Journal of Theoretical Biology 312 (2012) 87–95

Contents lists available at SciVerse ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Impact of weekday social contact patterns on the modeling of influenza transmission, and determination of the influenza latent period

S. Towers [a,b,*], G. Chowell [b,c]

[a] Purdue University, West Lafayette, IN, United States
[b] Arizona State University, Tempe, AZ, USA
[c] Division of Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

HIGHLIGHTS

- We assess the impact of weekday social contact patterns on influenza spread.
- Reduced contacts on weekends produce notable beats in epidemic curve.
- The weekday patterns are sensitive to the influenza latent period.
- 2009 A(H1N1) data from Santiago, Chile shows significant weekday patterns.
- We estimate the influenza latent period by fitting to the Chile data.

ARTICLE INFO

Article history:
Received 15 August 2011
Received in revised form
20 July 2012
Accepted 23 July 2012
Available online 1 August 2012

Keywords:
Epidemic model
Weekday contact patterns
Influenza latent period
Influenza incubation period
Pandemic influenza model

ABSTRACT

Human social contact patterns show marked day-of-week variations, with a higher frequency of contacts occurring during weekdays when children are in school, and adults are in contact with co-workers, than typically occur on weekends. Using epidemic modeling, we show that using the average of social contacts during the week in the model yields virtually identical predictions of epidemic final size and the timing of the epidemic incidence peak as a model that incorporates weekday social contact patterns. This is true of models with a constant weekly average contact rate throughout the year, and also of models that assume seasonality of transmission.

Our modeling studies reveal, however, that weekday social contact patterns can produce substantial weekday variations in an influenza incidence curve, and the pattern of variation is sensitive to the influenza latent period. The possible observability of weekday patterns in daily influenza incidence data opens up an interesting avenue of further inquiry that can shed light on the latent period of pandemic influenza. The duration of the latent period must be known with precision in order to design effective disease intervention strategies, such as use of antivirals. For a hypothetical influenza pandemic, we thus perform a simulation study to determine the number of cases needed to observe the weekday variation pattern in influenza epidemic incidence data. Our studies suggest that these patterns should be observable at 95% confidence in daily influenza hospitalization data from large cities over 75% of the time.

Using 2009 A(H1N1) daily case data recorded by a large hospital in Santiago, Chile, we show that significant weekday incidence patterns are evident. From these weekday incidence patterns, we estimate the latent period of influenza to be [0.04, 0.60] days (95% CI). This method for determination of the influenza latent period in a community setting is novel, and unique in its approach.

1. Introduction

Influenza, a viral respiratory disease, is associated with significant morbidity and mortality in the population, with a typical influenza epidemic in the United States killing around 40,000 people per year (Dushoff et al., 2006). During pandemic years, this number can be much larger. Accurate modeling of the spread of influenza within a population is crucial to the design of effective disease intervention strategies, such as the use of antivirals or social distancing measures.

During the course of the week, particularly comparing weekdays to weekends, human social contact patterns vary, tending to

* Corresponding author. Tel.: +1 765 423 7650.
E-mail address: stowers@purdue.edu (S. Towers).

0022-5193/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jtbi.2012.07.023

be higher during weekdays when schools are in session and adults are in contact with co-workers. To examine the effect of these contact pattern variations on the spread of influenza, we use the daily frequency of social contacts from the detailed sociological contact survey data of Mossong et al. (2008), where participants kept daily diaries recording the length and nature of their contacts. Using these data, and the results of a previous study of the effect of weekday social contact patterns on the spread of infectious disease (Hens et al., 2009), we determine age-structured contact matrices for weekdays and weekends. We then use these matrices in an age-structured deterministic influenza model (the *weekday model*) and compare the results to that of a model that does not take into account weekday contact patterns, and instead uses a weekly average contact matrix (the *average model*). We find that when the average weekly contact rate is assumed constant throughout the year (ie; no seasonality) the two models produce virtually identical predictions of epidemic peak incidence time and final size, in agreement with the results of previous studies (Bacaer and Abdurahman, 2008).

Influenza is a seasonal disease in countries with temperate climates, displaying markedly higher attack rates in winter months. There appear to be several reasons for this, including seasonality of host health (susceptibility to infection) and seasonal environmental effects on the transmissibility of the virus (Cauchemez et al., 2008; Alonso et al., 2007; Dowell, 2001; Lofgren et al., 2007; Lowen et al., 2007; Shaman et al., 2010). Periodicity of school holiday closures is also thought to play a role (Chowell et al., 2011; Chao et al., 2010; Cauchemez et al., 2009). Such seasonality can be important to take into consideration when modeling pandemic influenza to assess disease intervention strategies (Feng et al., 2011; Towers et al., 2011, 2012). When we include seasonal periodicity of the contact rate in the model, we also find that the average and weekday models produce nearly identical predictions of epidemic final size and incidence peak time(s). From a practical standpoint, it is thus sufficient to use the weekday average contact rate if one is primarily interested in modeling the overall progression of the epidemic, such as final size or incidence peak time.

However, our modeling studies demonstrate that weekday patterns in human social contacts produce a marked weekday beat structure in the epidemic incidence curve. As we will discuss in the following sections, inclusion of weekday patterns in the model has the potential to allow for determination of the latent period of influenza if the model is fit to temporal patterns in daily influenza incidence data. Accurate knowledge of the influenza latent period has important implications for optimization of disease intervention strategies, such as the use of antiviral treatment (see, for instance, Longini et al., 2004).

We thus perform a statistical study, simulating a hypothetical pandemic, that determines that the patterns in weekday social contacts would be expected to likely produce an observable effect in daily pandemic incidence data from a large city. Using daily incidence data from A(H1N1)-related emergency department visits in Santiago, Chile during the 2009 pandemic, we then show that significant weekday patterns are indeed evident. We fit an age-structured epidemic model to this incidence data, and determine that the latent period is less than half a day with 95% confidence. Our method for determination of the influenza latent period in a community setting is novel, and unique in its approach.

2. Data

Data used in these studies are the time series of emergency department visits per day during the 2009 A(H1N1) pandemic at a large hospital in Santiago, Chile (Torres et al., 2010). The population of Santiago is 5.9 million (CIA World Factbook, 2011). The visits are broken down into all visits, and visits made by patients whose symptoms met the clinical definition of influenza A(H1N1) infection.

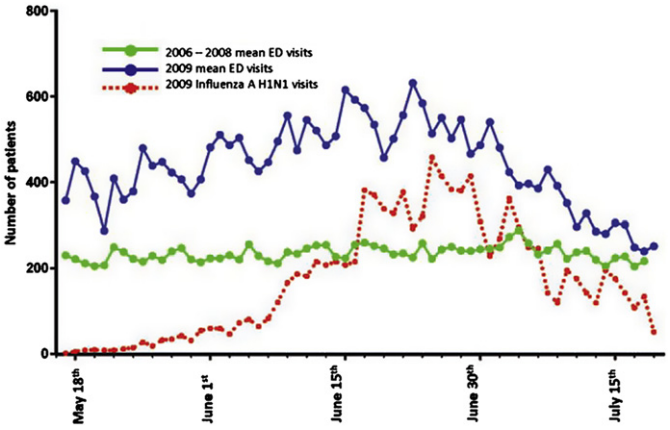

Fig. 1. The daily number of A(H1N1)-related emergency department visits at a large hospital in Santiago, Chile, during the 2009 A(H1N1) epidemic (red). Shown in blue are the daily number of emergency department visits during the same period for all causes. Reproduced from Reference (Torres et al., 2010). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Of the over 10,000 patients clinically diagnosed with A(H1N1), the average time from symptom onset to diagnosis is 1.4 days, with a median of 1 day, properties that are consistent with those of an Exponential distribution with mean 1.4.

In Fig. 1 we show the A(H1N1)-related visits recorded by day by the hospital. We fit a spline to the epidemic incidence curves, with enough degrees of freedom to capture the average behavior without capturing weekday patterns, as shown in Fig. 3. We then take the averages of the data minus spline fit within weekday; for instance, for data points that fall on a Monday, the average of the data minus the spline fit is calculated. The resulting averages by weekday are shown in the lower plot of Fig. 3.

3. Model

One of the simplest epidemiological models is the so called compartmental SEIR model (Hethcote, 1976), which keeps track of the number of susceptible (S), exposed but not yet infectious (E), infectious (I), and recovered (R) people in the population using the coupled deterministic ordinary non-linear differential equations:

$$\begin{aligned} S' &= -\beta(t)SI/N, \\ E' &= \beta(t)SI/N - \kappa E \\ I' &= \kappa E - \gamma I, \\ R' &= \gamma I, \end{aligned} \tag{1}$$

where $1/\kappa$ is the average latent period for influenza, $1/\gamma$ is the average infectious period, $\beta(t)$ is the time-dependent transmission rate, and population size is given by $N = S+E+I+R$. Because we will be considering epidemics of very short duration relative to human population dynamics, we do not include vital dynamics in the models we discuss here. We also assume no disease-induced mortality for simplicity.

Typically the transmission rate $\beta(t) = \beta$ is assumed to be a constant. However, for diseases such as influenza, seasonal variation in the transmission rate is an important consideration (Feng et al., 2011; Towers et al., 2011, 2012; Towers and Feng, 2009). Any periodic function can be expressed as a sum of harmonic

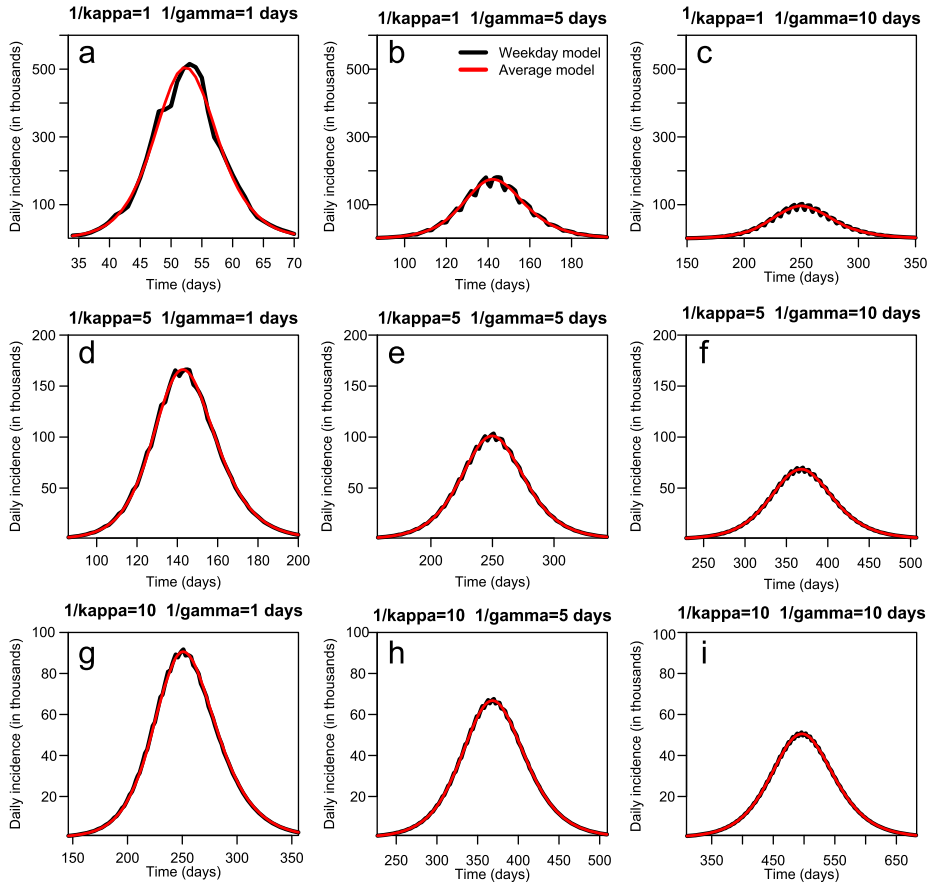

Fig. 2. Weekday and average model predictions of daily incidence in a population of 10 million under various assumptions of the latent and infectious periods, as estimated by Eq. (3). The transmission is assumed to be constant, with $\mathcal{R}_0 = 1.7$. A longer latent period reduces the relative variation in weekday patterns, but a longer infectious period does not. Smaller $\mathcal{R}_0$ produces equivalent relative weekday variation in the epidemic curve to that shown here, but the final size of the epidemic is smaller.

terms, and in this analysis we model periodicity of transmission due to seasonal effects on host health, school holidays, and seasonal environmental effects (such as temperature, humidity, etc.) using the first order harmonic

$$\beta(t) = \beta_0[1+\epsilon\cos(2\pi\omega t)], \tag{2}$$

where the period is $1/\omega = 365$ days, β_0 is the average transmission rate over one period, and ϵ is the degree of seasonal forcing (Bacaer and Gomes, 2009; Bacaer and Dads, 2010). We define the β function given in Eq. (2) to have its maximum at the beginning of a calendar year. The results of this analysis can be easily generalized to include a phase if it is believed that influenza transmission peaks at some other time of year, but inclusion of a phase does not affect the overall conclusions of the study.

The time of introduction of the virus to the population, t_0, is a parameter of the seasonal model when the transmission rate is expressed in a periodic form. Disease models with seasonal transmission rates are extremely non-linear, and the final size and dynamics predicted by such models are strongly dependent upon the model parameters and initial conditions (Feng et al., 2011; Towers et al., 2011, 2012; Towers and Feng, 2009; Bacaer and Gomes, 2009; Dietz, 1976; Hethcote and Levin, 1988; Shi et al., 2010). This can lead to interesting multiwave dynamics, such as those observed during the 2009 and 1918 pandemics (Miller et al., 2009). When there are no periodic parameters in the model the predicted final size and shape are independent of t_0, and the epidemic incidence curve can exhibit only one peak in the absence of vital dynamics in the model.

We expand the model in Eqs. (1) to include age heterogeneity, with mixing between $n=2$ different age-stratified classes representing children less than 19 years of age, and adults:

$$\begin{aligned} S_i' &= -\beta(t)S_i\sum_{j=1}^{n} C_{ij}(t)I_j/N_j,\\ E_i' &= \beta(t)S_i\sum_{j=1}^{n} C_{ij}(t)I_j/N_j - \kappa E_i,\\ I_i' &= \kappa E_i - \gamma I_i,\\ R_i' &= \gamma I_i, \end{aligned} \tag{3}$$

where the population size, N, is $N=\sum_i N_i = \sum(S_i+E_i+I_i+R_i)$. The matrix C_{ij} is known as the *contact matrix*, and is the average number of contacts made per day by people in class i with people in class j. In this analysis we use contact data from the detailed sociological contact survey data of Mossong et al. (2008), where participants kept daily diaries recording the length and nature of their contacts; using the methods of Medlock and Galvani (2009), we parameterize the contact matrix by averaging the Mossong survey data wherein respondents estimated the number of daily contacts by age of respondent and by age of contact.

For constant $\beta(t)=\beta_0$ and a time invariant contact matrix, Eqs. (3) can be linearized about the disease free equilibrium to yield the basic reproduction number:

$$\mathcal{R}_0 = \frac{\beta_0}{2\gamma}\left[(C_{11}+C_{22}) + \sqrt{(C_{11}-C_{22})^2+4C_{12}C_{21}}\right]. \tag{4}$$

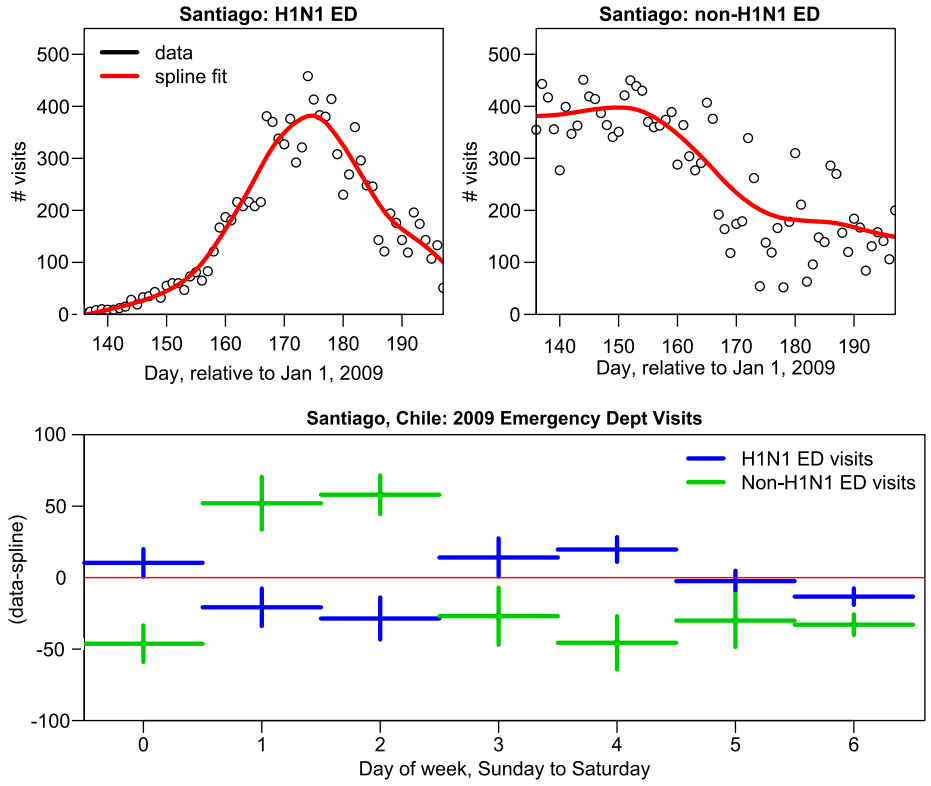

Fig. 3. The top two plots show the daily number of H1N1 and non-H1N1 related emergency department visits at a large hospital in Santiago, Chile, during the 2009 A(H1N1) epidemic. Shown in red are spline fits to the distributions. The bottom plot shows the mean of the data minus the spline fit, averaged within weekday. The error bars represent the one standard deviation uncertainties on the average. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

This is also a good approximation to the $\mathcal{R}_0$ of the weekday model, with or without seasonality of transmission (Williams and Dye, 1997; Moneim, 2007; Wang and Zhang, 2007; Bacaer, 2007).

To model weekday variations in social contacts, we separate the contact matrix into a weekday matrix, C^{week}, and a weekend matrix, C^{end}. Hens et al. (2009) point out that the peer-to-peer contacts between school children and co-workers are reduced during weekends by a factor sufficient to reduce the overall basic reproduction number, $\mathcal{R}_0$, by around 20%. We find that a 25% reduction in peer-to-peer contacts from weekday to weekends is sufficient to achieve a reduction in $\mathcal{R}_0$ by 19%. The elements of the weekly average contact matrix $C=(5C^{\text{week}}+2C^{\text{end}})/7$ are shown in Table 1.[1]

The model has initial conditions at the time of introduction, t_0, of the index case into the population

$$S_i(t_0)=N_i(1-\nu_i)-1,$$

$$I_i(t_0)=1,$$

$$E_i(t_0)=0,$$

$$R_i(t_0)=\nu_i, \tag{5}$$

where $i=1,2$ and ν_1 and ν_2 are the initial immune fractions due to prior infection or vaccination, and N_1 and N_2 are based on the demographics of the population. In this analysis we assume no pre-immunity for simplicity (ie; $\nu_1=\nu_2=0$).

The parameters of the model are shown in Table 1.

[1] Note that the off-diagonal elements of C^{week} and C^{end} are the same as that of the average matrix, C, and the diagonal elements can easily be obtained using the relationships $C_{ii}^{\text{week}}=(7/(5+2p))C_{ii}$ and $C_{ii}^{\text{end}}=(7p/(5+2p))C_{ii}$, where $p=0.75$.

4. Results

4.1. Comparison of weekday and average models

In this section we study and compare results of the model in Eqs. (3) under two scenarios:

1. Contact matrix varies from a weekday contact matrix to a weekend contact matrix (the *weekday model*).
2. Contact matrix remains constant at an average of the weekday and weekend contact matrices (the *average model*).

We first consider the case of a constant transmission rate. The latent and infectious periods for seasonal influenza have been estimated from challenge and observational studies to be $1/\kappa=1.1$ and $1/\gamma=4.8$ days, respectively (Carrat et al., 2008). Near the beginning of the 2009 A(H1N1) pandemic in the Northern Hemisphere, the $\mathcal{R}_0$ in community settings was estimated to be around 1.3–1.7 (Towers and Feng, 2009; Fraser et al., 2009; Yang et al., 2009). To model an equivalent pandemic to that of 2009, we assume $\beta(t)=\beta_0$, set $C_{ij}(t)$ constant at $C=(5C^{\text{week}}+2C^{\text{end}})/7$, and solve Eq. (4) for β_0 as a function of $\mathcal{R}_0$. Assuming $\mathcal{R}_0=1.7$, we then estimate β_0.

The resulting epidemic incidence curves are shown in the top plot of Fig. 4 for the average and weekday models, for a population of size $N=10{,}000{,}000$, and assuming no background immunity to the virus. The temporal profile of the weekday model shows clear weekly beats in the epidemic incidence curve, whereas the average model, as expected, does not. However, it should be noted that the two models yield virtually identical predictions of the peak times and final sizes versus $\mathcal{R}_0$, as seen in the bottom two plots in Fig. 4. This is in agreement with the results of Bacaer and Abdurahman (2008), which found little difference in the growth rate prediction of an SEIR model with

Table 1
Definition of symbols and parameter values used in simulations.

Variables	Definition	Value (range)
$S(t)$	Number of susceptible individuals at time t	
$E(t)$	Number of exposed individuals at time t	
$I(t)$	Number of infectious individuals at time t	
$R(t)$	Number of recovered individuals at time t	
Parameters		
$1/\kappa$	Avg sojourn in Exposed class	$1/\kappa = 1.1$ days (Carrat et al., 2008)
$1/\gamma$	Avg sojourn in Infectious class	$1/\gamma = 4.8$ days (Carrat et al., 2008)
t_0	time of introduction of the virus to the population	Varied
f	Fraction of population under 19 years of age	0.30
N	Population size	10,000,000
C_{11}	Avg number of child-to-child contacts per day	8.9
C_{12}	Avg number of child-to-adult contacts per day	5.5
C_{21}	Avg number of adult-to-child contacts per day	1.9
C_{22}	Avg number of adult-to-adult contacts per day	9.3
$\mathcal{R}_0$	Reproduction number of pandemic influenza	1.5
β_0	Average transmission rate	$\mathcal{R}_0\gamma/12.4$ (see Eq. (4))
ϵ	Oscillation magnitude of seasonal transmission rate, $\beta(t)$	0.30
p	Scale factor for peer-to-peer contact during weekends	0.75

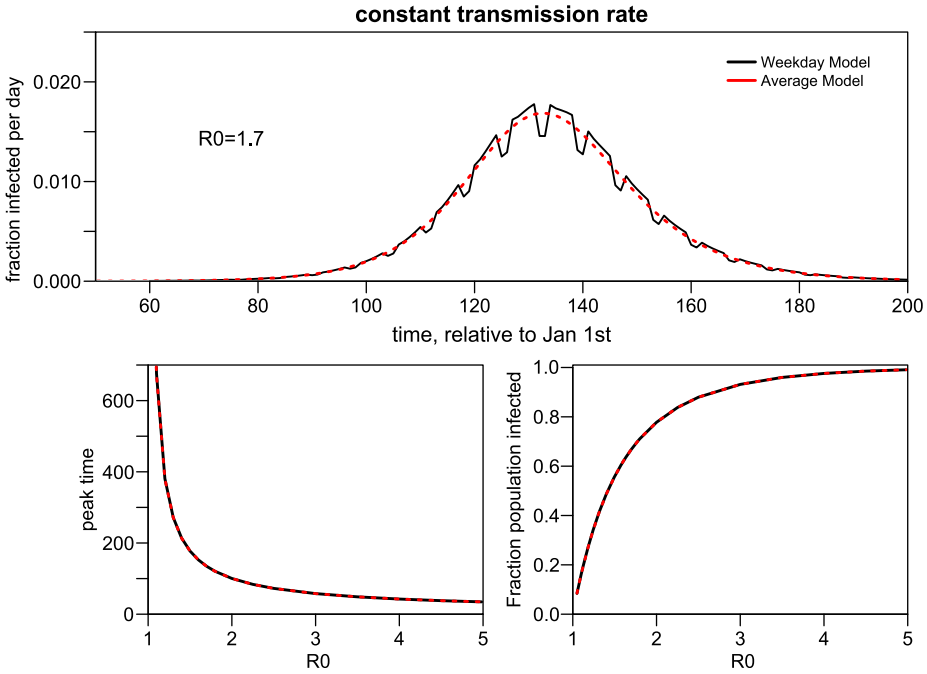

Fig. 4. The top plot shows the epidemic incidence curves for the average and weekday influenza models (see text for details), using a constant transmission rate, with a population of size $N = 10{,}000{,}000$, and assuming no pre-immunity to the virus. As seen in the bottom two plots, for all $\mathcal{R}_0$, the average and weekday model yield similar predictions for the epidemic final size and peak time.

a single age class and a constant contact rate compared to that predicted by a model that included weekly periodicity in the contact rate, as long as the relative variation in the contact rate was less than around 50% (which is the case here).

Turning our attention now to the seasonal model, we note that data and modeling studies indicate that values of 0.25–0.35 appear to be reasonable estimates for transmission rate seasonal forcing term, ϵ (Towers and Feng, 2009; Bacaer and Dads, 2010; Ferguson et al., 2003). Here we assume $\epsilon = 0.30$, and again assume $\mathcal{R}_0 = 1.7$, calculating β_0 as above. In Fig. 5 we show the number of new infections vs time for two different times of introduction of the virus to the population, for a population of size $N =$ 10,000,000, and assuming no pre-immunity to the virus. Again the weekday model exhibits obvious weekday structure, and the average model closely matches the overall average shape of the epidemic incidence curve.

In Fig. 6 we compare the weekday and average seasonal model predictions of final size versus time of introduction of the virus to the population, assuming $\mathcal{R}_0 = 1.7$. Fig. 6 also compares the predictions of the two models of the time average of the incidence vs time of introduction of the virus; if we have incidence measurements, $\mathcal{Y}_i$, at $i = 1, \ldots, k$ consecutive time points, t_i, along the epidemic curve, the time average of the incidence is defined as

$$\bar{t} = \frac{\sum_{i=1}^{k} t_i \mathcal{Y}_i}{\sum_{i=1}^{k} \mathcal{Y}_i}. \tag{6}$$

For a symmetric incidence distribution with one peak, $\bar{t}$ is the same as the peak time. There is marked agreement between the predictions of final size and $\bar{t}$ of the two models.

4.2. Observability of weekday patterns in data

An important consideration is observability of the phenomenon of weekday beats in influenza daily epidemic data. Observability is affected by various factors:

- The relative variation of weekday patterns compared to the null hypothesis of no weekday variation; larger relative weekly variation of the "wiggles" in the epidemic curve increases the probability of detecting significant evidence of weekday patterns in incidence.
- The number of cases detected during the course of the epidemic; if only a very small fraction of cases are detected, the relative stochasticity in the detected daily incidence will be large. Large relative stochasticity makes it less probable to detect significant evidence of weekday patterns in incidence.
- The final size of the epidemic; a smaller $\mathcal{R}_0$ results in a smaller final size, which increases the relative stochasticity in the detected daily incidence.
- The probability distribution assumed for the time elapsed between onset of infectiousness and detection of the case; a broad probability distribution smears out variations in the daily incidence due to weekday variations in contact, making those variations harder to detect.

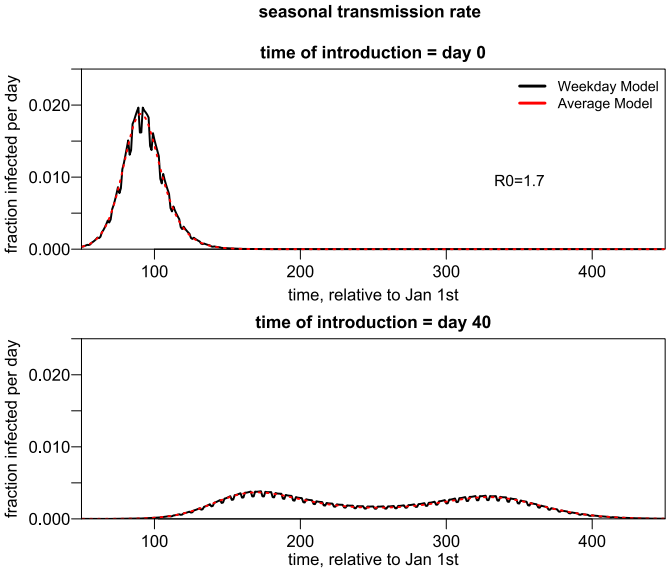

Fig. 5. The number of new infections vs time (incidence) for two different times of introduction of the virus to the population for a model with a seasonal transmission rate and a population of size $N = 10{,}000{,}000$, and assuming no pre-immunity to the virus. The black (red) line is the epidemic incidence curve predicted by the weekday (average) model. The upper and lower plot have differing y-axis scales. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In Fig. 2 we show the predicted total daily incidence in a population of 10 million for the weekday and average models under various assumptions of κ and γ. The model predictions are from Eqs. (3) under the assumption of $\mathcal{R}_0 = 1.7$ and constant transmissibility. The relative weekday variation in the curve is most affected by the latent period, $1/\kappa$, with a short $1/\kappa$ of only 1 day producing a relative variation of around 6–7% compared to the average model (regardless of the value of γ and $\mathcal{R}_0$), and a longer $1/\kappa$ of 10 days producing a relative weekday variation of only around 1% (again, insensitive to the value of γ and $\mathcal{R}_0$). Thus a short latent period would result in greater observability of weekday patterns compared to a long latent period.

We find that, for a given $\mathcal{R}_0$ and case detection fraction, the duration of the epidemic will not affect the observability if due to an increase in the length of the infectious period; a choice of γ that doubles the epidemic duration also halves the incidence in each bin, although the weekday variation relative to the average model remains the same (compare, for instance, plots b and c in Fig. 2). The longer epidemic does provide more observations of how the daily incidence on a particular day of the week differs from the expected incidence under the null hypothesis that no weekday patterns in incidence are present. However, the longer epidemic also reduces the daily incidence by an amount that increases the relative stochasticity of each observation, exactly

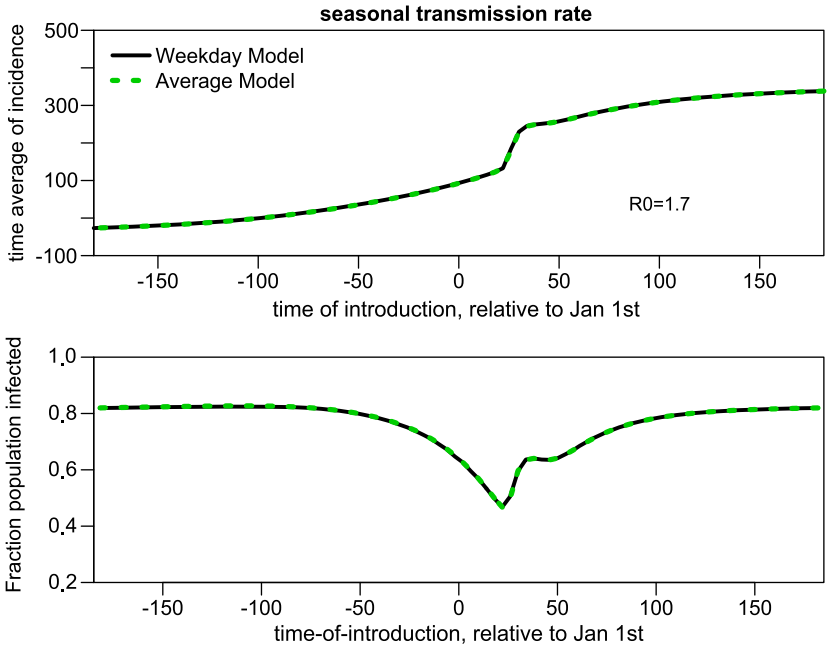

Fig. 6. Time average of the incidence and final size of the epidemic vs time of introduction of the virus to the population, as predicted by the weekday and average models with a seasonal transmission rate.

counterbalancing the effect of having more time points to observe the weekday variation. Thus the value of γ has little effect on the observability of weekday patterns. If the long duration of the epidemic is due to a long latent period, the observability of weekday patterns is reduced because of the decrease in the relative weekday variation in the daily incidence curve.

To study the effect of case detection rate on observability, we use the weekday model to estimate a hypothetical daily hospitalization incidence in a population of $N=10{,}000{,}000$, assuming a constant transmission rate with $\mathcal{R}_0=1.7$, and $1/\gamma=4.8$ days (Carrat et al., 2008). To simulate the difference between time of becoming infectious and time of detection, we convolve the incidence curve determined from Eqs. (3) with an Exponential distribution with mean 1.4 days (Torres et al., 2010). The convolution is achieved by numerically integrating

$$y^{\text{pred}}(t)=\int_0^\infty d\tau \mathcal{Y}^{\text{pred}}(t-\tau)p(\tau), \tag{7}$$

where $\mathcal{Y}^{\text{pred}}(t)$ is the model prediction of the number of people entering the symptomatic I class at time t, and $p(\tau)$ is the probability that an infectious case will be confirmed at time τ after entering the I class.

We scale $\sum_t y^{\text{pred}}(t)$ to simulate varying numbers of total detected cases over the course of the epidemic, from 1000 to 50,000 (with fewer detected cases implying a lower case detection fraction). To obtain the final simulated epidemic incidence curve, we randomly Poisson vary the resulting number of detected cases at each time bin, t_i, around the mean, $y^{\text{pred}}(t_i)$. We then examine an eight week period centered around the time of peak epidemic incidence. A spline curve was fit to this distribution, with degrees of freedom sufficiently high to capture the essential shape of the distribution, but not so high as to fit to the short term temporal patterns in the distribution. It was checked that the results of the analysis were insensitive to minor changes in the degrees of freedom of the spline fit.

We repeat the epidemic incidence curve simulation procedure 1000 times. We calculate the Pearson-χ^2 statistic comparing the weekday distribution of the simulated data minus the spline fit with the expected model. We count the number of simulated data sets for which this distribution is consistent with the expected pattern from the weekday model to a greater than 5% CL, and yield data minus spline inconsistent with no weekday pattern to greater than 95%.

When the size of the simulated data sample is 1000 (5000, 10,000, 50,000), weekday patterns are observable at a 95% CL around 15% (45%, 75%, $>99\%$) of the time.

The total number of infections predicted by the model is 7.2 million. If we assume a 65% symptomatic rate (Elder et al., 1996; King et al., 1988), a hospitalization rate of symptomatic cases of 0.0045 (Reed et al., 2009), and assume that 50% of all hospitalized cases are confirmed to be influenza, we predict that around 10,000 confirmed hospitalized cases will occur in the population. Thus weekday patterns would be expected to be observable at a 95% CL in such data around 75% of the time.

Using a larger mean in the Exponential convolution to simulate time to reporting (such as would perhaps be appropriate for hospitalization or ICU data compared to data based on visits to an emergency department or physician's office) will lower the observability of the weekday phenomenon. The average time between onset of symptoms and hospitalization has not been well studied, and may perhaps be longer for hospitalized cases than cases identified in an emergency department of physicians office; for instance, communication with the authors and the New South Wales Department of Health reveals that during the 2009 pandemic the average time between first symptoms and hospitalization of the 240 people admitted to the ICU with confirmed A(H1N1) infection was 5 days (range from 2 to 7.5 days). If this mean is used instead of the 1.4 day mean observed in the Santiago, Chile emergency department visits, the weekday beat phenomenon is only observable 50% of the time when the total number of cases in the sample is 50,000.

Similarly, as expected from the discussion above, the simulated data studies show that the value of the latent period, $1/\kappa$, affects observability; if we simulate time to reporting with an Exponential of mean 1.4, and use $1/\kappa=3$ days instead of $1/\kappa=1.1$ days, the weekday beat phenomenon is only observable around 80% of the time when the total number of cases in the sample is 50,000. Using $1/\kappa=5$ days reduces the observability to 50% for a sample of that size. It thus appears that the observation of weekday beat phenomenon in incidence data in and of itself implies that $1/\kappa$ is likely low.

4.3. Estimation of latent period using Santiago, Chile pandemic influenza data

In Fig. 3 we show the A(H1N1) and non-A(H1N1) related visits recorded by day by a large hospital in Santiago, Chile during the 2009 A(H1N1) influenza epidemic (Torres et al., 2010). We fit a spline to the epidemic incidence curves, with enough degrees of freedom to capture the average behavior without capturing weekday patterns. The lower plot in Fig. 3 shows the data minus spline fit, averaged by weekday. The results of the lower plot are insensitive to minor changes in the degrees of freedom used in the spline fits. In the non-A(H1N1)-related visits, significant patterns are evident that are likely due to temporal patterns in healthcare-seeking behavior: the most likely days to visit the emergency department for non-A(H1N1)-related reasons were Monday and Tuesday, a pattern that matches that observed in a study of weekday patterns in emergency department visits in New York (New York Department of Health, 2011). Interestingly, Monday and Tuesday were the least likely days to visit for A(H1N1)-related reasons. It is unknown how much the weekday patterns of A(H1N1)-related visits are affected by the general patterns of healthcare seeking behaviors, but it is clear that any correction for healthcare seeking behaviors would magnify the already significant dip in A(H1N1)-related visits on Monday and Tuesday.

We fit to the A(H1N1) incidence data using a weekday SEIR model; since the epidemic occurs over a relatively short period, we assume that the average transmission rate is approximately constant throughout the epidemic. There were three public holidays during the epidemic, occurring on May 1, 21, and June 29. We use a weekend contact matrix in the model on those days. We set $1/\gamma=4.8$ days (Carrat et al., 2008), and the optimal values of the time of introduction t_0, γ, κ, and the transmission rate, β, are determined by fitting the parameters of the model to the $M=62$ data points.

The two most common model parameter estimation methods are maximum likelihood and least squares estimation (MLE and LSE, respectively). MLE is perhaps not as widely recognized as LSE in the field of mathematical biology, but is a standard approach to parameter estimation and inference in statistics. MLE has many optimal properties in estimation. For instance, parameter estimates from MLE are consistent (the true parameter value that generated the data is recovered asymptotically, which means that for data of sufficiently large samples the estimate is unbiased), and efficient (which means that the lowest-possible variance of parameter estimates is achieved asymptotically) (Myung, 2003; Cowan, 1998). In contrast, parameter estimates from LSE methods in general have neither properties (Cowan, 1998).

The MLE method estimates model parameters by maximizing the probability of observing the data, under the assumption that

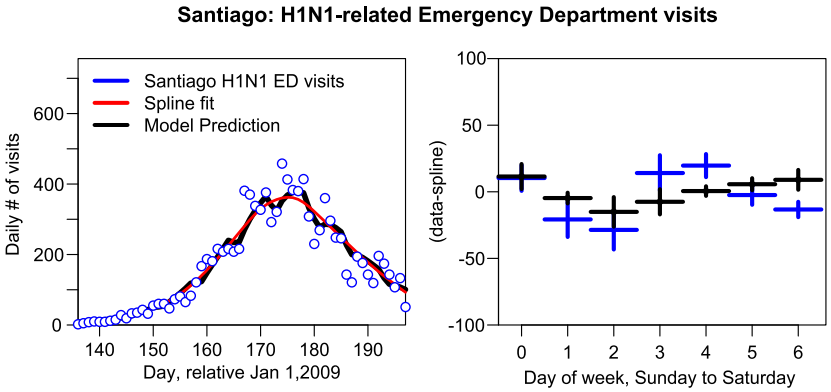

Fig. 7. The blue points in the left plot represent the daily number of H1N1 related emergency department visits at a large hospital in Santiago, Chile, during the 2009 A(H1N1) epidemic. Shown in black is the best fit to the data using an SEIR model with weekday contact patterns. The red line represents a spline fit to the model prediction (the spline fit to the data is shown in Fig. 3). The right plot shows the average within weekday of the model minus spline (black), and data minus spline (blue). The error bars represent the one standard deviation uncertainties on the averages. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the data are described by the model. In this case, if we assume that the average number of new infectious people at time t_m is the disease model prediction, $\lambda_m = y^{\text{pred}}(t_m|\kappa,\gamma,\beta)$, and that the probability distribution underlying the stochasticity in each bin is the Poisson distribution with that mean, the probability (likelihood) of observing the data given the model parameters κ, γ, and β is

$$\mathcal{L} = \prod_{m=1}^{M} (\lambda_m)^{y_m^{\text{data}}} e^{-\lambda_m} / (y_m^{\text{data}})!. \tag{8}$$

The best-fit disease model parameters are the ones that maximize this probability.

In practice, sums of logarithms are easier to work with than products of probabilities, and maximizing the sum of the probability logarithms is equivalent to maximizing their product (Cowan, 1998). We thus find the best fit κ and β by maximizing the log likelihood

$$\log\mathcal{L} = \sum_{m=1}^{M} y_m^{\text{data}} \log(y^{\text{pred}}(t_m|\kappa,\gamma,\beta)) - y^{\text{pred}}(t_m|\kappa,\gamma,\beta). \tag{9}$$

The best fit to the distribution is shown in Fig. 7. The 95% CI estimate of the latent period from the fit is $1/\kappa = [0.04,0.60]$ days. The 95% CI estimate of the time of introduction of the virus is $t_0 = [75,88]$ days (i.e.; near the end of March, 2009), in reasonable agreement with studies of data from Mexico, the origin of the 2009 A(H1N1) epidemic, that showed the virus was introduced there sometime around the end of February (Fraser et al., 2009). The 95% CI estimate of the infectious period is $1/\gamma = [5.3,6.3]$ days, slightly higher than the value of $1/\gamma = 4.8$ days estimated from volunteer challenge studies. The 95% CI estimate of $\mathcal{R}_0$ is [1.7,1.9], again in approximate agreement with the various predictions of previous studies of 2009 A(H1N1) data (Towers and Feng, 2009; Fraser et al., 2009; Yang et al., 2009; Nishiura et al., 2009).

5. Discussion

In this paper we examined an age-structured SIR epidemic model with and without seasonally varying transmission rates, and with and without contact rates that varied by weekday.

For both the constant and seasonal transmission rate model, we found that taking into account weekday variations did not change the model prediction of the peak incidence time and final size, but did cause notable weekday patterns in the epidemic incidence curve.

Because the weekday patterns predicted by the model depend on the latent period, $1/\kappa$, such patterns can be used to estimate $1/\kappa$ if it is possible to observe them in data. To explore the issue of observability of this phenomenon in pandemic influenza data, we assumed a population size of 10 million, and $\mathcal{R}_0 = 1.7$. We then simulated synthetic data samples of various sizes from 1000 to 50,000 cases. We convolved the incidence curve predicted by the disease model with an Exponential distribution to simulate the time between infection to detection of infection, and randomly Poisson varied the bins of the distribution about the mean. Under these assumptions, we found that in a population of 10 million people and a sample size of 10,000 (50,000), the phenomenon of weekly beats in the hospitalization data should be observable at a 95% CL over 75% ($>99\%$) of the time. The results of this study, of course, rely upon our model and all subsequent assumptions being applicable to the study population.

We then examined a sample of over 10,000 influenza related emergency department visits recorded daily by a large hospital in Santiago, Chile during the 2009 A(H1N1) epidemic. We found that significant weekday patterns were evident, and were inconsistent with weekday patterns in healthcare seeking behaviors. Fitting to the epidemic incidence curve using an SEIR weekday model determined that the latent period is $1/\kappa = [0.04,0.60]$ days (95% CI).

This method for determination of the latent period from cases in a community setting is novel and unique, and yields a value of $1/\kappa$ that is quite consistent with results of viral challenge studies; the meta-analysis of influenza virus challenge studies presented in Carrat et al. (2008) determines that 83% (97%, 100%) of 77 participants had detectable virus shedding after 24 h (48, 72). This distribution is consistent with that of an Exponential distribution with a mean of [0.4,0.75] days (95% CI).

Previous studies of spread of influenza in a community setting estimate the incubation period as the time from contact with an infected person to the onset of symptoms (Tuite et al., 2010; Moser et al., 1979; Ghani et al., 2009; Lessler et al., 2009), which is not necessarily the same as the time from contact to onset of infectiousness (latent period). The studies determine the incubation period to be from 1.4 to 4 days, which is longer than the latent period determined by this study. However, Carrat et al. (2008) note that viral shedding can be detected at least 24–48 h before symptom onset; in conjunction with the results of this study, it appears that the latent period of influenza is shorter than the incubation period and coincides with the onset of detectable viral shedding, not the later onset of symptoms. This knowledge is invaluable towards the development of models for the design of effective disease intervention strategies, such as the use of antivirals.

Acknowledgments

The authors wish to thank Dr. Michelle Cretikos of the New South Wales Department of Health, and Dr. Ping Yan and Dena Schanzer of the Public Health Agency of Canada for informative discussions related to this work.

We are also grateful to the reviewers for their many insightful comments and careful review of the analysis and manuscript.

This research is partially supported by NSF grant DMS-1022758 (ST) and by the Fogarty International Center, NIH (GC).

References

Alonso, W.J., et al., 2007. Seasonality of influenza in Brazil: a traveling wave from the Amazon to the SubTropics. Am. J. Epidemiol. 165, 1434–1442.

Bacaer, N., 2007. Approximation of the basic reproduction number R0 for vector-borne diseases with a periodic vector population. Bull. Math. Biol. 69, 1067–1091.

Bacaer, N., Abdurahman, X., 2008. Resonance of the epidemic threshold in a periodic environment. J. Math. Biol. 57, 649–673.

Bacaer, N., Dads, E.H.A., 2010. Genealogy with seasonality, the basic reproduction number, and the influenza pandemic. J. Math. Biol. 62 (5), 741–762.

Bacaer, N., Gomes, M.G.M., 2009. On the final size of epidemics with seasonality. Bull. Math. Biol. 71, 1954–1966.

Carrat, F., et al., 2008. Time lines of infection and disease in human influenza: a review of volunteer challenge studies. Am. J. Epidemiol. 167 (7), 775–785.

Cauchemez, S., et al., 2008. Estimating the impact of school closure on influenza transmission from Sentinel data. Nature 452, 750–754.

Cauchemez, S., et al., 2009. Closure of schools during an influenza pandemic. Lancet Infect. Dis. 9 (8), 473–481.

Chao, D.L., Halloran, E.M., Longini Jr., I.M., 2010. School opening dates predict pandemic influenza A(H1N1) outbreaks in the United States. J. Infect. Dis. 202, 877–880.

Chowell, G., Echevarría-Zuno, S., Viboud, C., Simonsen, L., Tamerius, J., et al., 2011. Characterizing the epidemiology of the 2009 Influenza A/H1N1 Pandemic in Mexico. PLoS Med. 8 (5), e1000436, http://dx.doi.org/10.1371/journal.pmed.1000436.

CIA World Factbook, 〈https://www.cia.gov/library/publications/the-world-factbook/geos/ci.html〉 (accessed Nov, 2011).

Cowan, G., 1998. Statistical Data Analysis. Oxford Science Publications.

Dietz, K., 1976. The incidence of infectious diseases under the influence of seasonal fluctuations. Mathematical Models in Medicine. Lecture Notes in Biomathematics, vol. 11. Springer-Verlag, New York, pp. 1–15.

Dowell, S.F., 2001. Seasonal variation in host susceptibility and cycles of certain infectious diseases. Emerg. Infec. Dis. 7 (3), 369–374.

Dushoff, J., Plotkin, J.B., Viboud, C., Earn, D.J., Simonsen, L., 2006. Mortality due to Influenza in the United States. Am. J. Epidemiol. 163 (2), 181–187.

Elder, A.G., O'Donnell, B., McCruden, E.A., et al., 1996. Incidence and recall of influenza in a cohort of Glasgow healthcare workers during the 1993–4 epidemic: results of serum testing and questionnaire. Br. Med. J. 313, 1241–1242.

Feng, Z., Towers, S., Yang, Y., 2011. Modeling the effects of vaccination and treatment on pandemic influenza. Am. Assoc. Pharm. Sci. J. 13 (3), 427–437.

Ferguson, N.M., Galvani, A.P., Bush, R.M., 2003. Ecological and immunological determinants of influenza evolution. Nature 422, 428–433.

Fraser, C., et al., 2009. Pandemic potential of a strain of influenza A(H1N1): early findings. Science 1557, 324.

Ghani, A.C., Baguelin, M., Griffin, J., et al., 2009. The early transmission dynamics of H1N1pdm influenza in the United Kingdom. PLoS Curr., RRN1130.

Hens, J., et al., 2009. Estimating the impact of school closure on social mixing behaviour and the transmission of close contact infections in eight European countries. BMC Infectious Diseases 2009 (9), 187.

Hethcote, W.H., 1976. Qualitative analysis of communicable disease models. Math. Biosci. 28 (3–4), 335–356.

Hethcote, H.W., Levin, S.A., 1988. Periodicity in epidemiological models. In: Levin, S.A., Hallam, T.G., Gross, I. (Eds.), Applied Mathematical Ecology. Biomathematics, vol. 18. Springer, Berlin, Heidelberg, New York, pp. 193–211.

King Jr., J.C., Haugh, C.J., Dupont, W.D., Thompson, J.M., Wright, P.F., Edwards, K.M., 1988. Laboratory and epidemiological assessment of a recent influenza B outbreak. J. Med. Virol. 25 (3), 361–368.

Lessler, J., et al., 2009. Outbreak of 2009 pandemic influenza A(H1N1) at a New York City school. N. Engl. J. Med. 361, 2628–2636.

Lofgren, E., Fefferman, N.H., Naumov, Y.N., Gorski, J., Naumova, E.N., 2007. Influenza seasonality: underlying causes and modeling theories. J. Virol. 81 (11), 5429–5436.

Longini, I.M., Halloran, M.E., Nizam, A., Yang, Y., 2004. Containing pandemic influenza with antiviral agents. Am. J. Epidemiol. 159, 623–633.

Lowen, A.C., Mubareka, S., Steel, J., Palese, P., 2007. Influenza virus transmission is dependent on relative humidity and temperature. PLoS Pathog. 4 (10), 151–158.

Medlock, J., Galvani, A.P., 2009. Optimizing influenza vaccine distribution. Science 325 (5948), 1705–1708.

Miller, M.A., Viboud, C., Balinska, M., Simonsen, L., 2009. The Signature Features of Influenza Pandemics. N. Engl. J. Med. 360 (25), 2595–2598.

Moneim, I.A., 2007. Seasonally varying epidemics with and without latent period: a comparative simulation study. Math. Med. Biol. 24, 1–15.

Moser, M.R., Bender, T.R., Margolis, H.S., Noble, G.R., Kendal, A.P., Ritter, D.G., 1979. An outbreak of influenza aboard a commercial airliner. Am. J. Epidemiol. 110, 1–6.

Mossong, J., et al., 2008. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 5 (3), e74, http://dx.doi.org/10.1371/journal.pmed.0050074.

Myung, J., 2003. Tutorial on maximum likelihood estimation. J. Math. Psychol. 47, 90–100.

New York Department of Health, Hospital Emergency Departments of New York State: 2005 Emergency Department Visits by Condition and Day of Week, 〈http://www.health.ny.gov/statistics/sparcs/ed/2005/table14.htm〉 (accessed Nov, 2011).

Nishiura, H., Wilson, N., Baker, M.G., 2009. Estimating the reproduction number of the novel influenza A virus (H1N1) in a Southern Hemisphere setting: preliminary estimate in New Zealand. J. N. Z. Med. Assoc. 122 (1299), 73–77.

Reed, C., et al., 2009. Estimates of the prevalence of pandemic (H1N1) 2009, United States, April–July 2009. Emerg. Infect. Dis. 15 (12), 2004–2007.

Shaman, J., Pitzer, V.E., Viboud, C., Grenfell, B.T., Lipsitch, M., 2010. Absolute humidity and the seasonal onset of influenza in the continental United States. PLoS Biol. 8 (2), e1000316.

Shi, P., Keskinocak, P., Lee, B.Y., 2010. Modelling seasonality and viral mutation to predict the course of an influenza pandemic. Epidemiol. Infect. 138, 1472–1481.

Torres, J.P., et al., 2010. Impact of the novel influenza A (H1N1) during the 2009 autumn–winter season in a large hospital setting in Santiago, Chile. Clin. Inf. Dis. 2010 (50), 860–868.

Towers, S., Feng, Z., 2009. Pandemic H1N1 influenza: predicting the course of a pandemic and assessing the efficacy of the planned vaccination programme in the United States. Eurosurveillance 14 (41).

Towers, S., Vogt Geisse, K., Zheng, Y., Feng, Z., 2011. Antiviral treatment for pandemic influenza: assessing potential repercussions using a seasonally forced SIR Model. J. Theor. Biol. 289, 259–268.

Towers, S., VogtGeisse, K., Chia-Chun, T., Han, Q., Feng, Z., 2012. The impact of school closures on pandemic influenza: assessing potential repercussions using a seasonal SIR Model. Math. Biosci. Eng. 9 (2), 413–430.

Tuite, A.R., et al., 2010. Estimated epidemiologic parameters and morbidity associated with pandemic H1N1 influenza. Can. Med. Assoc. J. 182 (2), 131–136.

Wang, W., Zhang, X-Q., 2007. Threshold dynamics for compartmental epidemic models in periodic environments. J. Dyn. Differential Equation 20, 699–717.

Williams, B.G., Dye, C., 1997. Infectious disease persistence when transmission varies seasonally. Math. Biosci. 145 (1), 77–88.

Yang, Y., et al., 2009. The transmissibility and control of pandemic influenza A(H1N1) Virus. Science 326 (5953), 729–733.