

2008-1-30

Transmissibility of the Influenza Virus in the 1918 Pandemic

White, Laura Forsberg, Marcello Pagano. "Transmissibility of the Influenza Virus in the 1918 Pandemic" PLoS ONE3(1): e1498. (2008)

<https://hdl.handle.net/2144/3103>

"Downloaded from OpenBU. Boston University's institutional repository."

Transmissibility of the Influenza Virus in the 1918 Pandemic

Laura Forsberg White*, Marcello Pagano

Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, United States of America

Background. With a heightened increase in concern for an influenza pandemic we sought to better understand the 1918 Influenza pandemic, the most devastating epidemic of the previous century. **Methodology/Principal Findings.** We use data from several communities in Maryland, USA as well as two ships that experienced well-documented outbreaks of influenza in 1918. Using a likelihood-based method and a nonparametric method, we estimate the serial interval and reproductive number throughout the course of each outbreak. This analysis shows the basic reproductive number to be slightly lower in the Maryland communities (between 1.34 and 3.21) than for the enclosed populations on the ships ($R_0=4.97$, $SE=3.31$). Additionally the effective reproductive number declined to sub epidemic levels more quickly on the ships (within around 10 days) than in the communities (within 30–40 days). The mean serial interval for the ships was consistent (3.33, $SE=5.96$ and 3.81, $SE=3.69$), while the serial intervals in the communities varied substantially (between 2.83, $SE=0.53$ and 8.28, $SE=951.95$). **Conclusions/Significance.** These results illustrate the importance of considering the population dynamics when making statements about the epidemiological parameters of Influenza. The methods that we employ for estimation of the reproductive numbers and the serial interval can be easily replicated in other populations and with other diseases.

Citation: White LF, Pagano M (2008) Transmissibility of the Influenza Virus in the 1918 Pandemic. PLoS ONE 3(1): e1498. doi:10.1371/journal.pone.0001498

INTRODUCTION

The emergence of the highly pathogenic avian influenza strain H5N1 has raised concerns of an imminent influenza pandemic. Public health workers, government officials and disaster planners have an increasing interest in better understanding the potential impact of an influenza pandemic and possible strategies for containment. Crucial in this planning is an understanding of the basic epidemiology of the disease in various settings. This has led to a growing interest in the analysis and understanding of past epidemics, particularly that of 1918, the most virulent and deadly influenza epidemic of the 20th century. Mortality has been estimated at 50–100 million people worldwide as a result of influenza in the 1918 pandemic [1]. It is reasonable to suppose that by better understanding the transmission dynamics of the highly pathogenic virus in 1918, we can gain greater insight into the dynamics, and thus potential methods of control, for a future pandemic [2].

Important parameters for understanding disease transmission are the reproductive number and the serial interval [3]. The basic reproductive number is defined as the average number of secondary infections created from a primary infection in an entirely susceptible population [4, see also 5]. A more complex, but perhaps meaningful parameter is the effective reproductive number which defines the average number of secondary infections an infected will create at a given point in the epidemic. This parameter takes into account that not all contacts of an infected individual are with susceptible persons, as well as the impact of public health control measures. Control strategies are typically targeted to drive this number below one and maintain it there, as this will lead to eventual extinction of the epidemic. An example of this is herd immunity, or immunity to a disease that is incurred from a sufficiently large proportion of the population being immune to a disease. Modeling techniques are often used to determine the proportion of the population that should be vaccinated in order to have the reproductive number low enough to avoid outbreaks of disease [6].

The serial interval can be defined as the time interval between a primary case presenting with symptoms and its infectee developing symptoms [7,8]. Thus this quantity is completely observable. This is a mixture of the incubation period and the infectious period,

both of which are useful to understand, but difficult to measure. The SARS outbreak of 2003 had a relatively long serial interval, estimated to be between 8 and 10 days on average and following a Weibull distribution [9] making case isolation extremely effective in containing the epidemic.

Methods for the estimation of basic epidemiological parameters are still in development phase. [10] provides a thoughtful summary of methods for estimating the reproductive number. One particularly interesting and useful method has been previously described by [7] for estimating the daily reproductive number, R_t , or the average number of cases an infected individual on day t would cause. One interesting feature of this method is that for days where no cases are observed, the estimated effective reproductive number is zero. Another observation is that this method essentially estimates a curve for the effective reproductive number that traces the epidemic curve, lagged by the average serial interval length. This nonparametric method presupposes information on the serial interval distribution. This is typical as most methods for estimating the reproductive number rely on knowledge of the serial interval.

Few have described analytical methods for estimating the serial interval, making most methodologies dependent on contact tracing data, which is often difficult and expensive to attain. [11] describe a method to estimate the reproductive number that relies on limited contact tracing information but not a full estimate of the serial

Academic Editor: Matthew Baylis, University of Liverpool, United Kingdom

Received November 9, 2007; **Accepted** December 20, 2007; **Published** January 30, 2008

Copyright: © 2008 White, Pagano. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Work funded by a grant from the National Institutes of Health, R01 EB0006195.

Competing Interests: The authors have declared that no competing interests exist.

* **To whom correspondence should be addressed.** E-mail: lfwwhite@bu.edu

interval. [12] have recently described a method to estimate the serial interval and then used this estimate with the estimator proposed in [7] of the daily reproductive number and have applied their method to data from outbreaks of avian influenza in poultry farms in Europe.

Several researchers have studied the 1918 pandemic and estimated some of these key epidemiological parameters. Estimates have ranged from 2–3 for the basic reproductive number, R_0 , when using an SEIR model with a mean latent period of 1.9 days and infectious period of 4.1 days [13,14]. Using an exponential model and assuming the serial interval to be four days (somewhat based on the assumptions of [13]), [15] estimated R_0 to be 2.6–10.6 for confined settings (such as prison and ships) and 2.4–4.3 for community settings. The estimates for the mean latent and infectious periods come from [16] and were used again by [17] and [18]. It appears that the original estimates were derived from epidemic data, although their source is not well documented.

In what follows, we introduce new methodology for the estimation of both the daily reproductive number and the serial interval. We apply this method to data from two outbreaks on military ships in the 1918 influenza outbreak, as well as well-documented outbreaks in five Maryland communities. The results from this method are compared to that of [12]. The results illustrate the differences in infectious disease dynamics between outbreaks in a closed population and a dynamic community.

METHODS

Data

We analyze data from several well-documented influenza outbreaks in 1918. First we consider data from two troop ships that embarked in the late fall of 1918 [19]. The Medic reported two initial cases on November 11. Out of 989 passengers (156 crew members, 829 soldiers, 4 civilians) 313 became sick with influenza over a 40 day period (Attack Rate, AR, = 0.32), though most of the cases occurred within the first fourteen days. The Boonah left Durban and in five days, on November 29, reported the first three definitive cases of influenza. Those who collected the data note that there were likely some initial cases that were not identified. Out of 1095 on board (164 crew members and 931 troops), 470 cases were reported (AR = 0.43) in the 40 days of the epidemic.

The United States Public Health Service created special surveys of 18 localities during the pandemic [20]. Reported results from six communities in Maryland are derived from house-to-house surveys requesting the date of onset of influenza for all infected, and the sex and age of each case of pneumonia and influenza. A summary of these populations is provided in Table 1.

Statistical Methods

We describe a likelihood based methodology for estimating the reproductive number at each day in the epidemic as well as the serial interval. The method builds on that described by [21]. We

assume that the population is closed, that all cases are observed, and use daily case counts only (i.e. number of new cases each day).

Let $\mathbf{N} = \{N_0, N_1, N_2, \dots, N_T\}$ represent the daily cases counts of influenza for the T days of the epidemic and X_{ij} represent the number of cases that appear on day j that are infected by individuals that appeared sick on day i . Following is a representation of the disease transmission model in the population.

$$\begin{aligned} N_0 & \\ N_1 &= X_{01} \\ N_2 &= X_{02} + X_{12} \\ N_3 &= X_{03} + X_{13} + X_{23} \\ N_4 &= X_{04} + X_{14} + X_{24} + X_{34} \\ &\vdots \\ N_T &= \sum_{i=T-\min(T,k)}^{T-1} X_{iT} \end{aligned}$$

We assume that the total number of cases produced by those on day i , X_i , are Poisson distributed with parameter $N_i R_i$, where R_i is the reproductive number for cases on day i . We further assume that $\mathbf{X}_i = \{X_{i,i+1}, X_{i,i+2}, \dots, X_{i,i+k}\}$ follows a multinomial distribution with parameters X_i , \mathbf{p} , k , where $\mathbf{p} = \{p_1, p_2, \dots, p_k\}$ represent the distribution of the serial interval. Using these assumptions we can construct a likelihood function (see details in the Supplemental Information), which, when simplified, yields the following convenient form

$$L(R_i, \mathbf{p} | N) = \prod_{i=1}^T \frac{\exp(\mu_i) \mu_i^{N_i}}{N_i!},$$

where $\mu_i = R_i (\sum_{j=1}^k p_j N_{i-j})$ [21].

Maximization of this likelihood with respect to R_i and \mathbf{p} yields estimates of these parameters. To further simplify this process and create a more parsimonious model, we parameterize \mathbf{p} by allowing it to follow a traditional parametric form for a serial interval (for instance a Weibull, Gamma, Log Normal, or Exponential distribution). Then the p_j are functions of the parameters of the density (for instance in the case of the Gamma distribution, the p_j only depend on the shape and rate parameters of the Gamma).

Similarly R_i can be modeled parametrically as a function of time. One example of a reasonable model for this is the four parameter logistic curve [22–24] given by

$$R_i = a + \frac{b}{1 + \exp\{c(i-d)\}}.$$

Table 1. Demographic and survey information on the Maryland communities surveyed in the 1918 Influenza pandemic.

Community	1917 Population	Number Surveyed	Percent Surveyed	Number of Cases	Attack Rate
Baltimore	594,637	33,776	5.7	7,868	0.23
Cumberland	26,686	5,234	19.6	2,147	0.41
Frederick	11,225	2,420	21.6	777	0.32
Salisbury	6,690*	1,735	25.9*	796	0.46
Lonaconing	1,553*	1,840	-	1,093	0.59

doi:10.1371/journal.pone.0001498.t001

The parameters of this curve describe the initial height of the curve (approximately $a+b$), the point of inflection (d), the curvature over the inflection (c) and the final height of the curve (a). These parameters have biological meaning in this setting where the initial height corresponds to the values of R_i prior to intervention and significant depletion of the susceptible population. The inflection point and its steepness would describe the timing of intervention and the rapidity with which it impacts transmission. The final height would describe the ultimate value of R_i , which typically is less than one, indicating that disease transmission is in a sub epidemic state.

In our analysis, we also implement the method described by [12] (hereafter referred to as the Garske et al. method) and compare the results of the two methodologies. This method first estimates the generation time distribution using a likelihood based method. Then the effective reproductive number is estimated using the method described by [7] (hereafter referred to as the WT method). We fit the likelihood for both methods using a Nelder-Mead maximization procedure and use 576 starting values in order to ensure that we reach the global maximum. All analyses were done using R 2.4.1.

Both methods assume homogenous mixing in the population, no missing data (clearly violated with the data from the Maryland communities), that a primary case experiences symptom onset prior to any cases that it infects and a completely closed system where all cases are infected by a case that has been observed. In the case of the Maryland data, where only a sample of the total number of cases was surveyed, we can observe the efficacy and robustness of these methods with sample data. Certainly results should be interpreted with caution, however, as we will show, the results that are obtained are consistent with previous estimates for influenza.

Error Estimates and Residuals

Standard errors were calculated for the MLE method using a parametric bootstrap. One thousand epidemics were simulated using the parameter estimates and estimates were obtained from each of these simulated epidemics. The standard deviation of the 1000 estimates was used as the standard error estimates. We used the method described in [12] to estimate the standard error for their estimates, however our simulations based on their assumption of asymptotic normality yielded a large number of negative estimates for the parameters. It is possible that this is due to the non-independence in the data and lack of theoretical underpinnings for the method that they propose. These results make their standard error estimates infeasible to estimate in this case. Therefore we do not present standard error estimates for the results obtained using their methodology.

In order to determine the accuracy and relative merit of the estimates obtained from each methodology, we compute one-step-ahead residuals and implement a cross validation approach to analyze the generalizeability of the estimates obtained. The one-step-ahead residuals were calculated by first using the estimates from a particular location along with the data to predict the next days' number of cases, \tilde{N}_i as follows:

$$\tilde{N}_i = \sum_{j=1}^{\min(k,i)} p_j R_{i-j} N_{i-j}.$$

Each \tilde{N}_i is calculated using N_0, N_1, \dots, N_{i-1} . Then the one-step-ahead residuals are calculated as

$$e_i = \frac{(N_i - \tilde{N}_i)^2}{\tilde{N}_i}.$$

We present these residuals averaged over the T days observed.

Generalizeability of the results was studied using an ad hoc cross validation (CV) technique. This is done by using the estimates obtained from one location to calculate the one step ahead residuals for another location. Specifically we use the Boonah ship estimates to calculate residuals with the Medic data and then use the Medic estimates to calculate the residuals for the Boonah data. For the Maryland communities, we report the average of the residuals obtained using the estimates from one community to predict the epidemics in each of the other four communities, creating five CV estimates (one for each community).

RESULTS

Serial Interval Estimates

Table 2 gives the results for the serial interval distribution estimates. Notable in these results is the striking consistency in the estimates of the first moment, with the exception of Cumberland. The second moments vary much more, however. In general they tend to be much larger for the ships when using the Garske et al. method compared to the MLE method. For the communities, we observe that they are consistently around 10 for the Garske et al. method and vary much more for the MLE method. Also of interest in these results are the large error estimates, particularly for Cumberland, but also to a smaller extent for Frederick. This is perhaps indicative of the model not fitting the data as well, for instance the logistic model may not be the best fit in this scenario, or that the lack of census data on cases might be more problematic here.

Reproduction Number

In Table 3 and Figure 1, we present the results for estimation of the effective reproductive number. Evident in these results, is the large initial reproductive number for the Boonah ship. This is likely due to some of the missing data at the beginning of the epidemic and thus the model attributing the large number of cases that rapidly develop to the few individuals who were initially reported. The logistic model fits this as accurately as possible, but perhaps the important message is the qualitative result, indicating that initial transmission in this susceptible, non-quarantined population was very high and rapidly decreased as many became infected. The result is similar for Medic though the initial value is not high. We also note that the reproductive number dropped to sub epidemic levels rapidly (around 10 days for both ships).

In the Maryland communities the initial reproductive number tended to be slightly lower (ranging from 1.34 in Salisbury to 3.21

Table 2. Serial interval estimates for the MLE method and the Garske et al. method.

Location	MLE		Garske et al	
	$\hat{\mu}$ (SE)	$\hat{\sigma}^2$ (SE)	$\hat{\mu}$	$\hat{\sigma}^2$
<i>Ships</i>				
Boonah	3.81 (3.69)	1.25 (2.83)	4.38	19.64
Medic	3.33 (5.96)	11.35 (4.60)	3.88	18.04
<i>Communities</i>				
Baltimore	2.83 (0.53)	2.30 (1.28)	2.90	8.45
Cumberland	8.28 (951.95)	25.00 (6143.28)	3.61	10.71
Frederick	4.66 (10.68)	28.65 (157.78)	3.09	11.57
Salisbury	3.31 (1.94)	9.08 (14.80)	3.76	12.40
Lonaconing	4.02 (14.10)	3.25 (26.74)	3.99	12.69

doi:10.1371/journal.pone.0001498.t002

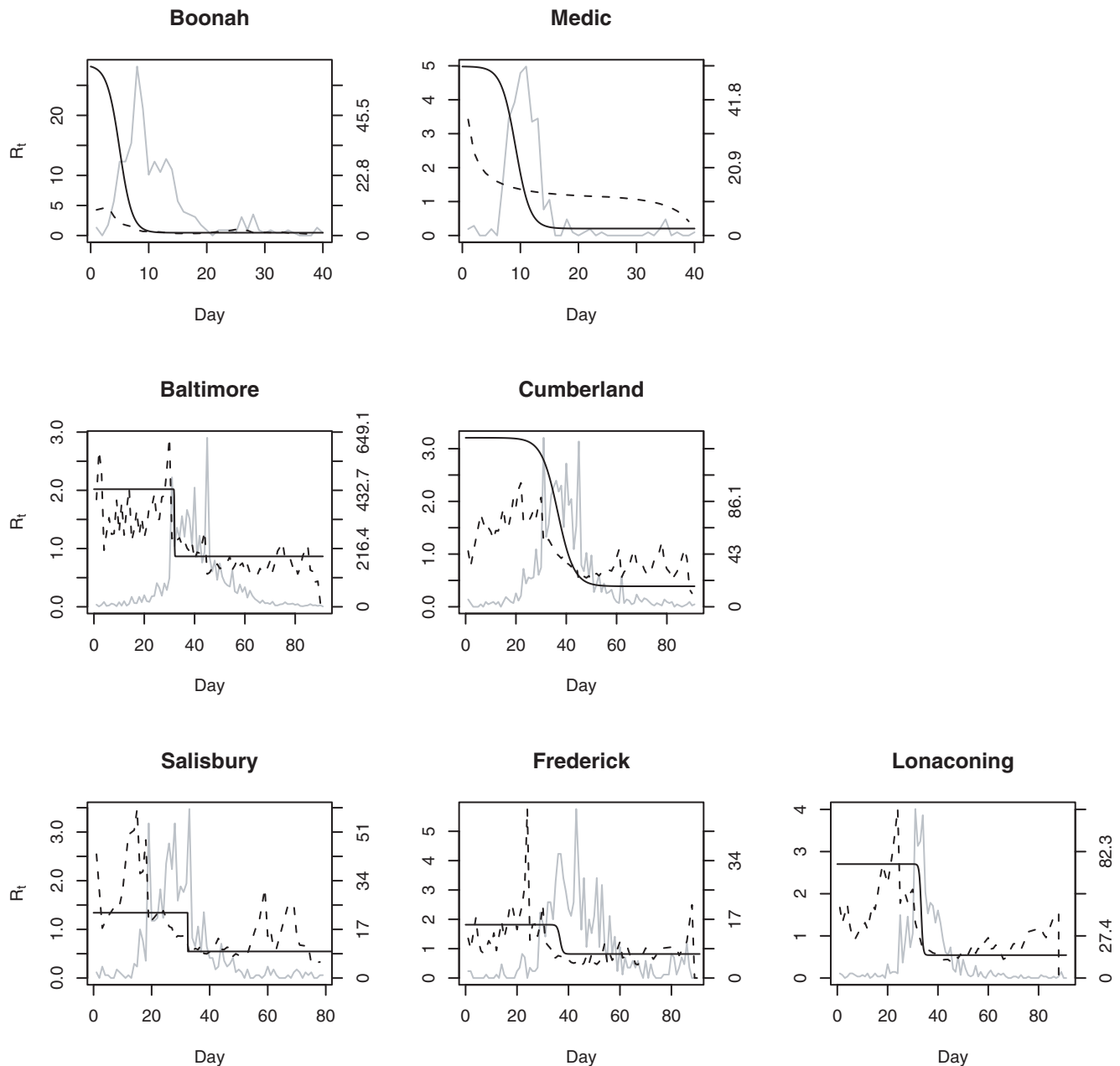


Figure 1. Estimated effective reproductive number for each location, using the MLE method (solid black line) and WT method (dashed line). The epidemic curve is shown in gray and its axis is on the right of the figure.
doi:10.1371/journal.pone.0001498.g001

in Cumberland). For the WT method, the initial values were also small, but if one considers the maximal values, these were much more varied (from 2.35 to 5.76), as shown in Table 2. In fact, we observed that the effective reproductive number peaked relatively late in the epidemic and assumed much higher values than that observed with the MLE method. Overall it took longer for the reproductive number to drop below one in these communities (typically between 30–40 days).

Generalizeability

In Table 4, we present the results of the residual analysis. We notice here that the Garske et al. method often does better than

the MLE method. It is important to point out that the WT method of fitting the effective reproductive model over fits the model and suffers from generalizeability. This method essentially traces the epidemic curve, lagged by the mean of the generation time distribution. Thus, according to the residuals, it appears that the WT method outperforms the MLE. However, considering the importance of external validation and reproducibility, the model suffers somewhat as evidenced by the CV measures. The exceptions to this are in the case of the Boonah where the CV measure is impacted by the large initial MLE estimate of the reproductive number and in Cumberland where it appears that either the parametric model chosen may not represent the best fit to the data or there were sensitivities to the survey data.

Table 3. Effective reproduction number estimates for the MLE method and the Wallinga and Tuenis (WT) method.

Location	Max WT \hat{R}_t (day)	\hat{R}_t			
		Day 1	Day 10	Day 30	Day 70
Ships					
<i>Boonah</i>					
MLE (SE)		27.71 (8.04)	0.74 (0.71)	0.47 (0.03)	
WT	4.74 (3)	4.27	0.68	0.53	
<i>Medic</i>					
MLE (SE)		4.98 (3.31)	1.83 (0.830)	0.21 (0.09)	
WT	3.42 (1)	3.42	1.36	1.06	
Communities					
<i>Baltimore</i>					
MLE (SE)		2.02 (0.12)	2.02 (0.12)	2.02 (0.12)	0.87 (0.12)
WT	2.90 (30)				
<i>Cumberland</i>					
MLE (SE)		3.21 (0.73)	3.21 (0.73)	2.88 (0.66)	0.39 (0.15)
WT	2.35 (22)	1.06	1.42	2.07	0.84
<i>Frederick</i>					
MLE (SE)		1.82 (0.14)	1.82 (0.14)	1.82 (0.14)	0.82 (0.04)
WT	5.76 (24)	1.35	1.11	2.49	1.02
<i>Salisbury</i>					
MLE (SE)		1.34 (0.18)	1.34 (0.18)	1.34 (0.18)	0.55 (0.13)
WT	3.47 (15)	2.55	1.79	0.86	0.87
<i>Lonaconing</i>					
MLE (SE)		2.70 (0.19)	2.70 (0.19)	2.70 (0.19)	0.54 (0.03)
WT	4.01 (24)	1.67	1.17	2.10	0.81

doi:10.1371/journal.pone.0001498.t003

DISCUSSION

We have presented results that are informative with regard to the dynamics of the 1918 influenza pandemic in different populations and provide insight into two methodologies for estimating basic epidemiological parameters. Both methods assume that the population is closed, there are no missing cases and no migration to or from the population. The second of these assumptions is clearly violated with the data from Maryland; however the results appear to be reasonably robust to this discrepancy, except in the case of Cumberland.

The purpose of this exercise determines to some extent which methodological approach we might favor. If the intent is to simply estimate the parameters for a specific epidemic and better understand what exactly was occurring in that setting, then the method presented by [12] (Garske et al.) appears to provide good fit. The caveat that we see in this method is that by estimating the effective reproductive number with the methodology of [7] (WT) there is an over fit of this parameter and it essentially traces the epidemic curve, lagged by the mean of the serial interval. It is not clear if this is a desirable or informative property. The MLE method has greater promise for generalizability. While it can be argued that adhering to a parametric definition of the shape of the effective reproductive number leads to a greater chance of lack of fit, it can also lead to a result that can be interpretable for other settings that are similar to that being studied.

One can choose any reasonable parametric form for modeling the effective reproductive number. Here we have only shown the four parameter logistic model, and feel that it is suitable in most

Table 4. One step ahead residuals for both methods fitting each epidemic.

Location	Residual		CV Measure	
	MLE	Garske et al	MLE	Garske et al
Ships				
<i>Boonah</i>				
	25.04	3.84	56.57	19.03
<i>Medic</i>				
	13.27	6.31	15.09	30.30
Communities				
<i>Baltimore</i>				
	29.36	20.54	5.81	5.95
<i>Cumberland</i>				
	10.01	5.50	18.37	13.23
<i>Frederick</i>				
	3.80	14.67	13.48	16.37
<i>Salisbury</i>				
	6.32	3.38	20.08	66.84
<i>Lonaconing</i>				
	8.21	4.51	16.69	19.14

The cross validation measure for the ships is the one step ahead residuals calculated from predicting one ship's data using the estimates from the other ship's data. For the communities, it is the sum of the residuals predicting the other four communities using the estimates from the community indicated. For instance the CV measure for Baltimore is the sum of all the residuals that come from using the estimates for Baltimore to predict the other four communities' outbreaks.

doi:10.1371/journal.pone.0001498.t004

cases where the epidemic curve has a single peak. It is feasible that this model may not apply well in all situations. Another approach might be to analyze the data using the Garske et al. method and then smooth the plot of the effective reproductive number and from this determine a parametric form that closely approximates the smoothed curve. Multiple models could be implemented, then the residual analysis that we have shown provides a valuable tool for model assessment and comparison.

The results of these models can be sensitive to underreporting initially in the epidemic. We see this clearly in Boonah, where it was acknowledged that there was underreporting early on and this led to us getting very high estimates for the initial reproductive number. Similarly, in Cumberland, if we remove the first five days of data (three cases on the first day, six cases on the second and then no cases the following three days) we get much more reasonable estimates ($\hat{\mu} = 6.00$, $\hat{\sigma}^2 = 10.32$) with smaller residuals (6.00). Therefore, it is important to note that unusual observations in the first few days can impact the estimates and one should pay careful attention to this possibility.

Overall both methodologies presented are valuable tools that can be used in tandem for understanding the dynamics of infectious disease epidemics. These methods are easy to implement and interpret.

The results that we have presented suggest that the average serial interval for pandemic influenza in 1918 was consistently between three and four, regardless of the setting. The standard deviation for the serial interval distribution varied much more for the MLE method depending on the location. Garske et al. estimates indicate that the value was consistently smaller in the communities than in the ships. It is not clear exactly how to interpret this result. Further, we consistently see a large initial value for the reproductive number. In the ships, this value is higher and rapidly drops off, perhaps due to the close quarters and extremely rapid transmission that could take place in these very vulnerable populations. In the communities, the reproductive number tended to drop off later, typically around day thirty. This could be due to a larger initial susceptible population and more complicated dynamics for the disease to spread, leaving large pockets of susceptible individuals unexposed for a longer period of time than in the ships.

These results confirm the high pathogenicity of influenza and its ability to rapidly spread through populations. It also appears that the greatest difference between the spread of influenza in a closed population without the ability to implement control measures is a large initial reproductive number that declines rapidly. In more diffuse communities with complicated dynamics, it is likely that the reproductive number will not decline as rapidly.

REFERENCES

1. Morens DM, Fauci AS (2007) The 1918 Influenza Pandemic: Insights for the 21st Century. *Journal of Infectious Diseases* 195: 1018–1028, DOI: 10.1086/511989.
2. Centers For Disease Control and Prevention (2006) Pandemic Influenza: Past, Present and Future, Workshop Proceedings. Available: <http://www.pandemicflu.gov/general/workshop.pdf>. Accessed 2007 Oct 4.
3. Fine PE (2003) The interval between successive cases of an infectious disease. *Am J Epidemiol.* 158: 1039–1047.
4. Anderson RM, May RM (1991) *Infectious Diseases of Humans*. Oxford: Oxford University Press.
5. Amundsen EJ, Stigum H, Rottingen J-A, Aalen OO (2004) Definition and estimation of an actual reproduction number describing past infectious disease transmission: application to HIV epidemics among homosexual men in Denmark, Norway and Sweden. *Epidemiol Infect.* 132: 1139–1149.
6. John TJ, Samuel R (2000) Herd immunity and herd effect: new insights and definitions. *Eur. J. Epidemiol.* 16 (7): 601–6.
7. Wallinga J, Teunis P (2004) Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures. *American Journal of Epidemiology* 160: 509–516.
8. Svensson A (2006) A note on generation times in epidemic models. *Mathematical Biosciences* 208: 300–311.
9. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, et al. (2003) Transmission Dynamics and Control of Severe Acute Respiratory Syndrome. *Science.* 300: 1966–1970.
10. Fraser C (2007) Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic. *PlosOne.* doi:10.1371/journal.pone.0000758.
11. Cauchemez S, Boelle P-Y, Thomas G, Valleron A-J (2006) Estimation in real time the efficacy of measures to control emerging communicable diseases. *Am J Epidemiol.* 164: 591–597.
12. Garske T, Clarke P, Ghani AC (2007) The Transmissibility of Highly Pathogenic Avian Influenza in Commercial Poultry in Industrialized Countries. *PlosOne* 4: Available: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0000349> via the internet. Access 21 September 2007.
13. Mills C, Robins J, Lipsitch M (2004) Transmissibility of 1918 pandemic influenza. *Nature* 16: 904–906.
14. Chowell G, Ammon CE, Hengartner NW, Hyman JM (2006) Transmission dynamics of the great influenza pandemic of 1918 in Geneva, Switzerland: Assessing the effects of hypothetical interventions. *J Theor Biol* 241: 193–204.
15. Vynnycky E, Trindall A, Mangtani P (2007) Estimates of the reproduction numbers of Spanish influenza using morbidity data. *Int J Epidemiol* 36: 881–889.
16. Elvebeck LR, Fox JP, Ackerman E, Langworthy A, Boyd M, Gatewood L (1976) An influenza simulation model for immunization studies. *Am J Epidemiol.* 103: 152–165.
17. Halloran E, Longini I, Cowart DM, et al. (2002) Community trials of vaccination and the epidemic prevention potential. *Vaccine.* 20: 3254–3262.
18. Longini IM Jr, Halloran ME, Nizam A, Yang Y (2004) Containing pandemic influenza with antiviral agents. *Am. J. Epidemiol.* 159: 623–633.
19. FluWeb Historical Influenza Database, <http://influenza.sph.unimelb.edu.au>, accessed 25 September 2007.
20. Frost WH, Sydenstricker E (1919) Influenza in Maryland. Preliminary statistics of certain locations. *Public Health Reports* 34: 491–504.
21. White LF, Pagano M (2007) A likelihood based method for real time estimation of the serial interval and reproductive number of an epidemic. *Statistics in Medicine.* In press.
22. Richards FJ (1959) A flexible growth function for empirical use. *J Exp Bot* 10: 290–300.
23. Finney DJ (1976) Radioligand Assay. *Biometrics* 32: 721–740.
24. Healy JJR (1972) Statistical analysis of radioimmunoassay data. *Biochemical Journal.* 130: 207–210.

ACKNOWLEDGMENTS

Author Contributions

Conceived and designed the experiments: LW MP. Performed the experiments: LW. Analyzed the data: LW. Wrote the paper: LW.