.** The migration, birth, and natural death rates were fixed to a common value $\hat{\mu} = 1/(60 * 365)$ days^{-1}, and the latent period was fixed

to $1/\hat{k_1} = 1/\hat{k_2} = 1.9$ days as in previous studies [19]. For each of the two influenza waves, we estimated seven parameters: The transmission rate β_i, the recovery rate γ_i, the diagnostic rate α_i, the relative infectiousness of asymptomatic cases q_i, the proportion of clinical cases ρ_i, and the initial numbers of exposed $E_i(0)$ and infectious $I_i(0)$ individuals using least squares fitting of $C(t, \mathbf{\Theta_i})$ in Model (1) ($\mathbf{\Theta_i}$ is the vector of fitting parameters for wave i) to the initial phase of the cumulative number of hospital notifications of influenza. With regard to other initial conditions, at the beginning of the spring wave, the initial number of diagnosed individuals ($J_1(0)$) was set according to the first number of reported cases in the time-series data, and the rest of the population was assumed to be completely susceptible to pandemic influenza. All other initial conditions for other epidemiological states were initialized to zero. The numbers of susceptible, recovered, and dead individuals at the end of the spring wave are set to be the corresponding initial conditions to model the fall wave. Similarly, for the fall wave the initial number of diagnosed individuals was set according to the first number of reported cases from the epidemic data of the fall wave. The advantage of using the cumulative over the daily number of new cases for estimation is that the former smoothes out known reporting delays on weekends and national holidays. For the least squares fitting procedure, we used the Levenberg-Marquardt method with line-search implemented in MATLAB (The Mathworks, Inc.) by the routine `lsqcurvefit` in the optimization toolbox. For notation purposes, $\hat{R}_i^{(t)}$ and $\hat{O}_i^{(t)}$ denote estimates of the reproduction number and the clinical reporting proportion, respectively, for the i-th pandemic wave and using the first t epidemic days of the cumulative number of influenza notifications. A summary of parameter definitions and baseline values are given in Table 1.

2.6. **Parameter uncertainty.** Confidence intervals for the model parameter estimates, the reproduction number, and the clinical reporting proportion were constructed via a simulation study to generate sets of realizations of the best-fit curve $C(t)$ using the parametric bootstrap [21]. Each realization of the cumulative number of case notifications $C_j(t)$ ($j = 1, 2, \ldots, m$) is generated as follows: for each observation $C(t)$ for $t = 2, 3, \ldots, n$ days generate a new observation $C_j^{'}(t)$ for $t \geq 2$ ($C_j^{'}(1) = C(1)$) that is sampled from a Poisson distribution with mean: $C(t) - C(t-1)$ (the daily increment in $C(t)$ from day $t-1$ today t). Sampling from a Poisson distribution where the mean equals the variance is appropriate since we only have information about the number of cases over the course of the epidemic. Knowing more information about the variability in the number of cases would allow us to model the variance as well using, for example, the negative binomial distribution. The corresponding realization of the cumulative number of influenza notifications is given by $C_j(t) = \sum_{h=1}^{t} C_j^{'}(t)$ where $t = 1, 2, 3, \ldots, n$. The reproduction number was then estimated from each of 1000 simulated epidemic curves. The distribution of estimated reproduction numbers can be used to construct 95% confidence intervals. Fitting a complex model (with 7 fitting parameters in our case) comes at the cost of increased variance for these estimates. The worst case would occur when the model parameters cannot be uniquely determined from the data leading to unbounded variances of the estimated parameters. This simulation study allowed us to explore the identifiability of model parameters. Lack of identifiability can be recognized when large perturbations in the model parameters generate small changes in the model output [22]. Our results indicate that our parameter estimates are stable to perturbations around the model output.

TABLE 1. Parameter definitions and baseline estimates obtained from the initial epidemic phase comprising the first 10 epidemic days of the corresponding wave using system (1) during the 1918 influenza pandemic in Geneva, Switzerland. The standard deviation of the model parameter estimates were obtained from our simulation study using the parametric bootstrap that consisted of 1000 realizations as explained in the parameter estimation section.

			Spring Wave		Fall Wave	
Parameter	Definition	Source	Estimate	S. D.	Estimate	S. D.
β	Transmission rate (days^{-1})	Estimated	2.79	0.30	8.59	0.37
γ_1	Recovery rate (days^{-1})	Estimated	0.55	0.03	0.58	0.03
α	Diagnostic rate (days^{-1})	Estimated	0.60	0.09	0.62	0.08
q	Relative infectiousness of the asymptomatic class	Estimated	0.04	0.07	0.01	0.04
ρ	Proportion of clinical infections ([0,1])	Estimated	0.57	0.09	0.39	0.04
k	Rate of progression to infectious (days^{-1})	[19]	1/1.9	–	1/1.9	–
μ	Birth and natural death rate (days^{-1})	[17]	$1/(60*365)$	–	$1/(60*365)$	–
$E(0)$	Initial number of exposed individuals	Estimated	86.44	17.92	101.42	5.91
$I(0)$	Initial number of infectious individuals	Estimated	36.36	7.3	60.72	3.62

3. **RESULTS.** We analyzed the first two pandemic waves of the Spanish flu pandemic in Geneva, Switzerland. Figure 2 shows the daily number of hospital notifications of influenza along with the monthly numbers of notifications in all of Switzerland for comparison.

3.1. **Spring wave.** The transmissibility of the infection for the spring influenza wave was significantly lower than that of the fall wave [6]. For the spring herald pandemic wave in Geneva, our estimates of the reproduction number and the clinical reporting proportion using the first 10 or 15 epidemic days of the epidemic growth phase are consistent with estimates obtained using the epidemic data of the whole pandemic wave (Table 2, Figure 3). In fact, the confidence intervals for $\hat{R}_1^{10}$, $\hat{R}_1^{15}$, and $\hat{R}_1^{all}$ overlap with each other (Table 2). Specifically, we estimated a reproduction number of $\hat{R}_1^{10} = 1.57$ (95% CI: 1.45, 1.70) and $\hat{R}_1^{15} = 1.42$ (1.37, 1.48) from the initial epidemic phase comprising the first 10 and 15 epidemic days, respectively. Compare these values to the estimate of $\hat{R}_1^{all} = 1.49$ (1.45, 1.53) obtained by fitting the model to the whole epidemic wave [6]. Figure 4 shows the distribution of estimated model parameters using a parametric bootstrap for our two observational windows are consistent with one another. Regarding the clinical reporting proportion (see Equation 3), the estimate obtained by fitting the model to the whole spring wave was found to be only slightly higher than estimates obtained from the initial epidemic phase (Table 2). Moreover, the resulting distributions of the clinical reporting proportion from the parametric bootstrap are skewed to the right (Figure 3).

TABLE 2. Estimates of the reproduction number and the clinical reporting percentage using increasing amounts of epidemic data of the initial phase of the spring and fall waves of the Spanish Flu Pandemic in Geneva, Switzerland.

	spring wave				fall wave			
Epidemic days	$\hat{R}_1$	$\hat{R}_1$ 95% CI	$\hat{O}_1$	$\hat{O}_1$ 95% CI	$\hat{R}_2$	$\hat{R}_2$ 95% CI	$\hat{O}_2$	$\hat{O}_2$ 95% CI
10 days	1.57	(1.45,1.70)	51.6	(46.7, 56.6)	3.10	(2.81, 3.39)	51.4	(46.3, 56.5)
15 days	1.42	(1.37, 1.48)	51.6	(48.0, 55.1)	2.72	(2.59, 2.86)	51.7	(47.6, 55.8)
20 days	n.a.	n.a.	n.a.	n.a.	3.29	(2.97, 3.62)	60.3	(50.4, 70.2)
25 days	n.a.	n.a.	n.a.	n.a.	3.97	(3.40, 4.54)	86.2	(78.3, 94.1)
30 days	n.a.	n.a.	n.a.	n.a.	3.62	(3.40, 3.84)	83.0	(79.9, 86.1)
35 days	n.a.	n.a.	n.a.	n.a.	3.75	(3.55, 3.95)	84.4	(80.7, 88.0)
All days	1.49	(1.45, 1.53)	59.7	(55.7, 63.7)	3.75	(3.57, 3.93)	83.0	(79.0, 87.0)

n.a., not applicable. Confidence intervals (CI) were estimated from a simulation study using the parametric bootstrap that consisted of 1000 realizations as explained in the parameter uncertainty section.

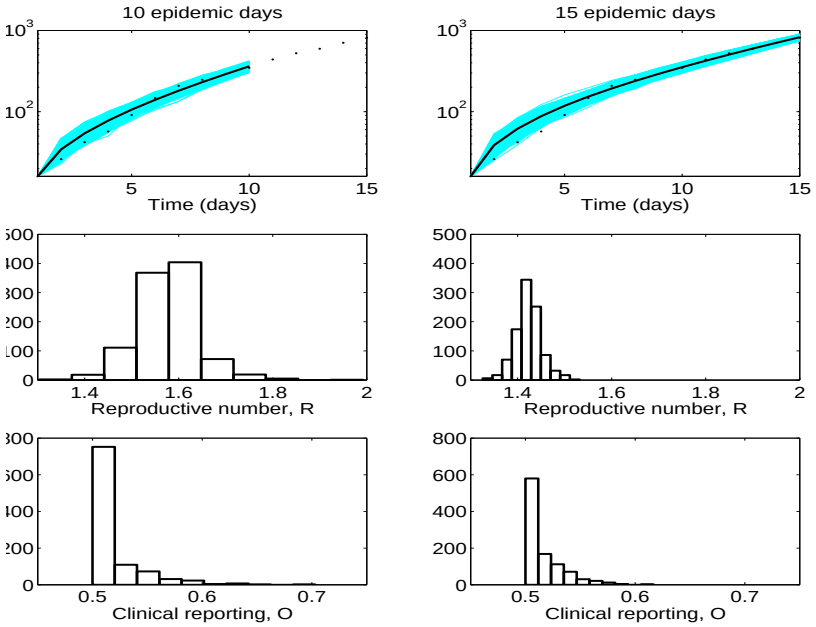

FIGURE 3. Model fits (log-lin scale) and the resulting distributions of the reproduction number and the proportion of the clinical reporting obtained after fitting the epidemic model to the initial phase of the spring wave using 10 and 15 epidemic days of the Spanish flu pandemic in Geneva, Switzerland. (top) The data are the dots, the solid line is the model best fit, and the lighter blue bands are 1000 realizations of the model fit to the data obtained through parametric bootstrapping as explained in the text. (middle) The distribution of the reproduction number and (bottom) the distribution of the clinical reporting proportion obtained from the simulation study with 1000 realizations.

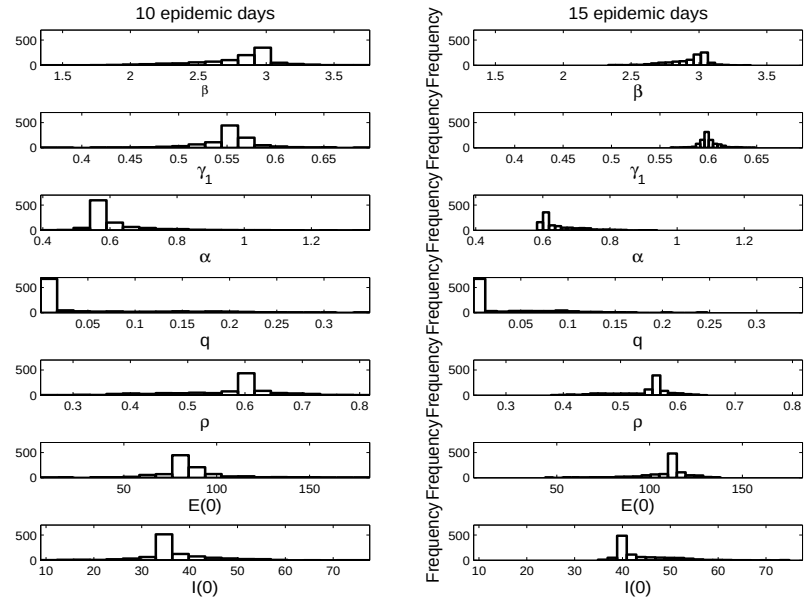

FIGURE 4. Histograms of the model parameter estimates for the spring wave of the Spanish flu pandemic in Geneva, Switzerland using 10 and 15 epidemic days of the initial phase of the pandemic wave used in the estimation procedure.

3.2. **Fall wave.** The high transmissibility of the fall pandemic wave can be observed from the start of the pandemic wave. The estimate of the reproduction number obtained from the initial rise of the cumulative number of hospital notifications comprising 10 epidemic days for the most deadly, fall wave [14] was 3.10 (95% CI: 2.81, 3.39). The reproduction number for the fall wave remained fairly stable with an increasing observational window used in the estimation (Table 2, Figure 5), albeit a significantly larger reproduction number of 3.29 (2.97, 3.62) was observed when fitting the model to the initial 20 epidemic days of data. Similarly, a significant increase in the clinical reporting proportion was observed after fitting the model to the initial 25 days of epidemic data, changing from $\hat{O}_2^{20} = 60.3$ (50.4, 70.2) to $\hat{O}_2^{25} = 86.2$ (78.3, 94.1). We found that the moderate variation in the reproduction number and the clinical reporting proportion can be explained as a result of a moderate systematic deviation of the model best fit to the epidemic data around epidemic day 9 (Figure 5). Such deviation of the model to the data becomes more significant as the epidemic model is fit to more than 20 days of epidemic data, as indicated by the snake shape of the residual plot (Figure 6), which can explain our larger estimate of the reproduction number when using the entire epidemic curve. A similar observation has been made for the case of the Spanish flu pandemic in the city of San Francisco, California [20]. The same significant variations can be observed in the model parameter estimates (Figure 7). In contrast to the spring pandemic wave, the distributions of the clinical reporting proportion for the fall

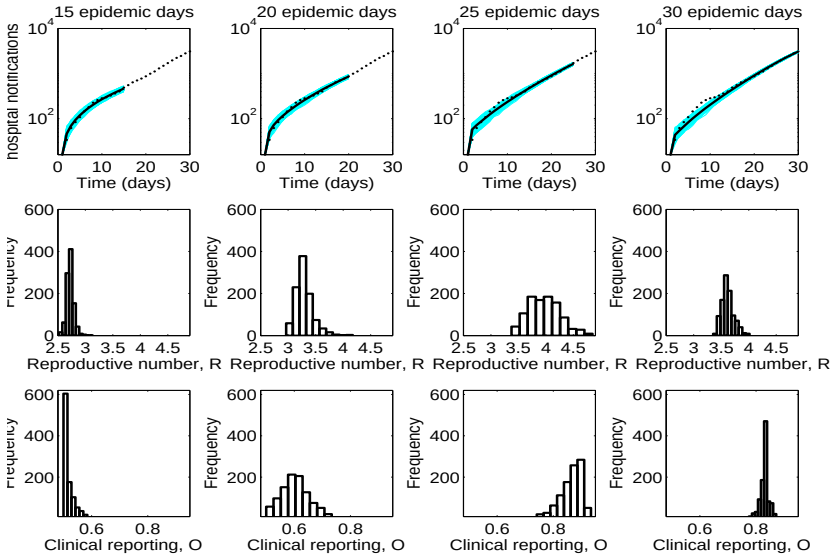

FIGURE 5. Model fits (log-lin scale) and the resulting distributions of the reproduction number and the proportion of the clinical reporting obtained after fitting the epidemic model to the initial phase of the second (fall) wave using 15, 20, 25, and 30 epidemic days of the Spanish Flu Pandemic in Geneva, Switzerland. (top) The data are the dots, the solid line is the model best fit, and the lighter blue bands are 1000 realizations of the model fit to the data obtained through parametric bootstrapping as explained in the text. (middle) The distribution of the reproduction number and (bottom) the distribution of the clinical reporting proportion obtained from the simulation study with 1000 realizations.

pandemic wave were found to be more symmetric except for the distribution using fifteen epidemic days (Figure 5).

4. **DISCUSSION.** Uncertainty bounds associated to (point) estimates of the reproduction number are critical in the decision-making process regarding the types and intensity of interventions necessary for epidemic control. In general the amount of uncertainty associated to estimates of the reproduction number will depend on the methods used, the amount and temporal scale of the data, and the epidemiology of the disease (e.g, length of the incubation and infectious periods, means of transmission). We used a parametric bootstrap approach to generate estimates of the variance of the reproduction number. The reproduction number was estimated using least-square fitting techniques from the initial phase of the first two waves of the Spanish flu pandemic in Geneva, Switzerland. We estimated the reproduction number with relatively small uncertainty after fitting the epidemic model to the initial epidemic take-off comprising 10 and 15 epidemic days of data for the spring wave, and 10, 15, 20, 25, 30, 35 epidemic days for the fall more severe wave. We were able to explain a moderate variation in the reproduction number of the fall

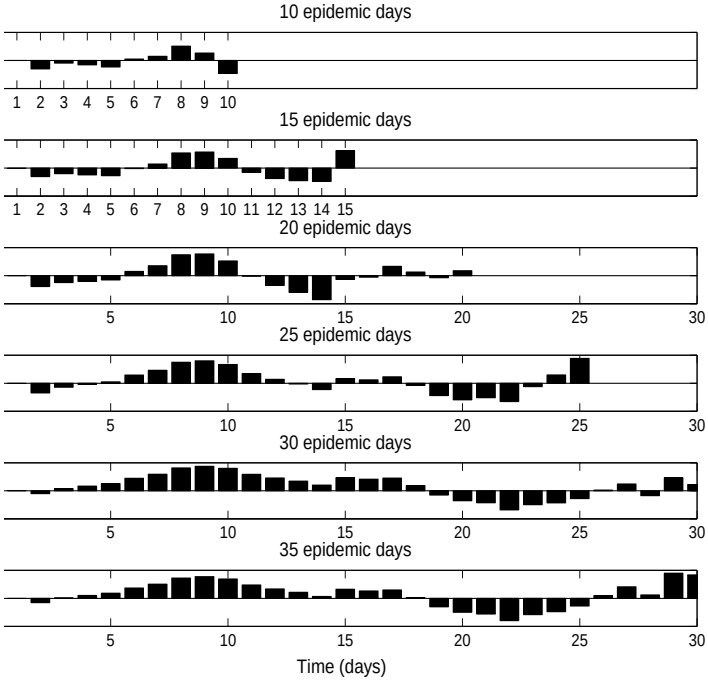

FIGURE 6. Scaled residual plots of the epidemic model best fit to the cumulative number of influenza notifications of the fall wave using 10, 15, 20, 25, 30, and 35 days of epidemic data. A moderate systematic deviation is observed around epidemic day 9. The residuals are scaled by the standard deviation of the distribution of residuals. These residuals are within $\pm$ 2 standard deviations (95% confidence level).

wave as a result of a moderate systematic deviation of the model to the data around epidemic day 9 (Figures 5 and 6). This can be the result of our modeling approach based on an autonomous system where the transmission rate, the diagnostic rate, and other epidemiological parameters are assumed to remain constant throughout the epidemic.

Our estimates of the reproduction number agree with previous estimates for the Spanish flu pandemic in other regions of the world [19, 18, 6, 20, 23, 24]. For example, the reproduction number has been estimated to be in the range 2-3 in US cities using mortality data for the fall wave of the Spanish flu pandemic in the largest cities in the United States. Using influenza morbidity data from the city of San Francisco, California, Chowell et al. [20] estimated R in the range 2-3 using four different estimation approaches. Gani et al. [18] estimated a basic reproduction number of 2 for the spring wave and 1.55 for the fall wave. Also, Rvachev and Longini [25] estimated a reproduction number of 1.9 for the influenza H3N2 pandemic of 1968 in Hong Kong. For comparison with seasonal flu, analyses of the A/H2N2 and A/H3N2 virus circulation have provided estimates of the reproduction

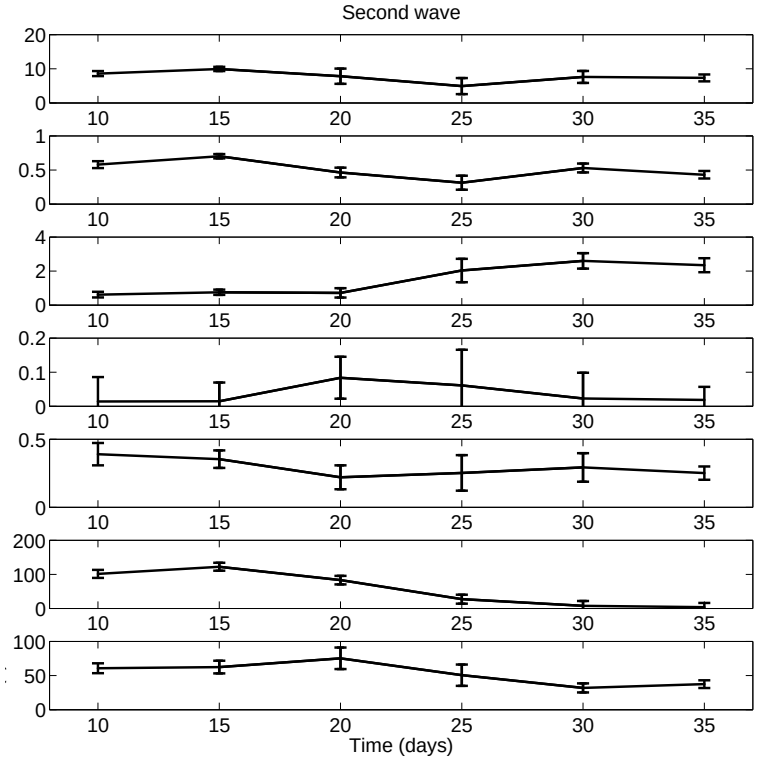

FIGURE 7. Model parameter estimates and their corresponding 95% confidence intervals (bars) for the fall wave of the Spanish flu pandemic in Geneva, Switzerland as a function of the number of days of the initial phase of the pandemic wave used in the estimation procedure.

number in the range from 1.43 in England and Wales during 1971-1972 to 2.61 in Greater London during 1960-1961 [26], and approximately 1.5 during the 1984-85 influenza season in France [27].

Our epidemic model maintains the minimal complexity necessary to estimate the reproduction number and the clinical reporting proportion using hospital notification data. Despite the simplifying assumptions on spatial homogeneity and the random mixing assumption (each individual has the same probability of infecting or being infected by any of the other individuals in the population), our epidemic model was able to fit well the epidemic course of the first two waves of the Spanish flu pandemic in Geneva, Switzerland. While estimates of the reproduction number remained fairly stable with respect to the assumption on social mixing, the individual model parameters were sensitive to the number of epidemic days used in the estimation (Figure 7). Therefore, the interpretation of model parameter estimates need to be addressed with caution because biases in model parameters could be introduced due to simplifying assumptions.

Here we fit the model separately to the spring and fall waves of infection, but the underlying processes that led to the multiple peaks of infection during 1918 flu pandemic were not elucidated. Possible contributing mechanisms to this phenomena include changes in population behavior and/or in the pathogenicity of the influenza virus. This is an open question for future research.

In summary, for the 1918 pandemic scenario in Geneva, Switzerland, estimates of the reproduction number obtained through least-square fitting of the model to increasing observational window during the initial epidemic rise were found to be in good agreement with each other. Nevertheless, control measures could generate significant changes in the effective proportion of susceptible individuals and the per capita transmission rate leading to a reduction in the reproduction number. The estimation of the reproduction number through least squares fitting from the initial growth phase of the epidemic curve has the potential of informing early into an epidemic the intensity of interventions necessary to contain a rising epidemic.

5. **Acknowledgments.** This work was carried out under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy at Los Alamos National Laboratory under Contract No. DE-AC52-06NA25396 and was partially supported by the NNSA's Laboratory Directed Research and Development Program.

REFERENCES

[1] N.P. Johnson and J. Mueller. Updating the accounts: Global mortality of the 1918-1920 "Spanish" influenza pandemic. *Bull. Hist. Med.* 76 (2002) 105-15.

[2] R.M Anderson and R.M. May. *Infectious Diseases of Humans*. Oxford, UK: Oxford University Press, 1991.

[3] F. Brauer and C. Castillo-Chavez. *Mathematical Models in Population Biology and Epidemiology*. New York, NY: Springer-Verlag, 2000.

[4] J.J. Potterat, H. Zimmerman-Rogers, S.Q. Muth, et al. Chlamydia transmission: concurrency, reproduction number, and the epidemic trajectory. *American Journal of Epidemiology* 150 (1999) 1331-1339.

[5] S. Riley, C. Fraser, C.A. Donnelly, et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science* 300 (2003) 1961-1966.

[6] G. Chowell, C.E. Ammon, N.W. Hengartner, J.M. Hyman. Transmission dynamics of the Great influenza pandemic of 1918 in Geneva, Switzerland: Assessing the effects of hypothetical interventions. *Journal of Theoretical Biology* 241 (2006) 193-204.

[7] M.A. Nowak, A.L. Lloyd, G.M. Vasquez, et al. Viral dynamics of primary viremia and antiretroviral therapy in simian immunodeficiency virus infection. *Journal of Virology* 71 (1997) 7518-7525.

[8] A.L. Lloyd AL. The dependence of viral parameter estimates on the assumed viral life cycle: limitations of studies of viral load data. *Proceedings of the Royal Society B* 268 (2001) 847-854.

[9] M. Lipsitch, T. Cohen, B. Cooper, J.M. Robins, S. Ma, L. James, G. Gopalakrishna, S.K. Chew, C.C. Tan, M.H. Samore, D. Fisman, M. Murray. Transmission dynamics and control of severe acute respiratory syndrome. *Science* 300 (2003) 1966-1970.

[10] M.J. Ferrari, O.N. Bjornstad, A.P. Dobson. Estimation and inference of R-0 of an infectious pathogen by a removal method. *Mathematical Biosciences* 198 (2005) 14-26.

[11] D.T Haydon, M. Chase-Topping, D.J. Shaw, et al. The construction and analysis of epidemic trees with reference to the 2001 UK foot-and-mouth outbreak. *Proceedings of the Royal Society B* 270 (2003) 121-127.

[12] J. Wallinga, P. Teunis. Different epidemic curves for Severe Acute Respiratory Syndrome reveal similar impacts of control measures. *American Journal of Epidemiology* 160 (2004) 509-516.

[13] J.M. Heffernan, R.J. Smith, L.M. Wahl. Perspectives on the basic reproduction ratio. *Journal of the Royal Society Interface* 2 (2005) 281-293.

[14] C.E. Ammon. Spanish flu epidemic in 1918 in Geneva, Switzerland. *Euro Surveillance* 7 (2002) 190-2.
[15] C.E. Ammon. The 1918 Spanish flu epidemic in Geneva, Switzerland. *International Congress Series* 1219 (2001) 163-168.
[16] A.H. Reid, J.K. Taubenberger, T.G. Fanning. The 1918 Spanish influenza: integrating history and biology. *Microbes Infect* 3 (2001) 81-7.
[17] J.M. Robine and F. Paccaud. Nonagenarians and centenarians in Switzerland, 1860-2001: a demographic analysis. *Journal of Epidemiology and Community Health* 59 (2005) 31-7.
[18] R. Gani, H. Hughes, D. Fleming, T. Griffin, J. Medlock, S. Leach. Potential impact of antiviral use during influenza pandemic. *Emerging Infectious Diseases* 11 (2005) 1355-62.
[19] C.E. Mills, J.M. Robins, M. Lipsitch. Transmissibility of 1918 pandemic influenza. *Nature* 432 (2004) 904–906.
[20] G. Chowell, H. Nishiura, L.M.A. Bettencourt. Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. *Journal of the Royal Society Interface* 4 (2007) 155-166.
[21] B. Efron, R. Tibshirani. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1 (1986) 54-75.
[22] G. Pillonetto, G. Sparacino, C. Cobelli. Numerical non-identifiability regions of the minimal model of glucose kinetics: superiority of Bayesian estimation. Mathematical Biosciences 184 (2003) 53-67.
[23] G. Sertsou, N. Wilson, M. Baker, P. Nelson, M.G. Roberts. Key transmission parameters of an institutional outbreak during the 1918 influenza pandemic estimated by mathematical modelling. Theor. Biol. Med. Model. 30 (2006) 3-38.
[24] E. Massad, M.N. Burattini, F.A. Coutinho, L.F. Lopez. The 1918 influenza A epidemic in the city of Sao Paulo, Brazil. Med. Hypotheses (2007) 68 442-5.
[25] L.A. Rvachev, I.M. Longini. A Mathematical model for the global spread of Influenza. *Mathematical Biosciences* 75 (1985) 3-22.
[26] I.M. Longini. The generalized discrete-time epidemic model with immunity: A synthesis. *Mathematical Biosciences* 82 (1986) 19-41.
[27] A. Flahault, S. Letrait, P. Blin, S. Hazout, J. Menares, J. Valleron. Modelling the 1985 influenza epidemic in France. *Statistics in Medicine* 7 (1988) 1147-1155.
[28] Statistiques des maladies transmissibles par mois, 1918-20. Bulletin de l'hygiène publique 1918-1919-1920, 1920.
[29] Relevé des cas de maladies transmissibles signalés (1918-19-20). Bulletin de l'hygiène publique:3 pages, 1920.

Received on November 20, 2006. Accepted on February 2, 2007.

E-mail address: chowell@lanl.gov
E-mail address: catherine-ammon@bluewin.ch
E-mail address: nickh@lanl.gov
E-mail address: hyman@lanl.gov