Theor Ecol (2011) 4:283–288
DOI 10.1007/s12080-011-0123-3

ORIGINAL PAPER

Mechanistic modelling of the three waves of the 1918 influenza pandemic

DaiHai He · Jonathan Dushoff · Troy Day · Junling Ma · David J. D. Earn

Received: 27 September 2010 / Accepted: 15 February 2011 / Published online: 12 March 2011

Abstract Influenza pandemics through history have shown very different patterns of incidence, morbidity and mortality. In particular, pandemics in different times and places have shown anywhere from one to three "waves" of incidence. Understanding the factors that underlie variability in temporal patterns, as well as patterns of morbidity and mortality, is important for public health planning. We use a likelihood-based approach to explore different potential explanations for the three waves of incidence and mortality seen in the 1918 influenza pandemic in London, England. Our analysis suggests that temporal variation in transmission rate provides the best proximate explanation and that the variation in transmission required to generate these three epidemic waves is within biologically plausible values.

D. He (✉) · D. J. D. Earn
Department of Mathematics & Statistics,
McMaster University, Hamilton, ON, Canada
e-mail: daihai@math.mcmaster.ca, hedaihai@gmail.com

J. Dushoff
Department of Biology, McMaster University, Hamilton,
ON, Canada

J. Dushoff · D. J. D. Earn
M.G. deGroote Institute for Infectious Disease Research,
McMaster University, Hamilton, ON, Canada

T. Day
Department of Mathematics & Statistics,
Queen's University, Kingston, ON, Canada

J. Ma
Department of Mathematics & Statistics,
University of Victoria, P.O. Box 1700 STN CSC,
Victoria, BC V8W 2Y2, Canada

Keywords Influenza · Pandemic · 1918 · London, England · Time series · Mathematical modelling · POMP

Introduction

The 1918 influenza pandemic displayed three distinct waves of infection and mortality within a 12-month period (Jordan 1927; Taubenberger and Morens 2006; Chowell et al. 2008; Miller et al. 2009). This fact has led many scientists and public health planners to anticipate multiple waves of infection whenever a new influenza variant causes a pandemic. The historical pattern of pandemic waves has not been consistent: Different parts of the world have experienced different numbers of waves of given pandemics (e.g. three clear waves in 1918 in England, but only two in the USA (Eggo et al. 2010)), and the number of waves in given locations has been different for different pandemics (e.g. three waves in England in 1918, but two in 1957, one in 1969 and two in 2009 (Miller et al. 2009; Ross et al. 2010; Mishra et al. 2010)). Making predictions about the occurrence of multiple waves during pandemics is hampered, moreover, by the fact that the cause of these epidemiological patterns remains unknown.

Potential causes of multiple waves include the effects of public health interventions (Ferguson et al. 2005; Hatchett et al. 2007; Bootsma and Ferguson 2007), social distancing as a result of individual behavioural responses to the outbreak (Markel et al. 2007; Liu et al. 2007; Caley et al. 2008), seasonality in social and/or climatic factors that affect transmission (Earn et al. 2002; Lipsitch & Viboud 2009; Shaman et al. 2010), viral

evolution (Earn et al. 2002; Day et al. 2006) or rapid loss of immunity in previously infected people.

From a theoretical standpoint, all of these are plausible explanations for the qualitative pattern of multiple waves. It remains unclear, however, whether any of them can provide an adequate quantitative match to observations and, if so, whether the parameter values required to do so are biologically reasonable.

In this paper, we approach this question with an explicit, likelihood-based analysis of mortality data from London, England, during the 1918 pandemic. The mechanisms discussed above act at the population level by changing the effective transmission rate, recovery rate or rate of loss of immunity (or some combination of these). Thus, we construct stochastic epidemic models (Dushoff et al. 2004) and ask how well we can fit observed data if the rates of transmission, recovery and/or loss of immunity change over time and whether the parameter values required to obtain good fits are plausible.

Methods

The data

Weekly pneumonia and influenza mortality reports for the period 1911–1921 were digitized from the Registrar General's Weekly Returns of Births and Deaths in London, England. The Weekly Returns list deaths from pneumonia and influenza separately, but following standard practice (Serfling 1963; Mills et al. 2004; Dushoff et al. 2006), we combine these two categories and begin our analysis from pneumonia and influenza (P&I) mortality. Data are available at the International Infectious Disease Data Archive (http://iidda.mcmaster.ca/).

The epidemic model

The pandemic is modelled using a simple SIRS epidemic model in which we also keep track of individuals who die from infection. The mean field limit of the model can be expressed as the following set of ordinary differential equations:

$$\dot{S} = \delta R - \beta \frac{SI}{N} \tag{1a}$$

$$\dot{I} = \beta \frac{SI}{N} - \gamma I \tag{1b}$$

$$\dot{R} = (1-\phi)\gamma I - \delta R \tag{1c}$$

$$\dot{M} = \phi \gamma I \tag{1d}$$

Here S, I and R are the numbers of susceptible, infected and recovered individuals at time t, and M denotes the cumulative number of individuals who have died from infection. N is the total population size (of those alive), and since $M/N < 0.01$ throughout, we assume for simplicity that N is constant. β is the transmission rate, γ is the recovery rate (i.e. the mean infectious period is γ^{-1}), δ is the immunity decay rate (i.e. mean duration of immunity is δ^{-1}) and ϕ is the case fatality ratio (CFR). We ignore the vital dynamics of the host population because the time frame of interest is a single year. We obtain individual stochastic realizations of the model using the standard fixed time-step binomial approximation (Bretó et al. 2008) of the Gillespie algorithm (Gillespie 1976), in which each term in Eq. 1 determines the rate of a given Markov transition.

With constant parameter values, Eq. 1 gives rise to an outbreak consisting of a single wave of infection and thus a single wave of mortality as well (Dushoff et al. 2004; the same is true of stochastic realizations of the model, except that some realizations fizzle). Therefore, in order to reproduce a pattern of multiple waves, one or more of the parameters must vary over time. For both the transmission rate, β, and the rate of recovery, γ, we consider two possibilities: (a) time independent or (b) time dependent, described by an exponentiated spline function. In particular, we assume these functions can be represented as:

$$\exp\left[b_0(t-t_s) + \sum_{i=1}^{N_B} b_i B_i(t-t_s)\right] \tag{2}$$

where t_s is the start time, b_0 is the magnitude of the linear trend, $B_i(\cdot)$ are periodic (period 1) cubic B-splines (i.e. piecewise-cubic, smooth, periodic basis functions; de Boor 1978) and N_B is the number of basis functions used. The parameters to be estimated are then b_i, $i = 0, 1, \cdots, N_B$. The use of an exponential B-spline, rather than a linear B-spline, is convenient because it ensures that the fitted function is always positive. We also fitted non-periodic (exponentiated) cubic B-splines and a piecewise log-linear function to $\beta(t)$ and obtained negligibly different results (hence we do not present the details of these additional fits).

For the rate of immunity loss, δ, we consider four possibilities: (a) $\delta = 0$ (i.e. no loss of immunity), (b) δ nonzero, time independent (i.e. constant rate of immunity loss), (c) δ time dependent (an exponential B-spline function plus a linear trend as above) or (d) δ zero except a single impulse on 30 August 1918 to mimic the sudden appearance of a new strain before the fall wave (Rios-Doria and Chowell 2009).

It is also believed that the case fatality ratio, ϕ, might have changed over the course of the 1918 pandemic, with the highest levels occurring during the second peak in incidence. Although this alone is unlikely to provide an explanation for the occurrence of multiple waves, it might nevertheless play an important role when fitting the model to mortality data. Consequently, we considered two possibilities with respect to case mortality: (a) ϕ time independent and (b) ϕ piecewise constant:

$$\phi(t) = \begin{cases} \phi_1, & t \in [\text{1918-Jun-07}, \text{1918-Aug-30}), \\ \phi_2, & t \in [\text{1918-Aug-30}, \text{1919-Jan-15}), \\ \phi_3, & t \in [\text{1919-Jan-15}, \text{1919-Jun-07}]. \end{cases} \quad (3)$$

We also explored other models: (a) the SEIRS model includes an exposed class, (b) the SI^3RS model uses two extra infectious classes to create a more realistic infectious period distribution and (c) the SId^3RS model has an explicit (three-stage) delay between recovery and death. Models (b) and (c) use a succession of identical exponential transitions to create Erlang-distributed residence times (Anderson and Watson 1980; Lloyd 2001).

Model fitting

We estimated parameters using maximum likelihood and compared models using the sample-size-corrected Akaike criterion (AICc; Akaike 1974; Burnham and Anderson 2002; King et al. 2008; He et al. 2009). We calculated likelihoods using sequential Monte Carlo methods, also known as *particle filters*, where "particles" refer to samples of a distribution of an unobserved variable and "filtering" means estimating the distribution at a given time based on "particles" available up to that time (Doucet et al. 2001) We maximized likelihoods via iterated filtering (i.e. solving a recursive sequence of filtering problems), using the R package POMP 0.16-9 (King et al. 2009) (technical parameters: number of particles, 10,000; scale coefficient for parameter modification, 2; time delay in initial condition, 52 weeks; cooling factor, 0.95). Observed mortality was modelled as the sum of a baseline process, $b(t)$, reflecting estimated pneumonia deaths *not* due to influenza, and estimated observed deaths from our influenza model, $z(t)$. Likelihoods were calculated using the negative binomial distribution, with a shape parameter that we estimated; this is a standard method for a Poisson process with over-dispersion (Bretó et al. 2008). We used the method of King et al. (2008) to calculate 95% confidence intervals based on the likelihood ratio test. For our model fits, we defined the pandemic period to be the 52 weeks from 1 June 1918 through 30 May 1919.

Seasonal pneumonia

Following the standard approach (Serfling 1963; Mills et al. 2004), we estimated influenza deaths by subtracting a "Serfling baseline" (seasonal pneumonia deaths presumed to have nothing to do with influenza) from

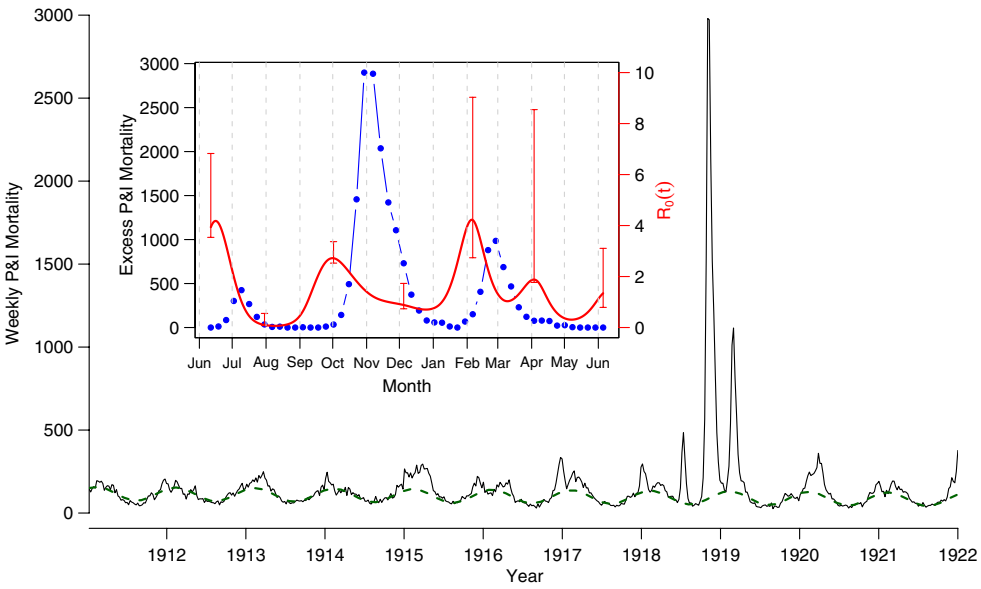

Fig. 1 Weekly P&I mortality in London, England and the reconstructed transmission rate. The *main panel* shows the reported weekly P&I mortality in London, England from 1911 to 1922 (*black*) and the annual baseline of pneumonia mortality (*dark green dashed curve*; see "Methods"). The *inset panel* shows excess P&I mortality from June 1918 to June 1919 (*dotted blue*), together with our reconstructed transmission rate in units of reproduction number $\mathcal{R}_0$ (*red*). Note that we expect disease incidence to increase when the *product* of $\mathcal{R}_0$ and the proportion susceptible exceeds 1

Table 1 Maximum likelihood estimates of parameters from the best-fit model (see Table 2)

Parameter	Our estimates	Our 95% CI	Previous results (reference)
γ^{-1} (days)	4.68	(3.09, 7.09)	3 to 6 (Mills et al. 2004)
Overall $\mathcal{R}_0$	1.52	(1.18, 2.2)	1.3 to 1.9 (Chowell et al. 2008)
ϕ, CFR	0.0078	(0.0058, 0.0213)	0.02 (Mills et al. 2004)

the P&I mortality time series. Thus, we fit the annual baseline function

$$b(t) = a_0 + a_1(t - t_0) + a_2 \cos[2\pi(t - a_3)] \quad (4)$$

to P&I mortality observations for all weeks with <150 cases between June 1913 and May 1918 (using least squares) and defined influenza deaths $z(t)$ to be the "excess" (i.e. total P&I deaths minus $b(t)$).

Results

Figure 1 summarizes our results. The main panel shows the weekly pneumonia and influenza (P&I) mortality data for London, England, from 1 January 1911 to 31 December 1921 (solid black) together with our fitted Serfling baseline for this period (dashed green). Thus, with the standard interpretation of the Serfling fit, deaths attributable to influenza are represented by the difference between the black and green curves. The pandemic period (1 June 1918 to 31 May 1919) had 24,158 P&I deaths, 19,533 (81%) of which we attribute to influenza. In contrast, in the years from 1911 to 1917, there was an average of 6,621 P&I deaths per year (1 June to 31 May), of which an average of 23% are attributable to influenza according to our Serfling fit.

The inset panel of Fig. 1 shows excess P&I mortality (blue) during the pandemic period, together with our fitted time-varying transmission rate $\beta(t)$ in units of reproduction number (red). The $\beta(t)$ curve that we show comes from the best model as judged by AICc, namely the model with constant mean infectious period ($1/\gamma$), constant CFR (ϕ), permanent immunity ($\delta = 0$) and 12 B-spline basis functions to represent $\beta(t)$. Point estimates and confidence intervals for $1/\gamma$, ϕ and the overall (time-averaged) $\mathcal{R}_0$ are given in Table 1, which also lists some previous estimates of these parameters from different 1918 influenza data sets for comparison. For several points on the plotted transmission rate curve, 95% confidence intervals are also shown. While the

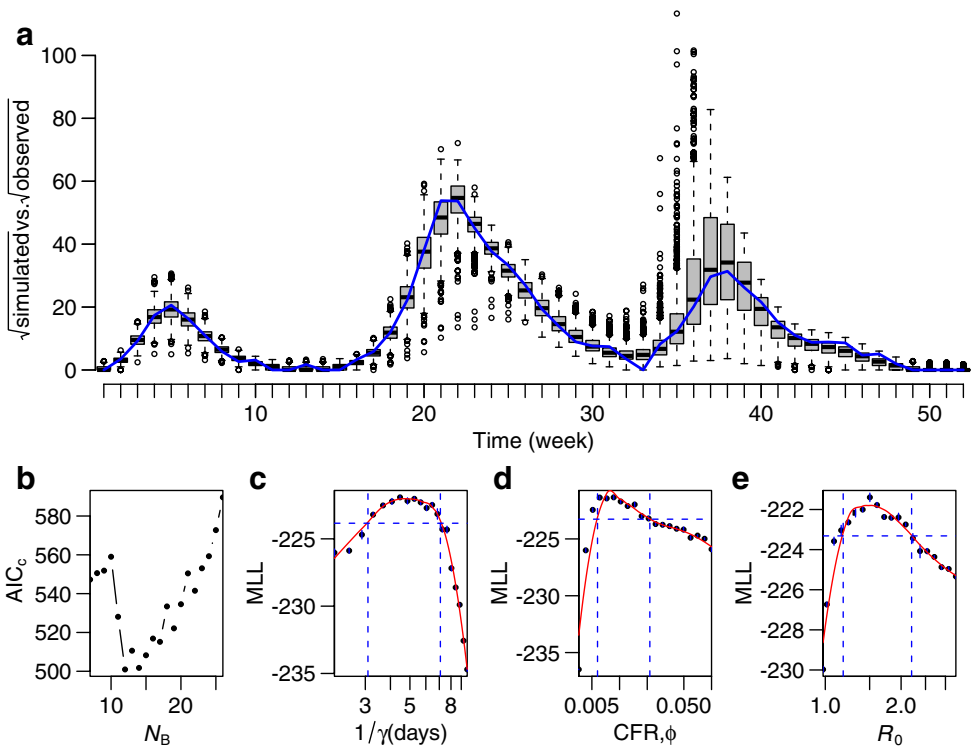

Fig. 2 Excess P&I mortality, realizations of the best model and likelihood profiles for four parameters. The *top panel* **a** shows the observed excess P&I mortality (*blue*, as in Fig. 1), together with box plots of 1,000 realizations of the best model. Panel **b** shows the AICc as a function of N_B, the number of B-spline basis functions. Panels **c**–**e** show likelihood profiles for $1/\gamma$, ϕ and $\mathcal{R}_0$ (with *horizontal lines* that define the 95% confidence intervals; see "Methods")

Table 2 All of the models considered here (ordered from best to worst according to AICc)

Model type	β	γ	δ	ϕ	N_p	N_B	MLL	AICc
SIRS	B-spline	Constant	Zero	Constant	18	12	−221.34	499.42
SI^3RS	B-spline	Constant	Zero	Constant	20	14	−218.35	503.79
SEIRS	B-spline	Constant	Zero	Constant	19	12	−221.16	504.06
SIRS	B-spline	Constant	Impulse	Constant	19	12	−221.86	505.47
SIRS	B-spline	Constant	Constant	Constant	21	14	−216.54	505.88
SIRS	B-spline	Constant	Zero	Step-function	20	12	−221.86	510.82
SIRS	Constant	B-spline	Zero	Constant	22	16	−217.94	514.78
SId^3RS	B-spline	Constant	Zero	Constant	21	14	−221.95	516.69
SIRS	Constant	Constant	B-spline	Constant	17	10	−260.23	572.46

The SEIRS model includes an exposed class. The SI^3RS model includes three stages in the infectious class. The SId^3RS model includes a three-stage delay before deaths are observed. N_B is the number of B-spline basis functions used while N_p is the total number of parameters in the model (including initial conditions). Our best model has time-varying β with 12 cubic B-spline basis, constant γ and ϕ and permanent immunity

confidence intervals are wide, the best model clearly suggests that each pandemic wave was preceded by a rapid rise in transmission rate and a rapid fall before the wave crest. The falls in transmission allow the epidemic to peak before the susceptible pool becomes too small and thus allow subsequent waves of infection.

The top panel of Fig. 2 shows the observed P&I mortality (blue), together with box plots of 1,000 realizations of the best model. The other panels show profiles for some of the key parameters of our fit (see figure caption).

Table 2 lists all the models we considered (ordered from best to worst according to AICc). The different models are defined primarily by which parameters were allowed to vary with time and how (either using B-splines, a step function or a sudden impulse). In this table, N_B is the number of B-spline basis functions used while N_p is the total number of parameters in the model (including initial conditions). Note that for each of the models listed in the table, we examined $N_B = 7$ to 26, but we show only the value of N_B with the best (lowest) value of AICc.

Discussion

We have attempted to obtain a better understanding of the underlying causes of multiple pandemic waves by fitting influenza transmission models to the weekly pneumonia and influenza (P&I) mortality in London, England, over the course of the three waves of the 1918 pandemic. We investigated the proximate mechanisms of multiple wave generation by considering temporal changes in the rates of transmission, recovery and loss of immunity. Based on an explicit likelihood-based analysis, we found that, of these three proximate mechanisms, temporal changes in transmission provide the best explanation for the observed weekly pattern of mortality in London in 1918–1919. Moreover, the magnitude of changes in the transmission rate that are required (as measured by temporal changes in the reproduction number $\mathcal{R}$) are within a biologically reasonable range.

Our analysis provides only indirect insight into the ultimate causes of multiple pandemic waves. Climatic factors (humidity and temperature) are generally believed to act through transmission rates. Behavioural changes can also influence transmission rates (in addition, if people are quarantining themselves based on symptoms, their behaviour could also lead to changes in the effective infectious period and hence the effective recovery rate). Evolution of the virus is most often thought to act through the rate of loss of immunity (antigenic evolution), but it is certainly possible that the transmission rate or recovery rate could evolve as well (particularly in a pandemic variant, which may not be well adapted to humans at first). Our results do argue against rapid loss of immunity in people as a mechanism for the observed waves, since our best models with changing immunity have $\Delta\text{AICc} \approx 6$ compared to the best-fit model. Our best-fit model also provides estimates for the lag time between changes in transmission rate and observed changes in P&I mortality patterns.

The changes in transmission rate that we infer are likely due to a combination of viral evolution (a shift before the first wave and drift or adaptation between waves), behavioural changes in response to the epidemic and changes in temperature and humidity. Although we can only speculate about these mechanisms, our analysis provides a starting point for investigating hypotheses that can be tested using additional data, such as newspaper reports and public health records.

Further work should explore model-fitting methods for combining various types of data (perhaps in a Bayesian framework) to better understand the mechanisms behind multiple waves of influenza pandemics.

Acknowledgements We were supported by the Canadian Institutes of Health Research and the Public Health Agency of Canada. Computations were performed on SHARCNET (http://www.sharcnet.ca). The views expressed in this paper do not reflect the views of the Public Health Agency of Canada.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Contr 19:716–723 doi:10.1109/TAC.1974.1100705

Anderson D, Watson R (1980) On the spread of a disease with gamma distributed latent and infectious periods. Biometrika 67:191–198

Bootsma MC, Ferguson NM (2007) The effect of public health measures on the 1918 influenza pandemic in U.S. cities. Proc Natl Acad Sci U S A 104:7588–7593

Bretó C, He D, Ionides EL King AA (2008) Time series analysis via mechanistic models. Ann Appl Stat 3:319–348

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach, 2nd edn. Springer, New York

Caley P, Philp DJ, McCracken K (2008) Quantifying social distancing arising from pandemic influenza. J R Soc Interface 5:631–639

Chowell G, Bettencourt LMA, Johnson N, Alonso WJ, Viboud C (2008) The 1918–1919 influenza pandemic in England and Wales: spatial patterns in transmissibility and mortality impact. Proc R Soc B 275:501–509

Day T, André J, Park A (2006) The evolutionary emergence of pandemic influenza. Proc R Soc Lond B Biol Sci 273:2945–2953

de Boor C (1978) A practical guide to splines. Springer, New York

Doucet A, de Freita N, Gordon NJ (2001) Sequential Monte Carlo methods in practice. Springer, New York

Dushoff J, Plotkin J, Levin S, Earn DJD (2004) Dynamical resonance can account for seasonality of influenza epidemics. Proc Natl Acad Sci U S A 101:16915–16916

Dushoff J, Plotkin JB, Viboud C, Earn DJD, Simonsen L (2006) Mortality due to influenza in the United States—an annualized regression approach using multiple-cause mortality data. Am J Epidemiol 163:181–187

Earn DJD, Dushoff J, Levin SA (2002) Ecology and evolution of the flu. Trends Ecol Evol 17:334–340

Eggo RM, Cauchemez S, Ferguson NM (2010) Spatial dynamics of the 1918 influenza pandemic in England, Wales and the United States. J R Soc Interface. doi:10.1098/rsif.2010.0216

Ferguson NM, Cummings DA, Cauchemez S, Fraser C, Riley S, Meeyai A, Iamsirithaworn S, Burke DS (2005) Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature 437:209–214 doi:10.1038/nature04017

Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys 22:403–434

Hatchett RJ, Mecher CE, Lipsitch M (2007) Public health interventions and epidemic intensity during the 1918 influenza pandemic. Proc Natl Acad Sci U S A 104:7582–7587

He DH, Ionides EL, King AA (2009) Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. J R Soc Interface. doi:10.1098/rsif.2009.0151

Jordan EO (1927) Epidemic influenza: a survey. Chicago American Medical Association, Chicago

King AA, Ionides EL, Bretó CM, Ellner S, Kendall B (2009) POMP: statistical inference for partially observed Markov processes (R package). Available at http://pomp.r-forge.r-rproject.org

King AA, Ionides EL, Pascual M, Bouma MJ (2008) Inapparent infections and cholera dynamics. Nature 454:877–879. doi:10.1038/nature07084

Lipsitch M, Viboud C (2009) Influenza seasonality: lifting the fog. Proc Natl Acad Sci U S A 106:3645–3646

Liu R, Wu J, Zhu H (2007) Media/psychological impact on multiple outbreaks of emerging infectious diseases. Comput Math Methods Med 8:153–164

Lloyd AL (2001) Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics. Theor Popul Biol 60:59–71

Markel H, Lipman HB, Navarro JA, Sloan A, Michalsen JR, Stern AM, Cetron MS (2007) Nonpharmaceutical interventions implemented by U.S. cities during the 1918–1919 influenza pandemic. JAMA 298:644–654

Miller MA, Viboud C, Balinska M, Simonsen L (2009) The signature features of influenza pandemics implications for policy. N Engl J Med 360:2595–2598

Mills CE, Robins JM, Lipsitch M (2004) Transmissibility of 1918 pandemic influenza. Nature 432:904–906

Mishra AC, Chadha MS, Choudhary ML, Potdar VA (2010) Pandemic influenza (H1N1) 2009 is associated with severe disease in India. PLoS ONE 5:e10540. doi:10.1371/journal.pone.0010540

Rios-Doria D, Chowell G (2009) Qualitative analysis of the level of cross-protection between epidemic waves of the 1918–1919 influenza pandemic. J Theor Biol 261:584–592

Ross T, Zimmer S, Burke D, Crevar C, Carter D, Stark J, Giles B, Zimmerman R, Ostroff S, Lee B (2010) Seroprevalence following the second wave of pandemic 2009 H1N1 influenza. PLoS Curr influenza, p RRN1148

Serfling RE (1963) Methods for current statistical analysis of excess pneumonia-influenza deaths. Publ Health Rep 78:494–506

Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M (2010) Absolute humidity and the seasonal onset of influenza in the continental United States. PLoS Biol 8:e1000316. doi:10.1371/journal.pbio.1000316

Taubenberger JK, Morens DM (2006) 1918 influenza: the mother of all pandemics. Emerg Infect Dis 12:15–22