Epidemics 3 (2011) 135–142

Contents lists available at ScienceDirect

Epidemics

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

Comparing three basic models for seasonal influenza

Stefan Edlund [a,*], James Kaufman [a], Justin Lessler [b], Judith Douglas [a], Michal Bromberg [c], Zalman Kaufman [c], Ravit Bassal [c], Gabriel Chodick [d], Rachel Marom [d], Varda Shalev [d], Yossi Mesika [e], Roni Ram [e], Alex Leventhal [f]

[a] IBM Almaden Research Center, 650 Harry Road, San Jose, CA 95120, USA
[b] Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, USA
[c] Israel Center for Disease Control, Gertner Institute, Sheba Medical Center, Tel Hashomer 52621, Israel
[d] Maccabi Healthcare Services, 27 Ha'Merad Street, Tel Aviv 50493, Israel
[e] IBM Haifa Research Labs, Haifa University Campus, Mount Carmel, Haifa 31905, Israel
[f] Israel Ministry of Health, 20 King David Street, Jerusalem 91010, Israel

ARTICLE INFO

Article history:
Received 4 December 2009
Received in revised form 13 April 2011
Accepted 13 April 2011
Available online 13 May 2011

Keywords:
Simulation
Compartmental disease models
Predictive validity
Epidemics

ABSTRACT

In this paper we report the use of the open source Spatiotemporal Epidemiological Modeler (STEM, www.eclipse.org/stem) to compare three basic models for seasonal influenza transmission. The models are designed to test for possible differences between the seasonal transmission of influenza A and B. Model 1 assumes that the seasonality and magnitude of transmission do not vary between influenza A and B. Model 2 assumes that the magnitude of seasonal forcing (i.e., the maximum transmissibility), but not the background transmission or flu season length, differs between influenza A and B. Model 3 assumes that the magnitude of seasonal forcing, the background transmission, and flu season length all differ between strains. The models are all optimized using 10 years of surveillance data from 49 of 50 administrative divisions in Israel. Using a cross-validation technique, we compare the relative accuracy of the models and discuss the potential for prediction. We find that accounting for variation in transmission amplitude increases the predictive ability compared to the base. However, little improvement is obtained by allowing for further variation in the shape of the seasonal forcing function.

Introduction

In temperate regions, seasonal influenza epidemics occur every year, generally peaking in the winter and virtually disappearing during summer. Reasons behind the seasonal variation of influenza are not clearly understood, but factors such as temperature and absolute humidity, indoor crowding, and variations in host immune response have been put forward as potential explanations (Lipstich and Viboud, 2009; Lowen et al., 2007; Shaman and Kohn, 2009). Because influenza virus is constantly adapting through antigenic drift, immunity is not permanent as it is with some other viral diseases such as measles (Smith et al., 2004). Individuals typically become at least partially susceptible to new variants of the virus within a few years. Consequently, influenza falls into a category of diseases that can be studied with a compartmental model representing the passage of individuals between three states, Susceptible (S), Infectious (I) and Recovered (R), and eventually returning to the Susceptible state as immunity wanes (an SIRS model).

In this paper we compare three models for influenza based on 10 years of seasonal influenza data collected by the Israeli Centers for Disease Control for 49 locations in Israel. Over this period, influenza A was the most frequently isolated strain in 8 of the years and influenza B was the most frequently isolated strain in 2 of the years. Each of these models represents a hypothesis about differences in transmission dynamics between influenza A and influenza B dominant years. In the first model we consider there to be identical transmission dynamics in influenza A and influenza B dominant years. The second model allows peak transmissibility (as characterized by the seasonally forced reproductive number) to vary between A and B dominant years. The third model also allows variation in the duration of the influenza season and strength of seasonality. Each of these hypotheses is captured by the parameterization of a deterministic SIRS meta-population (patch) model representing the population and geography of Israel. Each model is fit to the 10 years of reference data using a Nelder–Mead simplex (Nelder and Mead, 1965), and compared based upon predictive performance using cross validation. The epidemiological parameters obtained from the best-fit models are compared in detail with values from the literature. (Anderson and May, 1991; Arino and van den Driessche, 2003; Colizza et al., 2007; Crépey and

* Corresponding author. Tel.: +1 408 927 1766.
E-mail address: edlund@almaden.ibm.com (S. Edlund).

1755-4365/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.epidem.2011.04.002

Barthémey, 2007; Fulford et al., 2002; Keeling and Rohani, 2008; Ken et al., 2009; Longini et al., 2005; Molinari et al., 2007; Riley, 2007; Bootsma and Ferguson, 2007; Carrat et al., 2006; Flahault et al., 1988; Pease, 1987; Hufnagel et al., 2004; Keeling et al., 2001; Longini, 1988; Mills et al., 2004; Rvachev and Longini, 1985; Sattenspiel and Dietz, 1995; Sattenspiel and Herring, 1998, 2003).

Methods

Data

The Israel Center for Disease Control (ICDC), Ministry of Health, operates an ongoing influenza surveillance system that is based on various clinical and virological data sources. Maccabi Health Care Services, the second largest health maintenance organization in Israel that serves about 25% of the Israeli population (approximately 1,700,000 members), provides the ICDC's surveillance system with daily data regarding visits of its members to outpatient clinics for various diagnoses, including influenza like illness (ILI). Each record includes the patient's area of residence. The method used to obtain the data was described previously (Kaufman et al., 2007). For the purpose of this study, the incidence of influenza was considered to be proportional to the number of ILI visits reported by physicians in Maccabi Health Care Services. Physician reported ILI visits appear to match fairly well to measures of influenza activity based on virological data gathered since 1996 by the network of sentinel community-based clinics operated by ICDC during the winter months. This analysis is based upon 10 years of aggregate daily ILI visits. Areas of residence were mapped to 50 "natural regions" in Israel, one of which, the Dead Sea Region, has essentially zero population and was excluded from analysis. The division of the State of Israel into administrative districts, sub-districts, and natural regions is published by the Israeli Central Bureau of Statistics (CBS, 2008).

The number of visits due to ILI represents only a fraction of the incidence. We assume that reported ILI represents a constant proportion (the reporting fraction) of actual influenza incidence. Since the actual (average) reporting fraction is unknown, we fit the aggregate data for all of Israel with several reporting fractions in the range 0.5–5% and compare the resulting estimates of epidemiological parameters to the literature (see the 'Results' section). A reporting fraction on the order of 3% yields parameters consistent with the literature and is used throughout subsequent analysis.

The rate of immigration (through birth and otherwise) and the rate of emigration (through death and otherwise) were estimated using data from the Israeli Central Bureau of Statistics (CBS). Immigration and emigration were considered to be constant across the 49 natural regions.

Transmission model

We model influenza transmission using a geographically explicit SIR(S) compartmental model with a seasonally modulated transmission coefficient $\beta(t)$ (Anderson and May, 1991). Eqs. (1a)–(1c) describe the model.

$$\frac{dS_j(t)}{dt} = -\beta(t)S_j(t)\frac{\sum_{k=1}^{K} m_{jk}I_k(t)}{\sum_{k=1}^{K} m_{jk}P_k(t)} + \alpha R_j(t) + bP_j(t) - \mu S_j(t) \quad (1a)$$

$$\frac{dI_j(t)}{dt} = \beta(t)S_j(t)\frac{\sum_{k=1}^{K} m_{jk}I_k(t)}{\sum_{k=1}^{K} m_{jk}P_k(t)} - \gamma I_j(t) - \mu I_j(t) \quad (1b)$$

$$\frac{dR_j(t)}{dt} = \gamma I_j(t) - \alpha R_j(t) - \mu R_j(t) \quad (1c)$$

where S_j, I_j, and R_j are the size of the susceptible, infectious, and recovered populations at location j; $\beta(t)$ is the seasonal transmission coefficient; P_j is the population at location j; γ is the recovery rate; α is the immunity loss rate; b is the birth rate; μ is the death rate; and m_{jk} is the rate of mixing between locations j and k (when $j = k$, $m_{jk} = 1$).

The seasonal transmission coefficient, $\beta(t)$, is modeled as a mixture of Gaussian distributions. We prefer this to the sinusoidal function often used to model seasonal forcing (e.g., $\beta(t) = \beta_1 \sin(\omega t + \phi) + \beta_0$) because it provides more flexibility in controlling the duration of the flu season. We define $\beta(t)$ as:

$$\beta(t) = \beta_i \left\{ \sum_i \left[\rho_i + (1 - \rho_i)\frac{g(t - t_i, \sigma_i)}{g(0, \sigma_i)}\xi(t, t_i) \right] \right\} \quad (2)$$

where i is the index of the yearly influenza season; β_i is the peak transmission coefficient for season i; $(1 - \rho_i)$ is the fraction by which transmission varies seasonally in season i ($0 < \rho_i < 1.0$); t_i is the day of peak transmission for season i; $g(x, \sigma_i)$ is the probability distribution function of a Gaussian distribution with mean 0 and standard deviation σ_i; and $\xi(t, t_i)$ is the mixing function. The day of peak transmission, t_i, is defined so that the transmission function has a period of exactly 1 year, with constant phase, φ. That is transmission peaks φ days after January 1 of season i. The joining function, $\xi(t,t_i)$, smoothly joins adjacent Gaussians so that $\beta(t)$ is only dependent upon a single Gaussian during the period of season i where most transmission occurs (see supplemental material).

Using this transmission model, we can specify our three hypotheses in terms of β_i, ρ_i and σ_i. That is:

- **Hypothesis 0**: Transmission is identical in both A and B dominant years ($\beta_i = \beta_0$, $\sigma_i = \sigma_0$, $\rho_i = \rho_0$ for all years).
- **Hypothesis 1**: Only the maximum rate of transmission varies between A and B dominant years. That is, $\beta_i = \beta^A$ in A dominant seasons, and $\beta_i = \beta^B$ in B dominant seasons ($\sigma_i = \sigma_0$, $\rho_i = \rho_0$ for all seasons).
- **Hypothesis 2**: The maximum rate of transmission and the effect of seasonal forcing vary between influenza A and influenza B. That is $\beta_i = \beta^A$, $\sigma_i = \sigma^A$, $\rho_i = \rho^A$ in A dominant years and $\beta_i = \beta^B$, $\sigma_i = \sigma^B$, $\rho_i = \rho^B$ in B dominant years.

The control hypothesis (H0) requires seven independent parameters. Hypothesis H1 uses eight parameters, and hypothesis H2 allows two more degrees of freedom (ten independent parameters).

All models are subject to the (clearly false) assumption of cross protection between influenza A and B. We are unable to fit a full multi-serotype model that accounts for variations in all of the epidemiological parameters for each serotype and tracks the immune state of the population with respect to each serotype (Murphy and Clements, 1989). This would require *a priori* knowledge of cross strain immunity or data of greater detail than is available at this time.

The mixing parameter, m_{jk}, represents the fraction of infectious people in region j that individuals in region k are exposed to (see Eq. (1)). For the purposes of this study, we assume mixing only occurs between adjacent regions. Because not all administrative regions in the geographic meta-population model are the same area, the mixing rate is scaled by a characteristic mixing distance, δ_0, for each region. This parameter models the distance people travel, on average, in a day. In this model the fraction of people leaving site j for any neighboring site k becomes:

$$m_{jk} \sim \min\left(1.0, \frac{\delta_0}{\sqrt{A_j}}\right) \quad (3)$$

where the characteristic mixing distance, δ_0, is determined by model fitting.

The differential equations are numerically integrated using Runge Kutta Cash-Karp (RKCK) adaptive step size integration to a precision of $1::10^{-9}$ (Cash and Karp, 1990).

Model fitting

For each hypothesis above, we fit the model parameters using a Nelder–Mead Simplex algorithm. The Nelder–Mead algorithm searches model parameter space to minimize a functional error measurement, typically the root mean squared error. To make the error more easily interpretable across models we use the normalized root mean square error (NRMSE) between the incidence, I_s, predicted by the simulation and the observed incidence, I_r, obtained from the historic reference data (Ahlburg, 1992; Hamad, 2006). The Nelder–Mead optimization is completely independent of normalization as there is only one normalization constant for the data set. We calculate the NRMSE as the root mean squared error over all time normalized by the difference between the maximum and minimum observed country-wide incidence.

$$\mathrm{NRMSE}(I_s, I_r) = \frac{1}{\max_t \sum_{j\in L} I_{r,j}(t) - \min_t \sum_{j\in L} I_{r,j}(t)} \times \sqrt{\frac{\sum_{t\in T}\left(\sum_{j\in L} I_{s,j}(t) - \sum_{j\in L} I_{r,j}(t)\right)^2}{|T|}} \quad (4)$$

where L is the set of all locations and T is the set of all times for which we have observations (~3650 time steps).

Model comparison

It is not appropriate to compare model performance by comparing the relative NRMSE on the data used to fit the models, as models with more free parameters (i.e., H1 and H2) will generally have lower NRMSE. To compare the performance of the three models based upon their performance in predicting data not used in model fitting, we use a cross validation framework (Duda et al., 2001). Cross validation is different from a typical biostatistical approach. As opposed to using probabilistic measures of statistical fit, cross validation is a measure of the deterministic "predictive" power of a model or method that tests how well a model performs against unseen data.

Each model is fit ten times holding out a different influenza season each time, fitting the parameters only considering the NRMSE in the other nine influenza seasons. After each fit, we compare model predictions against the observed data in the withheld season. We then calculate the summary NRMSE for each model as the average NRMSE across the ten predictions. The model with the lowest summary NRMSE is considered to have the best predictive ability of the three and hence represent the hypothesis most supported by the data.

Spatiotemporal Epidemiological Modeler

All models were implemented using the Spatiotemporal Epidemiological Modeler (STEM). STEM is an open source framework for building and evaluating spatiotemporal epidemiological models (Kaufman et al., 2008). STEM has denominator data for the entire world, provides users with a choice of solvers, i.e., a fast finite difference algorithm or Runge Kutta Cash-Karp (RKCK) adaptive step size integration (Cash and Karp, 1990), and implements the Nelder–Mead simplex optimization in an automated experiment feature (Edlund et al., 2009). As an open source project under the Eclipse Foundation, STEM is meant to support collaboration within a community. All the mathematical models used in the creation of this work are available as open contributions to STEM and can be freely used and extended. STEM provides a variety of models, tools for creating models, and analytical components and makes them available to any researcher interested in studying influenza or other infectious diseases at www.eclipse.org/stem.

Results

To determine the reporting fraction to use in the subsequent analysis, we fit the base (H0) SIR(S) model to the aggregate data for all of Israel with various reporting fractions. We found that a reporting fraction of 3% led to parameters estimates most consistent with the published literature. The parameter most sensitive to the assumed reporting fraction was the duration of immunity ($1/\alpha$), which increased with increasing reporting (Table S1). Table 1 summarizes parameter estimates from complete model fits using a reporting fraction of 3% in comparison to values reported in the literature.

To explore differences in transmission between influenza A and B dominant years, we fit each of the three models against all 10 years of historic data using Nelder–Mead optimization (Fig. 1). The incidence predicted by each model fits the data with a NRMSE less than 10% in all three cases (Table 2). Model H2 has the lowest NRMSE. However, in all models incorporating differences between influenza A and influenza B dominant years, we see off-season peaks not reflected in the data. In all models we estimate the peak reproductive number to be in the range $2.6 < R_0 < 2.8$ with a seasonal modulation of less than 10% (Table 2). Both model H1 and H2 estimate a lower reproductive number in influenza B years and model H2 suggests that B seasons are shorter with weaker seasonality; $1-\rho$ is 6% for B vs. 8% for A (Fig. 1B). The estimated length of immunity ranges from 4.9 years (model H1) to 5.9 years (model H0). The mean infectious period is estimated to be 1.1 days in all models, lower than typical literature values. The characteristic mixing distance is estimated at approximately 55 km in all models, that is full mixing occurs between neighboring regions less than $55^2 = 3025$ km^2 in area and the degree of mixing decreases for larger regions (Eq. (3)).

Fig. 2 shows a concatenation of the predicted incidence obtained for the unfitted years only in each of the ten cross validation trials. We found that NRMSE averaged across predictions was 0.244 (range 0.081,0.568) for model H0, 0.238 (range 0.105,0.556) for model H1, and 0.249 (range (0.085,0.559) for model H2. Based upon these results, model H1 (variation of only transmissibility between A and B dominant years) appears to be the best supported by the data. To ensure that our results were not dominated by prediction on any particular year, we also compared which model had the lowest NRMSE in each excluded year. We found that model H1 made the best prediction in 6 of 10 years, compared to models H0 and H2 each of which performed best in 2 years. The estimated parameters as well as the NRMSE results for cross validation of all three models are shown in Table 2.

The NRMSE between the unfitted model years and the reference is high whenever the predicted epidemic amplitude is incorrect *and/or* the timing of the epidemic is phase shifted from the reference (Fig. 2). If one examines the 2 years where influenza B was dominant, one can see the model for H2 (green) had the worst performance.

To assess the value of the spatially explicit model, we compared the parameter estimates obtained from the spatially explicit model (H1) with a model representing all of Israel as a single fully mixed population. We found significant differences in the estimated model parameters between the two approaches (e.g., a maximum R_0 of 3.0 for influenza A versus 2.8 in the spatially

Table 1
Comparison of estimates from the literature with model estimates for important epidemiological parameters of influenza.

Reported parameter values (literature)	Literature			Model		
	Lowest	Typical	Upper	H0	H1(A,B)	H2(A,B)
Mean period of infection [days]	1.0[a]	∼2[b]	4.8[c]	1.1	1.1	1.1
Mean period of immunity [years]	1[d]	∼4–6[e]	8[f]	5.9	4.9	5.4
Reproductive number	1.1[g]	1.5–2.0[h]	∼20[i]	2.6	2.8,2.7	2.6,2.6

[a] Gojovic et al. (2009) includes a latent period of ∼1.5 day.
[b] Flahault et al. (2009), Truscott et al. (2009), Longini et al. (2005).
[c] Carrat et al. (2008).
[d] Xia et al. (2005).
[e] Koelle et al. (2006), Couch and Kasel (1983).
[f] Dushoff et al. (2004).
[g] Basta et al. (2009).
[h] Halloran et al. (2008), Truscott et al. (2009).
[i] Mills et al. (2004).

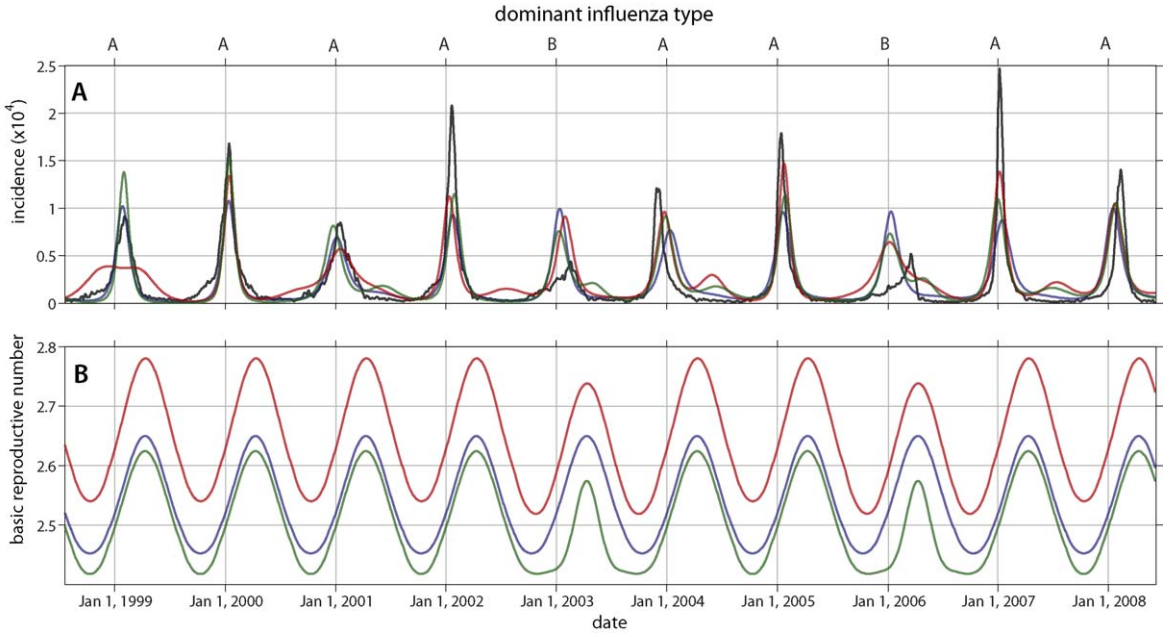

Fig. 1. (A) Actual (black) and estimated (colored) incidence of influenza in Israel; (B) the estimated time varying reproductive number, for each of three hypotheses: Blue represent hypothesis H0, no variation in transmission between influenza A and B dominant seasons. Red represents hypothesis H1, only the maximum rate of transmission varies between A and B dominant seasons. Green represents hypothesis H2, the maximum rate of transmission and the effect of seasonal forcing vary between influenza A and influenza B. The dominant strain for each year is indicated along the top of the figure.

Table 2
The best-fit epidemiological parameters and NRMSE for each hypothesis.

Parameter	H0		H1		H2	
	10 year fit	Cross valid range (low, high)				
Max $R_0(t)$	2.65	(2.63,2.66)	2.78	(2.64,2.80)	2.63	(2.62,2.78)
Max $R_0(t)$ (B)	n/a		2.74	(2.59,2.76)	2.58	(2.53,2.73)
Modulation $1-\rho$	0.077	(0.066,0.086)	0.091	(0.077,0.098)	0.083	(0.078,0.104)
Modulation $1-\rho$ (B)	n/a		n/a		0.059	(0.049,0.063)
σ [years]	0.19	(0.15,0.20)	0.20	(0.19, 0.22)	0.21	(0.20,0.25)
σ (B) [years]	n/a		n/a		0.10	(0.08,0.13)
Length of immunity $(1/\alpha)$ [years]	5.87	(5.29,6.36)	4.87	(4.74,5.03)	5.39	(4.70,5.74)
Infectious period $(1/\gamma)$ [days]	1.13	(1.11,1.14)	1.13	(1.07,1.14)	1.11	(1.11,1.15)
ϕ = day of peak transmission [day of year] {date[a]}	100 {Apr10}	(95,105)	102 {Apr12}	(99,107)	100 {Apr10}	(97,106)
Characteristic mixing distance δ_0 [km]	49	(50,55)	55	(55,62)	55	(55,62)
NRMSE	0.092	(0.081,0.100)	0.082	(0.077,0.095)	0.083	(0.065,0.081)
NRMSE (cross validation)	0.244[b]	(0.081,0.568)	0.238[b]	(0.105,0.556)	0.249[b]	(0.085,0.559)

The best fit is obtained by fitting all 10 years of historic data. In parentheses the table shows the range of values observed in each of ten cross validation runs (fitting only 9 years of data).
[a] The phase is shown in terms of day of year (DOY 0 = Jan 1), and as calendar date in the first year.
[b] The average NRMSE over ten cross validation trials.

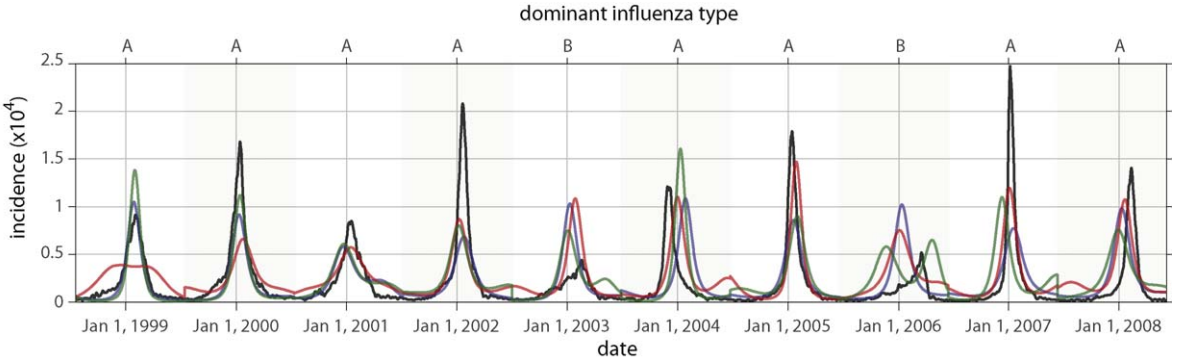

Fig. 2. Composite plot of predicted (held out) seasonal epidemics from the cross validation experiment. This plot combines the ten separate predictions into one graph for each hypothesis: H0 = blue, H1 = red, and H2 = green. Reference incidence is shown in black. Results from fitted years are not shown.

explicit model). Further comparison of the spatial and non-spatial H1 models may be found in the supplemental material.

Discussion

In this paper we compared models based on three hypotheses about the effects of the dominant circulating serotype on the seasonal transmission of influenza: identical transmission dynamics in influenza A and B dominant year (H0), differing amplitude of seasonal forcing in A and B dominant years (H1), and differing amplitude, width, and modulation of seasonal forcing in A and B dominant years (H2). The models were optimized against 10 years of historic ILI data from Israel. Based on its ability to predict withheld years in a cross validation experiment, we found Hypothesis H1 to be best supported by our analysis with 24% error (NRMSE) on average. That is, our analysis suggests a small reduction (~2%) in the peak transmissibility of influenza in B dominant years and no other differences in the shape of the seasonal forcing function.

The epidemiological parameters obtained from all three optimized models are within the range expected for seasonal influenza (see Table 1). Reported values for the mean infectious period have been based on studies of viral shedding as well as measurements of the average generation time or serial interval for the virus. The generation time is equivalent to the infectious period *if the latent period is 0 days* (the SIRS model approximation). Some authors used an SEIR(S) model with latent periods ~1.5 days. Flahault et al. (2009) report a mean generation time of 2 days; Truscott et al. (2009) find a similar result (2.2 days). Longini et al. (2005) estimate the mean infectious period at 1.9 days. Nicholson et al. (1998) observe viral shedding over 4.1 days; Cauchemez et al. (2004) report 3.8 days; Carrat et al. (2008) find 4.8 days; and Cauchemez et al. (2008) use 2.4 days. Ferguson et al. (2005) estimate the period of infection (with 95% confidence) between 2.1 and 3.0 days. Carrat et al. (2006) model influenza with an infectivity distribution that begins 0.5 days after infection, peaks at 2.5 days, and extends beyond 8 days. In a study of over 10,000 individuals, Gojovic et al. (2009) report that, on average, asymptomatic individuals were infectious for 1 day (following a 1.5 day latent period) while symptomatic individuals were infectious for 1.5 days. Our estimate is at the very low end of this range; to the extent that it is inconsistent with laboratory studies (Carrat et al., 2006), it may reflect the approximations inherent in using a simple SIR(S) compartmental model. We did explore the use of an SEIR(S) model as an alternative for this optimization. Introducing a latent period (of 1.5 days) does not have a significant effect on the other parameters reported here. Since the latent period is small, we decided to conduct the following analysis based on an SIR(S) model to minimize the total number of parameters being optimized.

The duration of immunity predicted by all our models is in the middle of the range found in the literature. The period over which recovered individuals reenter the susceptible state is in large part due to influenza strain dynamics and the rate of antigenic drift of the influenza virus (Smith et al., 2004). Koelle et al. (2006) found that, between 1968 and 2003, clusters with unique antigenic properties were replaced on a timescale of 2–8 years. Couch and Kasel (1983) concluded the average duration of immunity was longer than 4 years. A case study by Xia et al. (2005) concluded that, after an initially fast decay (on a timescale of 8–12 months), a low level of immunity persisted from 4 to 5 years. Dushoff et al. (2004) concluded a reasonable range for the average duration of immunity was 4–8 years.

The basic reproductive numbers estimated from our models are at the high end of estimates for seasonal influenza in the literature. Basta et al. (2009) reports R_0 in the range $1.1 < R_0 < 1.6$ for seasonal influenza with lower values corresponding to milder seasons. Flahault et al. (2009) report R_0 no greater than 1.5 for seasonal observations in the city of Paris. Halloran et al. (2008) suggest historic values in the range $1.5 < R_0 < 1.9$, and Cauchemez et al. (2008) suggest $1.2 < R_0 < 1.4$. Estimates for pandemic influenza are in a similar range or slightly higher (Lessler et al., 2007). For example, Mills et al. (2004) report estimates for the 1918 flu in the range of $1.6 < R_0 < 20.0$. Truscott et al. (2009) estimate the range at $1.7 < R_0 < 2.1$. Mathews et al. (2007) report $R_0 \sim 3$ with an effective reproductive number near 2.0.

Depending on the method used to estimate the reproductive number, it is not always clear the extent to which it is the basic reproductive number (R_0) or the effective reproductive number (R) being measured. While our estimates for R_0 are high, this reflects the high levels of population immunity at the time of peak $R_0(t)$ predicted by our models. The *effective* reproductive number is never estimated to go above 1.1 in our model and the time at which it peaks is more consistent with perceptions about when the majority of influenza transmission actually occurs (Fig. 3). The effective reproductive number is defined as:

$$R_e(t) = s(t)\,R_0(t) \tag{5}$$

where $s(t)$ is the fraction of the population susceptible. The effective reproductive number peaks in phase with the susceptible fraction but out of phase with $R_0(t)$. Similar behavior is reported by Mathews et al. (2007) for a simple sinusoidal forcing function. Certainly the peak $R_e(t)$ and the peak incidence must occur where the basic reproductive number is rising sharply. The late peak of $R_0(t)$ (in early April) and its correspondingly late nadir (early October) may also reflect an attempt by the model to fit a seasonal forcing function that is actually more "square" in shape than a Gaussian. One might expect such a shape if school terms are the dominant driver of seasonality.

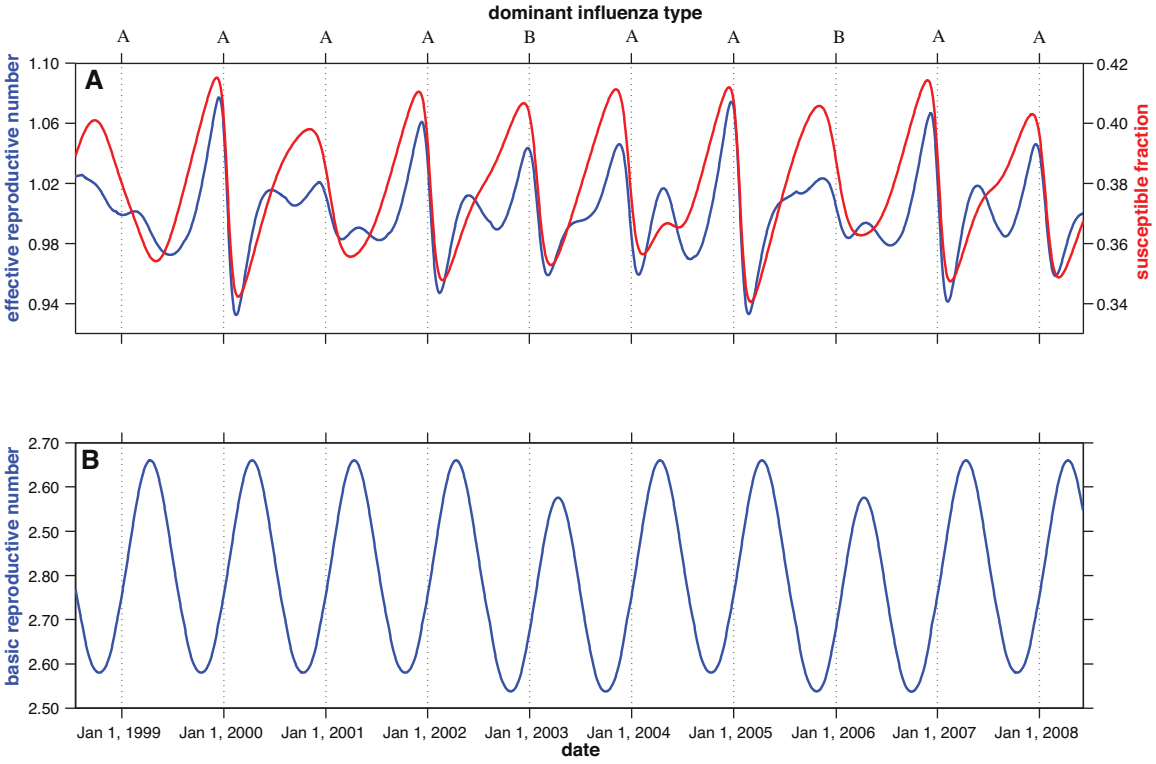

Fig. 3. (A) Variation of the effective reproductive number $R_e(t)$ for model H1 (blue) and the fraction of the population susceptible (red) vs time. (B) The basic reproductive number for the same model. Note that the peak $R_e(t)$ and the peak incidence occur where the basic reproductive number is rising sharply.

The Nelder–Mead optimization also provides an estimate for the characteristic mixing distance, $\delta_0 \sim 55$ km, used to model the circulation of populations between regions. Conceptually, we can think of this parameter as representing the distance people travel, on average, in a day. Hence, as the area of a patch gets larger than δ_0^2, a smaller fraction of the population will mix with populations in neighboring patches. It is informative to compare δ_0 with known geographic distances and commute patterns in the state of Israel. Residents of Israel often commute between the largest Israeli cities with in/out commute rates as high as 60% for some regions (Presman and Arnon, 2006). The driving distance from Haifa to Jerusalem (131 km) is just over twice the characteristic mixing distance found in this study. The average distance between Tel Aviv and all other districts in Israel is ~63 km (Presman and Arnon, 2006). These numbers are consistent with the derived mixing distance, and with a relatively high circulation of the population within the state of Israel.

Given the high circulation of the population within the state of Israel, it is reasonable to compare the current results with a purely non-spatial analysis. In the supplemental material we also evaluate Hypothesis H0 with a model that treats the entire nation of Israel as a single well mixed region. The Nelder–Mead optimization works equally well for the spatial and non-spatial models fitting the data with a NRMSE of 9%. However, when the (fixed) epidemiological parameters deduced from the non-spatial analysis are applied to each independent region, the prediction fails completely (NRMSE ~ 30%).

The models used involve several assumptions that may influence the accuracy of the transmission dynamics implied by this model. The Nelder–Mead optimization minimizes an error function in order to fit model parameters. Given that the period of incubation for influenza is small, we chose an SIR(S) model instead of an SEIR(S) model to limit the number of free parameters. The seasonal forcing function may neither be smooth or symmetrical, and it is unlikely that there is a constant probability of leaving the infectious or recovered compartments. However, the models to fit the three hypotheses are subject to the same limitations (in fact Model 3 contains all sub-models), hence major differences in the transmission dynamics should be apparent despite these limitations. More worrisome are possible year-to-year differences in the proportion of cases detected, whether due to changes surveillance or differences in symptom severity between years (or between influenza A and B themselves). Differences in reporting between influenza A and B could also lead to apparent differences in the reproductive number between years not reflective of changes in transmission dynamics. Misspecification of the reporting fraction in general would lead to misestimates of both the reproductive rate and the period of influenza immunity. Using a 3% reporting fraction (used by Israeli public health officials) does lead to parameter estimates consistent with the literature. Finally, the assumption of long term cross-protection between influenza A and B (and between A/H1N1 and A/H3N2 for that matter) is clearly incorrect, and some of the differences seen between A and B positive years may be a function of this assumption. An obvious next step in this work is to attempt to fit a full multi-strain model of influenza; however it is unclear if the available data is adequate to fit such a model.

The modeling paradigm reported here involved optimization of several models of influenza to a historic data set, and then using the control model as a standard against which to test hypotheses and implement improvements. Using this paradigm, we find support for differences in the reproductive number between influenza A and B dominant years. However, we do not suggest that *any* of the models evaluated are sufficient for influenza forecasting. Rather we hope to demonstrate a paradigm and a tool for the rapid development, fitting, and evaluation of epidemiological models. STEM provides world-wide denominator data, mechanisms for running spatially explicit compartmental models, and support for computational experiments making it an excellent starting point for

such endeavors. With its use of the Eclipse Equinox Framework (www.eclipse.org), STEM is both customizable and extensible and is intended to support collaborations in the modeling community. A paradigm that allows a community to develop, test, and compare epidemiological models in general could accelerate advances in the field.

Acknowledgments

This project was developed under Contract Number FA7014-07-C-0004, with the U.S. Air Force Surgeon General's Office (AF/SG) and was administered by the Air Force District of Washington (AFDW). The Air Force has not accepted the products depicted and issuance of a contract does not constitute Federal endorsement of the IBM Almaden Research Center.

The authors also thank Eclipse for its support of the STEM Project; the Israeli Center for Disease Control, the Israel Ministry of Health, and the Maccabi Health Care Service for providing the data used to validate the model and Professor Dani Cohen, Tel Aviv University.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.epidem.2011.04.002.

References

Ahlburg, D.A., 1992. Error measures and the choice of a forecast method. Int. J. Forecast. 8 (1), 99–100.

Anderson, R.M., May, R.M., 1991. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford, New York.

Arino, J., van den Driessche, P., 2003. A multi-city epidemic model. Math. Popul. Stud.: Int. J. Math Demogr. 10 (3), 1785–1793.

Basta, N.E., Chao, D.L., Halloran, M.E., Matrajt, L., Longini, I.M., 2009. Strategies for pandemic and seasonal influenza vaccination of school children in the United States. Am. J. Epidemiol. 170 (6), 679–686.

Bootsma, M.C., Ferguson, N.M., 2007. The effect of public health measures on the 1918 influenza pandemic in U.S. cities. Proc. Natl. Acad. Sci. U.S.A. 104 (18), 7588–7593.

Carrat, F., Luong, J., Lao, H., Sallé, A., Lajaunie, C., Wackernagel, H., 2006. A 'small-world-like' model for comparing interventions aimed at preventing and controlling influenza pandemics. BMC Med. 4 (2), doi:10.1186/1741-7015-4-26, PMID 17059593.

Carrat, F., Vergu, E., Ferguson, N.M., Lemaitre, M., Cauchemez, S., Leach, S., 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, doi:10.1093/aje/kwm375.

Cash, J.R., Karp, A.H., 1990. A variable order Runge-Kutta method for initial value problems with rapidly varying right-hand sides. ACM Trans. Math. Softw. 16, 201–222, doi:10.1145/79505.79507.

Cauchemez, S., Carrat, F., Viboud, C., Valleron, A.J., Boëlle, P.Y., 2004. A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. Stat. Med. 23 (22), 3469–3487, doi:10.1002/sim.1912.

Cauchemez, S., Valleron, A.J., Boëlle, P.Y., Flahault, A., Ferguson, N.M., 2008. Estimating the impact of school closure on influenza transmission from sentinel data. Nature 452 (7188), 750–754, doi:10.1038/nature06732.

Central Bureau of Statistics (CBS), 2008. See http://gis.cbs.gov.il/shnaton59/all_israel2008.pdf and http://gis.cbs.gov.il/shnaton59/keymap_e.htm.

Colizza, V., Barrat, A., Barthelemy, M., Valleron, A.-J., Vespignani, A., 2007. Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. PLoS Med. 4 (1), e13, doi:10.1371/journal.pmed.0040013.

Couch, R.B., Kasel, J.A., 1983. Immunity to influenza in man. Annu. Rev. Microbiol. 37, 529–549.

Crépey, P., Barthémey, M., 2007. Detecting robust patterns in the spread of epidemics: a case study of influenza in the United States and France. Am. J. Epidemiol. 166, 1244–1251.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification, 2nd ed. Wiley Interscience, John Wiley & Sons, New York, NY.

Dushoff, J., Plotkin, J.B., Levin, S.A., Eam, J.D., 2004. Dynamic resonance can account for seasonality of influenza epidemics. Proc. Natl. Acad. Sci. U.S.A. 101 (48), 16915–16916.

Edlund, S., Bromberg, M., Chodick, G., Douglas, J., Ford, D., Kaufman, Z., Lessler, J., Marom, R., Mesika, Y., Ram, R., Shalev, V., Kaufman, J., 2009. A spatiotemporal model for influenza. In: HISA HIC 2009. Frontiers of Health Informatics, Canberra, Australia, August 19–21, 2009.

Ferguson, N.M., Cummings, D.A., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., Iamsirithaworn, S., Burke, D.S., 2005. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature 437 (7056), 209–214, doi:10.1038/nature04017.

Flahault, A., Letrait, S., Blin, P., Hazout, J., Ménarés, J., Valleron, A.J., 1988. Modelling the 1985 influenza epidemic in France. Stat. Med. 7, 1147–1155.

Flahault, A., Vergu, E., Boëlle, P.Y., 2009. Potential for a global dynamic of influenza A (H1N1). BMC Infect. Dis. 12 (9), 129.

Fulford, G.R., Roberts, M.G., Heesterbeek, J.A.P., 2002. The metapopulation dynamics of an infectious disease: tuberculosis in possums. Theor. Popul. Biol., Available at igitur-archive.library.uu.nl.

Gojovic, M.Z., Sander, B., Fishman, D., Krahn, M.D., Bauch, C.T., 2009. Modeling Mitigation Strategies for Pandemic (H1N1) 2009., doi:10.1503/cmaj.091641.

Halloran, M.E., Ferguson, N.M., Eubank, S., Longini, I.M., Cummings, D.A.T., Lewis, B., Xu, S., Fraser, C., Vullikanti, A., Germann, T.C., et al., 2008. Modeling targeted layered containment of an influenza pandemic in the United States. Proc. Natl. Acad. Sci. U.S.A. 105 (12), 4639–4644, doi:10.1073/pnas.070689105.

Hufnagel, L., Brockmann, D., Geisel, T., 2004. Forecast and control of epidemics in a globalized world. Proc. Natl. Acad. Sci. U.S.A. 101, 15124–15129.

Hamad, H., 2006. A new metric for measuring metamodels quality-of-fit for deterministic simulations. In: Proceedings of the Conference on Winter Simulation. ACM Press, New York, pp. 882–888.

Kaufman, J.H., Conant, J.L., Ford, D.A., Kirihata, W., Jones, B., Douglas, J.V., 2008. Assessing the accuracy of spatiotemporal epidemiological models. In: Zeng, D., et al. (Eds.), BioSecure 2008, LNCS, vol. 5354, pp. 143–154.

Kaufman, Z., Wong, W., Peled-Leviatan, T., Cohen, E., Lavy, C., Aharonowitz, G., Dichtiar, R., Bromberg, M., Havkin, O., Kokia, E., Green, M., 2007. Evaluation of a syndromic surveillance system using the WSARE algorithm for early detection of an unusual, localized summer outbreak of influenza B: implications for bioterrorism surveillance. Isr. Med. Assoc. J. 9 (1), 3–7.

Keeling, M.J., Rohani, P., 2008. Modeling Infectious Diseases in Humans and Animals. Princeton University Press, Princeton, NJ.

Keeling, M.J., Woolhouse, M.E.J., Shaw, D.J., Matthews, L., Chase-Topping, M., Haydon, D.T., Cornell, S.J., Kappey, J., Wilesmith, J., Grenfelt, B.T., 2001. Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. Science 294 (5543), 813–817, doi:10.1126/science.1065973.

Ken, T.D., Read, J.M., Edmunds, W.J., 2009. Epidemic prediction and control in weighted networks. Epidemics 1, 70–76, doi:10.1016/j.epidem.2008.12.001.

Koelle, K., Cobey, S., Grenfell, B., Pascual, M., 2006. Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. Science 316, 1898–1903.

Lessler, J., Cummings, D.A.T., Fishman, S., Vora, A., Burke, D.S., 2007. Transmissibility of swine flu at Fort Dix, 1976, J.R. Soc. Interface 4, 755–762, doi:10.1098/rsif.2007.0228.

Lipstich, M., Viboud, C., 2009. Influenza seasonality: lifting the fog. PNAS 106, 3645–3646.

Longini Jr., I.M., 1988. A mathematical model for predicting the geographic spread of new infectious agents. Math. Biosci. 90, 367–383.

Longini Jr., I.M., Nizam, A., Xu, S., Ungchusak, K., Hanshaoworakul, W., Cumming, D.A.T., Halloran, M.E., 2005. Containing pandemic influenza at its source. Science 309 (5737), 1083–1087.

Lowen, A.C., Mubareka, S., Steel, J., Palese, P., 2007. Influenza virus transmission is dependent on relative humidity and temperature. PLoS Pathog. 3 (10), e151, doi:10.1371/journal.ppat.0030151.

Mathews, J.D., McCaw, C.T., McVernon, J., McBrydey, E.S., McCaw, J.M., 2007. A biological model for influenza transmission: pandemic planning implications of asymptomatic infection and immunity. PLoS ONE 2 (11), e1220, doi:10.1371/journal.pone.0001220 (see Appendix S1).

Mills, C.E., Robins, J.M., Lipsitch, M., 2004. Transmissibility of 1918 pandemic influenza. Nature 432 (7019), 904–905.

Molinari, N.A., Ortega-Sanchez, I.R., Messonnier, M.L., Thompson, W.W., Wortley, P.M., Weintrau, E., Bridges, C.B., 2007. The annual impact of seasonal influenza in the US: Measuring disease burden and costs. Vaccine 25 (27), 5086–5096.

Murphy, B.R., Clements, M.L., 1989. The systemic and mucosal immune response of humans to influenza A virus. Curr. Top. Microbiol. Immunol. 146, 107–116.

Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. Comput. J. 7, 308–313.

Nicholson, K., Webster, R.G., Hay, A.J., 1998. Textbook of Influenza. Blackwell Science, Malden, MA.

Pease, C.M., 1987. An evolutionary epidemiological mechanism with applications to type A influenza. Theor. Popul. Biol. 31, 422–452.

Presman, N., Arnon, A., 2006. Commuting Patterns in Israel 1991–2004. Bank of Israel, www.bankisrael.gov.il/deptdata/mehkar/papers/dp0604e.htm.

Riley, S., 2007. Large-scale spatial-transmission models of infectious disease. Science 316 (5827), 1298–1301, doi:10.1126/science.1134695.

Rvachev, L.A., Longini, I.M., 1985. A mathematical model for the global spread of influenza. Math. Biosci. 75 (1), 3–22.

Sattenspiel, L., Dietz, K., 1995. A structured epidemic model incorporating geographic mobility among regions. Math. Biosci. 128 (1–2), 71–91, doi:10.1016/0025-5564(94)00068-B.

Sattenspiel, L., Herring, D.A., 2003. Simulating the effect of quantine on the spread of the 1918–19 flu in central Canada. Bull. Math. Biol. 65 (1), 1–26, doi:10.1006/bulm.2002.03171.

Sattenspiel, L., Herring, D.A., 1998. Structured epidemic models and the spread of influenza in the central Canadian subarctic. Hum. Biol. 70 (1), 91–115.

Shaman, J., Kohn, M., 2009. Absolute humidity modulates influenza survival, transmission, and seasonality. Proc. Natl. Acad. Sci. U.S.A., doi:10.1073/Pnas.0806852106.

Smith, D.J., Lapedes, A.S., de Jong, J.D., Bestebroer, T.M., Rimmelzwaan, G.F., Osterhaus, A.D.M.E., Fouchier, R.A.M., 2004. Mapping the antigenic and genetic evolution of influenza virus. Science 305, 371–376.

Truscott, J., Fraser, C., Hinsley, W., Cauchemez, S., Donnelly, C., Ghani, A., Ferguson, N., Meeyai, A., 2009. Quantifying the transmissibility of human influenza and its seasonal variation in temperate regions. PLoS Curr. Influenza, PMCID: PMC2771768.

Xia, Y., Gog, J.R., Grenfell, B.T., 2005. Semiparametric estimation of the duration of immunity from infectious disease time series: influenza as a case-study. Appl. Statist. 54 (3), 659–672.