Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/MBE

MBE, 16(5): 3753–3770.
DOI: 10.3934/mbe.2019186
Received: 21 December 2018
Accepted: 09 April 2019
Published: 29 April 2019

Research article

Parameterizing a dynamic influenza model using longitudinal versus age-stratified case notifications yields different predictions of vaccine impacts

Michael A. Andrews[1] **and Chris T. Bauch**[2,*]

[1] Department of Mathematics and Statistics, University of Guelph, 50 Stone Rd. E. Guelph, Ontario, Canada N1G 2W1

[2] Department of Applied Mathematics, University of Waterloo, 200 University Ave. W. Waterloo, Ontario, Canada N2L 3G1

* **Correspondence:** Email: cbauch@uwaterloo.ca.

Abstract: Dynamic transmission models of influenza are sometimes used in decision-making to identify which vaccination strategies might best reduce influenza-associated health burdens. Our goal was to use laboratory confirmed influenza cases to fit model parameters in an age-structured, two-type (influenza A/B) dynamic model of influenza. We compared the fitted model under two fitting methodologies: using longitudinal weekly case notification data versus using cross-sectional age-stratified cumulative case notification data. The longitudinal data came from a Canadian province (Ontario) whereas the cross-sectional data came from the national level (all of Canada). We find that the longitudinal fitting method provides best fitting parameter sets that have a higher variance between the respective parameters in each set than the cross-sectional cumulative case method. Model predictions–particularly for influenza A–are very different for the two fitting methodologies under hypothetical vaccination scenarios that expand coverage in either younger age classes or older age classes: the cross-sectional method predicts much larger decreases in total cases under expanded vaccine coverage than the longitudinal method. Also, the longitudinal method predicts that vaccinating younger age groups yields greater declines in total cases than vaccinating older age groups, whereas the cross-sectional method predicts the opposite. We conclude that model predictions of vaccination impacts under different strategies may differ at national versus provincial levels. Finally, we discuss whether using longitudinal versus cross-sectional data in model fitting may generate further differences in model predictions (above and beyond population-specific differences) and how such a hypothesis could be tested in future studies.

Keywords: epidemic modelling; influenza; vaccination; parameter estimation

1. Introduction

Seasonal influenza imposes a significant health burden each year, reducing the quality of life for many across the globe [1]. Although often viewed as a mild illness typically causing school or workplace absenteeism, influenza can cause significant complications for vulnerable populations such as the elderly or those with weakened immune systems. In order to combat influenza, health jurisdictions may implement vaccination programmes (such as the Universal Influenza Immunization Program in Ontario, Canada) that may target certain age groups, professions, or make vaccines widely available to the public.

Nonlinear dynamical systems is an enduring topic of research in population dynamics, including both ecology and epidemiology [2–6]. The seasonality and nonlinearity that characterize influenza and many other vaccine-preventable infectious diseases have made complex disease dynamics a frequent topic of research among mathematical biologists [2, 3, 7, 8]. Mathematical models of disease transmission have been used to assess vaccination policies. For instance, dynamic transmission models have been used to evaluate the effectiveness of control strategies for seasonal influenza, such as targeted vaccinations and vaccine types [9–20]. These models are increasingly essential for decision-making regarding vaccine implementation, since *in silico* experiments regarding optimal age of vaccination can be done when experiments in real populations are impossible or impractical. For example, a frequent problem addressed in the literature is finding an optimal approach to distributing vaccines [21]. Some research has found that targeting younger age groups produces the most benefit in limiting influenza spread and improving health outcomes across a population [22–27]. However, other research has also shown this may not be the case in all circumstances [9, 28]. In the past, influenza transmission models have either chosen parameters without a fitting process [29, 30], have been fitted to a single year using cross-sectional cumulative cases for that season [31, 32], to weekly reports of influenza like illness (ILI) data, under the assumption that ILI incidence follows the same patterns as seasonal influenza [17, 31, 33, 34], or to laboratory confirmed influenza cases [17, 32]. Available data usually specify the timing of infection cases but not their age, or the age of infection cases but not their timing. To our knowledge, no modelling studies have compared results when fitting to longitudinal time series data versus cross-sectional cumulative case data over multi-year time spans.

Here, we create an age stratified dynamic transmission model of seasonal influenza following an approach similar to Thommes et al. [18], and use positive influenza specimen tests for parameter estimation. Our research questions are (1) to determine whether a dynamic two-strain (influenza A/influenza B) age-structured transmission model can be fitted to longitudinal time series data of weekly laboratory confirmed influenza cases spanning multiple years, and (2) to determine whether different vaccine programs are predicted to have different impacts at the provincial level (longitudinal data from Ontario) versus the national level (cross-sectional data from Canada). We compare the model fit and model predictions of the impact of vaccination when fitting to the longitudinal lab confirmed influenza cases versus when the model is fitted only to non-longitudinal, cross-sectional data on age-stratified cumulative attack rates instead.

2. Methods

Our model is a compartmental age structured model [35], and we will fit important transmission parameters to longitudinal influenza case data, as well as cross-sectional age stratified case data. Incorporating age structure is a critical factor, as population contact patterns, and therefore influenza transmission, depend on age [36]. Details of the model development are given in the following sections. All data used in our study are publicly available.

2.1. Population demographics

Our model uses age compartments one year in size, starting from 0–1 years, and ending at 99+ years [18, 33]. The population size and age structure are modelled after the province of Ontario, Canada, to remain consistent with our data on influenza incidence. We also chose Ontario as our study population on account of its relatively large population size and presence of a universal influenza immunization program in the province. When we age the population, we use yearly population projections given by Ontario's Ministry of Finance, which are based on census data, birth/death rates, immigration, and emigration [37–39], or census data (for the model's 2011 population) [40]. Due to the model's high age resolution, we are able to specify age dependent contact rates. These contact rates play a crucial role in influenza transmission, and we use a contact matrix which specifies the mean daily duration of contact time in minutes between age groups [41]. These contact data are based on studies conducted in the United States, and thus we are making the assumption that contact rates in the region we are modelling are similar.

2.2. Influenza incidence data and epidemiology

Data on confirmed influenza cases are available for the province of Ontario, Canada from the years 2010 to 2015 [42]. The data give the weekly number of confirmed cases in the province for the specified years. For fitting our model to age stratified cumulative cases, we use the years 2011 to 2016 due to these years having the required data available. The age categories used in the fitting are 0–19, 19–65, and 65+. In our model, we will consider influenza cases caused by both the A and B strains.

The influenza virus in our model has a susceptible-infected-recovered-vaccinated natural history. For transmission, we use the contact hypothesis [41] where our contact matrix C taken from Table 1 in Zagheni et al. [41] defines the average daily time of contact between age groups. We define β_i to be the probability that an individual in age group i becomes infected after being in contact with an infectious individual, which in our case is constant across age groups. The time varying force of infection for age group i is given by

$$\lambda_i(t) = \sum_{j=1}^{100} \beta_i \, C_{ij} \left(\frac{I_j}{N_j} \right), \tag{2.1}$$

where I_j is the number of infected individuals in age group j and N_j is the size of age group j. Additionally, influenza incidence shows a prominent annual recurrence in the winter months, which has been thought to be caused by a variety of factors such as temperature, humidity, and changes in contact patterns [43–45]. To ensure this seasonal variation in our model, we use a sinusoidal

function [53] and multiply the force of infection by

$$1 + A\ cos\left(\frac{2\pi(t+\delta)}{365}\right), \tag{2.2}$$

where A is the amplitude of the seasonality function which determines the variation of the basic reproductive number $\mathcal{R}_0$, and δ determines on what day the maximum value of the seasonality occurs ($\delta = 0$ corresponds to January 1). This formulation is similar to previous work modelling the same dynamic [18, 29, 33], and we use the derivation found in Thommes et al. [18] to relate β_i to $\mathcal{R}_0$.

Finally, infected individuals recover at a constant rate γ. Also, to model the antigenic drift of the influenza virus [59], we force individuals that have been infected to lose their immunity at a constant rate. In our model, natural immunity loss occurs at rate ρ_N.

2.3. Vaccination

In Ontario, the primary types of vaccines used are the trivalent inactivated vaccine, the quadrivalent inactivated vaccine, and the quadrivalent live-attenuated vaccine [46]. In this region, the recommended individuals to receive vaccination are those aged 6 months and older, and especially individuals in high risk groups or those who may directly transmit to high risk groups [46].

In our model, we specify a proportion of individuals in each age group to become vaccinated each year. At the time of vaccination, a vaccinated individual in age group i receives vaccine induced immunity according to the vaccine's efficacy with probability ϵ_i, and remains susceptible with probability $1 - \epsilon_i$. Vaccine efficacy is set to 65% for ages $<$ 65 and 55% for older age groups [50–52]. We also assume there is no partial immunity conferred with an inefficacious vaccination [18, 33]. For vaccination coverage rates, we use data from the studies by [55, 62–64], and based on the age ranges given, we use linear interpolation to restore our yearly age resolution. The baseline coverages are 0-1 years: 3.7%, 1–2 years: 7.4%, 2–11 years: 29.48%, 12–19 years: 36%, 20–49 years: 25.5%, 50–64 years: 48%, 65–74 years: 73%, 75–84 years: 84%, and 85+ years: 82% [55, 62–64]. Much like natural immunity, vaccine acquired immunity wanes at a constant rate of ρ_V. In our model, we choose ρ_V to be a fitted parameter rather than choosing it as a fixed value or assuming it to be equal to ρ_N, as was used in previous studies [18, 29, 33]. Finally, those who become infected regardless of vaccinating will not show a reduction in infectiousness.

2.4. Model structure

Our system of differential equations consists of susceptible $S_i(t)$, infected $I_i(t)$, recovered $R_i(t)$, and vaccinated $V_i(t)$ individuals where i denotes the respective age class to which an individual belongs:

$$\frac{dS_i}{dt} = -\lambda_i(t) + \rho_N R_i + \rho_V V_i \tag{2.3}$$

$$\frac{dI_i}{dt} = \lambda_i(t) - \gamma I_i \tag{2.4}$$

$$\frac{dR_i}{dt} = \gamma I_i - \rho_N R_i \tag{2.5}$$

$$\frac{dV_i}{dt} = -\rho_V V_i \tag{2.6}$$

A diagram showing the compartments and primary transitions of our model is shown in Figure 1. Each year we choose a day near the end of summer (August 31), to age the population [18,29,33]. Individuals are moved to the next age class in one time step, and those in the 99+ age category remain. Then, the population is scaled to match the demographics of the next year's population, as projected by Statistics Canada and Ontario Ministry of Finance. If these more in depth metrics are not available, population birth and death rates may simply be used. Newborns entering the first age category all populate the S_0 compartment. We age the population in this discrete manner to avoid artifacts introduced by a continuous aging process [65]. We reduce the impact of the discontinuities introduced through discrete aging by applying the process in the summer when infected cases are low, and also by re-initializing the adaptive stepsize numerical solver after each aging event. Other approaches which could also capture a discrete aging process include use of delay differential equations.

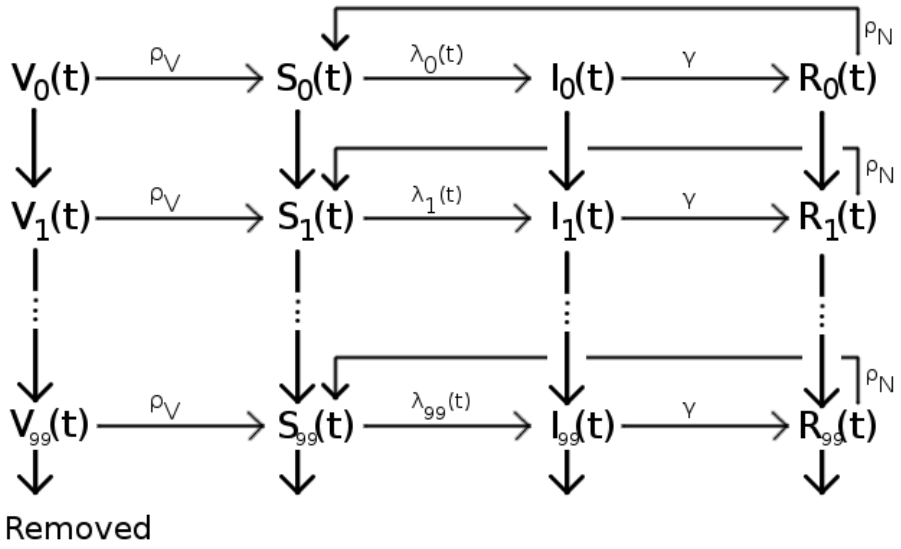

Figure 1. Diagram of the age-stratified SIRS compartmental model with vaccination. See Methods for definitions of parameters and variables.

Next, vaccination occurs on October 1 of each year because in our selected region the majority of vaccination occurs in the fall. In our model, we make the approximation that vaccination of the population occurs before each influenza epidemic begins. Then, at a point t_{seed} we add an external value λ_{ext} to the force of infection for the remainder of the influenza season. This is a hybrid between models by Goeyvaerts et al. [33] and Thommes et al. [18] as we find this small addition to the force of infection grants a smoother transition into each new influenza season as opposed to infecting a bulk amount of individuals all at once on t_{seed}. In addition, for the period of time between seasons, we remove λ_{ext} so no additional new cases arise. Vertical arrows represent aging, and on the day of the year the population is aged, members in each compartment are added to the corresponding compartment in the next age group.

The system is integrated with a time step of one day allowing for precise calculation the the daily force of infection as well as sufficient numerical solution accuracy. We use the MATLAB package ODE4 to fulfill our fixed time step requirement. In addition to the 5 year time period for which we

have empirical influenza incidence data, we run our model with a 10 year burn in period. During the burn in period, we use the 2010 population demographics and maintain the same vaccine uptake rates that were used during our period of interest.

2.5. Parameter fitting

We compared two methods of fitting our model's parameters: fitting the parameters to longitudinal weekly case notification data spanning multiple years (we will call this the "longitudinal method") and fitting the parameters to cross-sectional age-stratified data that lack a temporal variable (we call this the "cross-sectional method").

2.5.1. Longitudinal method

We aim to fit the parameters of our model to multi-year longitudinal time series data taken from [42] in a similar manner to Goeyvaerts et al. [33]. However, we use laboratory confirmed influenza specimen cases instead of ILI incidence data used by previous models which are based on reported influenza-like symptoms rather than laboratory confirmed cases.

In order to quantify the goodness of fit for a given parameter set, we use a least squares approach: empirical weekly incidence of the number of positive influenza cases is compared to our model's corresponding output. We define the empirical number of reported cases in week w to be I_w^H, and the number of cases given by the model in week w to be I_w^M. In order to directly compare the two quantities, we use the parameter α introduced by Goeyvaerts et al. [33] to scale the model incidence. Here, α captures the probability that an infected individual is symptomatic, visits a medical practitioner and gets tested for the influenza virus which returns a positive result. The sum of squares error is then

$$\sum_{\forall w} \left(I_w^H - (\alpha I_w^M) \right)^2 . \tag{2.7}$$

To evenly sample the parameter space, we use Latin hypercube sampling [54] to generate 35,000 parameter combinations. Parameter descriptions and fitting ranges (that is, Latin hypercube sampling ranges) are given in Table 1. We then determine each parameter set's sum of squares score over a simulation run. Next, we utilize MATLAB's GlobalSearch algorithm to search for optimal parameter combinations using the parameter sets that offered the lowest sum of squares values. GlobalSearch attempts to find a function's global minimum, and initializes its search over the parameter space from a user defined start point. In our case, the function we are seeking to minimize is the sum of squares score of our system of differential equations. The input points are used by the solver to determine an initial estimate for a basin of attraction, and the algorithm also generates a set of trial points to be used in finding the minimum. Additionally, upper and lower bounds may be specified for each parameter, which we define as the same bounds used in the Latin hypercube sampling. Any number of runs of the GlobalSearch algorithm may be performed, using a different starting point corresponding to the parameter sets obtained from the Latin hypercube samples for each run. Moreover, maximum runtimes may be specified as well.

Table 1. Parameter descriptions, fitting ranges and literature sources.

Parameter	Description	Fitting Range	Source
A	Amplitude of seasonality	0–1.0	Maximum range
t_{seed}	Days after vaccination when infected are seeded into the population	1–90 1–120**	Assumption*
δ	Timing of seasonality function peak (days)	-10–45 -60–10**	Assumption*; forces transmission peak to fall between November and January
$\mathcal{R}_0$	Basic reproduction number	1.0–2.5	[47], peak range also encompasses estimates of pandemic influenza strains [49, 58]
γ	Mean latent plus infectious period	4 days (fixed)	[48]
ρ_N	Natural immunity waning rate	1.0–2.5 years 1.0–4.0 years**	Assumption*
ρ_V	Vaccine immunity waning rate	0.5–1.5 years	[18, 33, 61]
λ_{ext}	Case importation rate	0.01–0.2	Assumption
α	Scaling factor of model incidence	0.0005–0.15	Estimated based on [60].

*Also based on preliminary Latin hypercube sampling. Wider ranges were originally used, but the best results were contained within ranges shown above.
**Ranges used for influenza B.

Due to the stochastic nature of the process, more runs may result in lower least squares fits, and the available computational resources will be a determinant of how many initial points, and therefore runs, of GlobalSearch are used. In our analysis, we use the 50 best performing parameter sets obtained from the Latin hypercube sampling to use as initial points for the GlobalSearch algorithm. We also tested a group of random initial points gathered from the top 15% of parameter sets from the Latin hypercube sampling, but they did not provide better results (lower sum of squares) than the aforementioned top 50 sets.

2.5.2. Cross-sectional method

For fitting age-stratified cumulative cases over the 5-year period, the available data is from Canada as a whole [42] (Ontario level data does not include age-stratified cases). Thus, we scaled the cases by the proportion of the Canadian population that lives in Ontario in order to remain consistent with the longitudinal method's fitting.

The fitting for the cross-sectional method was identical to the longitudinal method, except we did not compute a difference of squares from the model output to the empirical data for each week. Instead, the difference of squares was computed over total cases by age over the entire 5-year period.

3. Results

3.1. Parameter fitting comparison

Time series of the best parameter combinations resulting from the the fitting processes for influenza A and B for the longitudinal method (Ontario data) are shown in Figure 2 and Table 2. The plotted results are compared to the empirical laboratory confirmed cases over the time period. We used a separate fitting process for each strain, although we assume that the vaccine efficacy and the infectious periods are the same for both. The largest differences in our model emanate from the 2012 season for influenza B. Most parameter sets undershoot the peak in this season, although some achieve much closer fits.

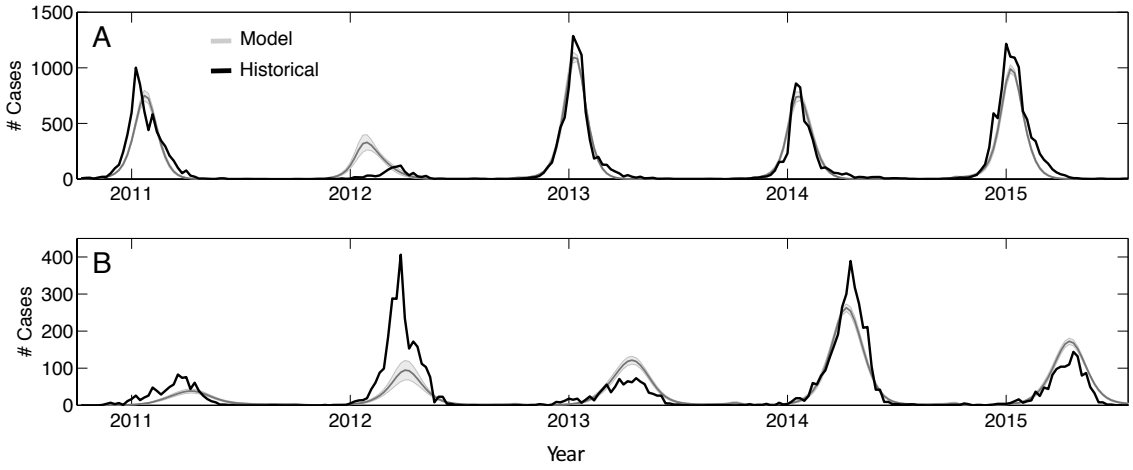

Figure 2. Time series of confirmed influenza cases (black) and the fitted model (grey) for (A) Influenza A and (B) Influenza B. Shaded region represents 95% confidence intervals.

Table 2. Best fitting parameter values (mean and standard deviation) for the longitudinal method (Ontario data).

Parameter	Mean Value, A Strain	Std. Dev.	Mean Value, B Strain	Std. Dev.
A	0.6304	0.2211	0.3115	0.1269
t_{seed} (day)	35.43	23.89	71.50	32.06
δ (day)	21.15	10.85	-25.14	19.26
$\mathcal{R}_0$	1.424	0.3195	1.322	0.3302
ρ_N (days)	593	237	1408	133
ρ_V (days)	468	88	427	169
λ_{ext}	0.0965	0.0731	0.1108	0.0731
α	0.002608	0.0008535	0.004065	0.006953

Simulations using the cross-sectional method (Canada data) produce the parameter combinations shown in Table 3, with the results for each age category shown in Figure 3A,B. Age-stratified results from the longitudinal model are included for comparison in panels C and D. In the cross-sectional method results, the variance of total cases produced by the parameter sets for influenza A (Figure3A) in each age category are much lower than that of influenza B (Figure 3B). This could stem from the fitting process, and how Globalsearch attempts to find optimal parameter combinations to match the age-stratified data. In the case of influenza A, each search has parameters converge to very similar values, whereas the final values for influenza B have much higher variance in comparison. This could be due to how the infections are spread out across each age category. For example, influenza A has many more cases in the ages 19+ than the ages 0-18. However, influenza B's cases are evenly spread across all ages.

For influenza A, the primary differences in parameters for the two fitting methods stem from the parameters A (seasonality amplitude), $\mathcal{R}_0$ (average basic reproduction number), and α (incidence

Table 3. Best fitting parameter values (mean and standard deviation) for the cross-sectional method (Canada data).

Parameter	Mean Value, A Strain	Std. Dev.	Mean Value, B Strain	Std. Dev.
A	0.4405	0.06101	0.4904	0.1131
t_{seed} (day)	46.16	26.14	59.84	36.14
δ (day)	12.02	14.64	-25.70	20.69
$\mathcal{R}_0$	1.003	0.01747	1.044	0.09240
ρ_N (days)	416	68	464	238
ρ_V (days)	548	11	550	0
λ_{ext}	0.06387	0.03514	0.1411	0.06227
α	0.0935	0.02440	0.004088	0.0008010

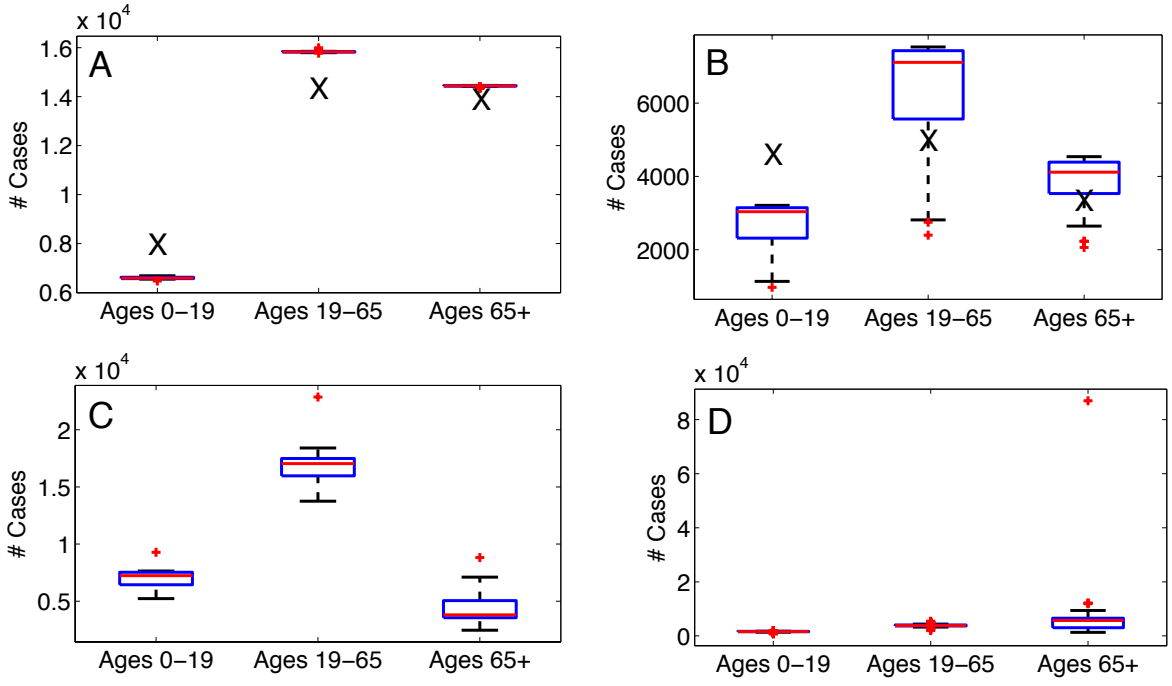

Figure 3. Age-stratified cumulative cases for influenza A and B compared to empirical targets in our model. Target number of cases from the emprical data are given by X's where applicable. (A) Cross-sectional method for influenza A. (B) Cross-sectional method for influenza B. (C) Longitudinal method with age-stratified results for influenza A. (D) Longitudinal method with age-stratified results for influenza B. From bottom to top, each line in each boxplot shows the following information: minimum value, first quartile, median, third quartile, maximum value. Red crosses are considered outliers.

scaling factor). For the seasonality amplitude, we notice that when not required to meet multiple varying seasonal peaks as we did in the longitudinal method, the average amplitude is lower with less

variance in the cross-sectional method fits than in its longitudinal counterparts. Similarly, the average $\mathcal{R}_0$ value amongst the parameter sets follows the same pattern: in the cross-sectional method's fits, the average value and variance amongst the sets is lower than the values seen in the longitudinal method's fits. Finally, α is much larger in the cross-sectional method's fits.

For influenza B, the biggest differences in parameters for the two fitting methods stem from $\mathcal{R}_0$ and the waning immunity rates ρ_N and ρ_V. Similarly to influenza A, the average basic reproduction number is smaller and has less variance amongst the sets in the cross-sectional method compared to the longitudinal method's parameters. Also, the average natural waning immunity rate is smaller as well. An interesting note is that the vaccine conferred waning immunity rate takes on its maximum allowed value in all of the sets for the cross-sectional method. A similar, but less extreme, shift towards the maximum ρ_V value occurs in the influenza A parameter sets as well.

3.2. Projected impact of expanded vaccination coverage

These results may be further tested by observing the impact of implementing different vaccination scenarios or strategies for vaccine allocation approaches. Here, we test the changes in outcomes of our model with a targeted vaccine allocation in a scenario where a health jurisdiction is expanding their influenza vaccination program.

When expanding vaccination coverage, an important consideration is targeted distribution of vaccines. The two main strategies are to target children, who are believed to be responsible for the majority of transmission, or to target high risk individuals and age groups, such as the elderly. To test these two scenarios, we will increase the vaccine uptake of younger aged age groups in our model (ages 0–18) by 30% for each age, a strategy which has been believed to indirectly protect other age groups as well [25–27]. Then, we compare these results to increasing the total number of vaccines administered by the same amount in older age groups instead, which in our population is the ages 55+.

With the longitudinal fitting method for influenza A, vaccinating younger age groups produces a 24.53% drop in total cases on average from baseline vaccination (Figures 4A and 5C). When targeting older age groups, we see an average reduction in total cases of 13.86%. Total mean confirmed cases and their 95% CIs are found in Table 4. Thus, the vaccination program aimed at the younger age classes provides a small benefit in total case reduction on average compared to a similar program targeting older ages. This stems from the low baseline vaccine uptake in children and their high contact rates with each other as well as middle aged adults. In the case of influenza B, targeting the younger ages gives an average 19.52% drop in total cases, whereas targeting the older age groups gives an average 24.27% drop in the mean (Figures 4B and 5D). Total mean confirmed cases and their 95% CIs are found in Table 4. In this case, vaccinating older age groups produces a small but largely negligible average reduction in the mean of total cases across parameter sets used.

Using the cross-sectional method, influenza A results differ from the longitudinal method. In this case, vaccinating older age groups results in the best case reduction (Figure 5A with comparison to the longitudinal model in Figure 5C). When increasing vaccination rates in the ages 55+, we see less than half the total cases than when expanding vaccination amongst ages 0-18. Total mean confirmed cases from the simulations are found in Table 4. For influenza B, vaccinating the younger age groups yields a 31.89% reduction in mean cases compared to baseline, and vaccinating older age groups provides a 50.16% reduction (Figure 5B with comparison to the longitudinal model in Figure 5D). Total mean

Table 4. Mean number of cases for influenza strains A and B under different vaccination scenarios.

Strain	Baseline (95% CI)	Vac. Prog. 0–18 (95% CI)	Vac. Prog. 55+ (95% CI)
Longitudinal Method (Ontario)			
Influenza A	28,787 (± 2,652)	21,725 (±2, 392)	24,797 (±3, 134)
Influenza B	7,483 (±914)	6,022 (±935)	5,667 (±1, 072)
Cross-Sectional Method (Canada)			
Influenza A	35,833 (±390)	24,173 (±597)	10,947 (±607)
Influenza B	12,861 (±785)	8,760 (±505)	6,410 (±646)

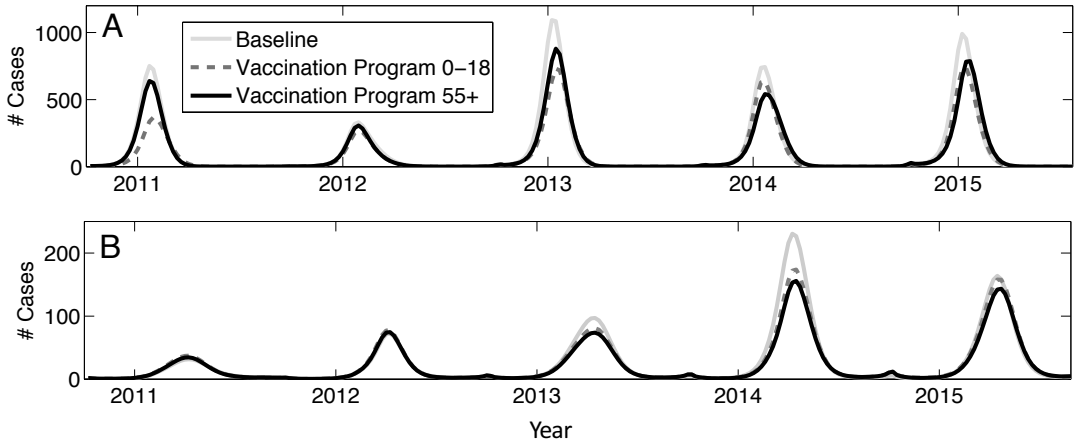

Figure 4. Time series of confirmed influenza A and B cases in the model under different vaccination scenarios for (A) Influenza A and (B) Influenza B.

confirmed cases from the simulations are found in Table 4. In general, the cross-sectional method's fitting predicts a much larger decrease in total cases for any vaccination expansion strategy than the longitudinal time series fitting method. These results reveal that the varying types of data that can be used to fit a predictive model of influenza transmission can produce very different results.

4. Discussion

We have designed and implemented an age-stratified dynamic transmission model of seasonal influenza for both A and B influenza types. The model parameters were fit to laboratory confirmed influenza cases from the years 2010–2015 in the province of Ontario, Canada, as well as age-stratified cumulative case data from the years 2011–2016 in Canada as a whole. We used this model to evaluate vaccine expansion strategies which target certain age groups, finding significant differences in predicted vaccine program impacts at the provincial and national level.

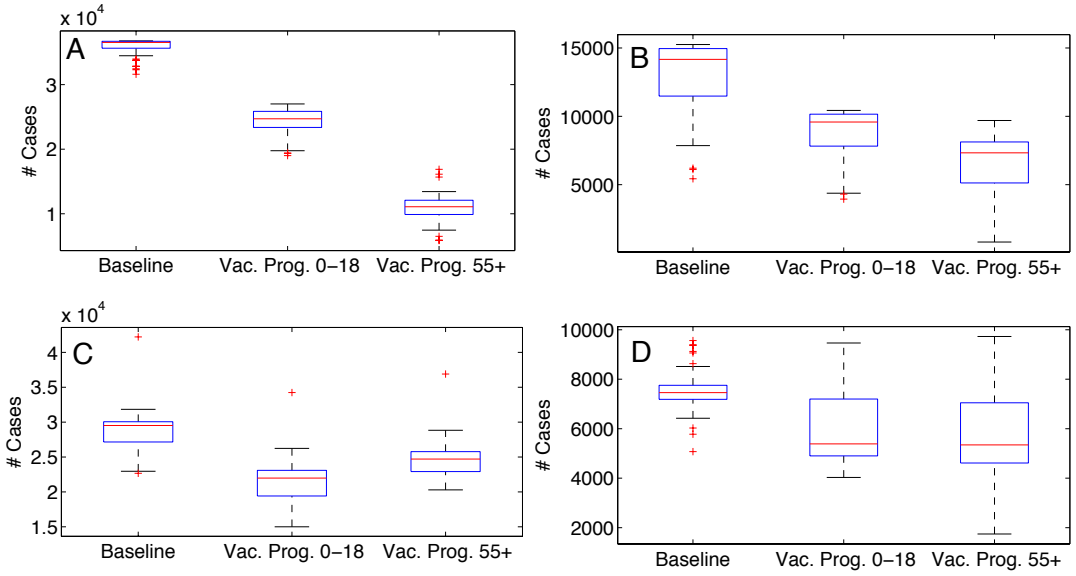

Figure 5. Cross-sectional method results of model predictions for influenza A and B cases under different vaccination scenarios, including comparison to the longitudinal method's corresponding results. Subpanels show number of cases for (A) Influenza A, (B) Influenza B, (C) Longitudinal method with age-stratified results for influenza A, (D) Longitudinal method with age-stratified results for influenza B. From bottom to top, each line in each boxplot shows the following information: minimum value, first quartile, median, third quartile, maximum value. Red crosses are considered outliers.

Using the cross-sectional method (for Canada as a whole), the variance amongst the respective parameters in each of the 50 best sets is generally smaller than that of the variance amongst the parameters found using the longitudinal method (for Ontario only). Also, when introducing vaccination scenarios targeting different age groups, outcomes from using parameters derived from the two types of data differ, particularly for influenza A. For example, the cross-sectional method's data predicts much larger decreases in total cases from baseline vaccination coverage than the time series data. Additionally, those simulations show that vaccinating older age groups will provide the most benefit in reducing the total number of cases in the population. Using the longitudinal method, results show that vaccinating younger age groups provides a moderate total case reduction for influenza A, and vaccinating older age groups provides a slight total case reduction for influenza B.

Our model makes some simplifying assumptions. For example, we assume that vaccine efficacy and the infectious period are the same for both A and B strains, which may not hold in general [18]. Similarly, the parameter α, which represents the rate at which an infected individual is symptomatic and visits a physician who in turn administers a laboratory test for influenza which returns positive, is constant across all age groups. In reality, this may not be the case as some age groups may be more likely to visit a physician after becoming ill, or physicians may be more likely to administer tests for certain age groups. We also assume that the laboratory confirmed case data is a consistently uniform sample of all influenza cases. However, physicians may send in more tests depending on the time of year or when they perceive the prevalence of influenza is higher.

A related simplifying assumption that is implicit in the use of a least-squares approach for fitting is that case reporting has a normal distribution with mean αI_w^M. This means that the variance in reporting is assumed to be the same for all values of αI_w^M, *i.e.* it is the same throughout the year. However, this may bias in parameter estimates for an infection like influenza where incidence can change dramatically throughout the year from very low (such as in the summer months in the northern hemisphere) to very high (such as during the peak of an outbreak). Future research could explore parameter estimates under an approach to modelling case reporting variance where each case has a probability α of being reported (the assumption used elsewhere in the paper): this leads to a binomial model which has variance $\alpha(1-\alpha)I_w^M$ which scales linearly with I_w^M, thereby allowing for more variance at the peaks.

An additional simplifying assumption we made is that seasonal forcing is sinusoidal and has the same magnitude for all age classes. In reality, the forcing function may be non-sinusoidal, or its magnitude may differ across age classes. Although schools appear to be fairly important for influenza transmission [57], it is also possible (although not firmly established) that susceptibility to influenza infection varies seasonally across a broad swathe of ages on account of vitamin D deficiency in winter months [56, 66]. The combined effects of different sources of seasonality on influenza transmission–and how they vary across populations and time–is an outstanding question with implications for how we model influenza transmission. Other models have taken alternative approaches of using the school term model in the transmission function [67] or using more complicated seasonal forcing functions such as B splines [68].

There are some differences in the data that hindered our comparisons. Perhaps most crucially, the age-stratified cumulative case data gives country-wide cases, whereas the time series data is for the province of Ontario only. Thus it is difficult to disentangle whether the observed effects are due to the different methodologies for fitting the model, or the fact that the data come from different populations. Although we scaled the countrywide number of cases by the proportion of the Canadian population that lives in Ontario, the epidemiology may still not be directly comparable due to environmental and other demographic differences. Moreover, the age-stratified cumulative case data available covers a one year difference from the time series data, causing some discrepancy in the number of cases over each 5 year span. Future research could explore this issue more systematically.

In the same vein, the implications of fitting an age-structured model to data that are not age-structured could be explored in order to better understand the implications for parameter identifiability. This problem arises on account of the unfortunate fact that available data often only describe the timing of new cases or the age of infected hosts, but not both. The lack of data for both timing and age has implications for our ability to infer sets of parameter values that have small uncertainty intervals. To better understand the implications for our model, parameter estimates under the approach described in this paper could be compared to two alternatives approaches. Under one approach, the existing age-structured model could be fit to synthetic age-structured data where both timing and age of infections are available. Under a second approach, model parameters such as the amplitude of seasonal forcing and the case reporting rate could be allowed to vary with age and the model could then be fitted to the synthetic age-structured data. Inferred parameter values from the two approaches and their uncertainty intervals could then be compared to the inferred parameter values under the existing approach.

Vaccine programs are known to have different impacts in different populations. Here we illustrated how the possibility of differing impacts at provincial and federal levels is tied up with the question of what kind of data are available and how the data are used to parameterize transmission models. Future work can more systematically analyze the impact of using longitudinal versus cross-sectional data for influenza model parameterization, to enable modellers to understand the effects of data availability and parameter identifiability on influenza model predictions.

Acknowledgments

This research was funded by a Natural Sciences and Engineering Research Council (NSERC) Discovery Grant to CTB.

Conflict of interest

The authors declare that there are no conflicts of interest.

References

1. L. Simonsen, The global impact of influenza on morbidity and mortality, *Vaccine*, **17** (1999), S3–S10.
2. C. Bauch and D. Rand, A moment closure model for sexually transmitted disease transmission through a concurrent partnership network, *Proc. R. Soc. Lond. B*, **267** (2000), 2019–2027.
3. D. Rand and H. Wilson, Chaotic stochasticity: a ubiquitous source of unpredictability in epidemics, *Proc. R. Soc. Lond. B*, **246** (1991), 179–184.
4. C. Innes, M. Anand and C.T. Bauch, The impact of human-environment interactions on the stability of forest-grassland mosaic ecosystems, *Sci. Rep.*, **3** (2013), 2689.
5. van B. Minus and R. A. David, The unit of selection in viscous populations and the evolution of altruism, *J. Theor. Biol.*, **193** (1998), 631–648.
6. D. A. Rand and H. B. Wilson, Using spatio-temporal chaos and intermediate-scale determinism to quantify spatially extended ecosystems, *Proc. R. Soc. Lond. B*, **259** (1995), 111–117.
7. D. J. Earn, P. Rohani, B. M. Bolker, et al., A simple model for complex dynamical transitions in epidemics, *Science*, **287** (2000), 667–670.
8. M. Keeling, D. Rand and A. Morris, Correlation models for childhood epidemics, *Proc. R. Soc. Lond. B*, **264** (1997), 1149–1156.
9. J. Dushoff, J. Plotkin, C. Viboud, et al., Vaccinating to protect a vulnerable subpopulation, *PLOS Med.*, **4** (2007), e174.
10. M. Alexander, C. Bowman, S. Moghadas, et al., A vaccination model for transmission dynamics of influenza, *SIAM J. Appl. Dyn. Sys.*, **3** (2004), 503–524.
11. J. Glasser, D. Taneri, Z. Feng, et al., Evaluation of targeted influenza vaccination strategies via population modeling, *PLOS One*, **5** (2010), e12777.

12. J. M. Tchuenche, N. Dube, C. P. Bhunu, et al., The impact of media coverage on the transmission dynamics of human influenza, *BMC Public Health*, **11** (2011), S5.

13. S. P. Tully, A. M. Anonychuk, D. M. Sanchez, et al., Time for change? an economic evaluation of integrated cervical screening and hpv immunization programs in Canada, *Vaccine*, **30** (2012), 425–435.

14. C. R. Wells and C. T. Bauch, The impact of personal experiences with infection and vaccination on behaviour–incidence dynamics of seasonal influenza, *Epidemics*, **4** (2012), 139–151.

15. C. Wells, E. Klein and C. Bauch, Policy resistance undermines superspreader vaccination strategies for influenza, *PLOS Comput. Biol.*, **9** (2013), e1002945.

16. Y. Hsieh, Age groups and spread of influenza: Implications for vaccination strategy, *BMC Infect. Dis.*, **10** (2010), 106.

17. M. Baguelin, S. Flasche, A. Camacho, et al., Assessing optimal target populations for influenza vaccination programmes: An evidence synthesis and modelling study, *PLOS Med.*, **10** (2013), e1001527.

18. E. Thommes, A. Chit, G. Meier, et al., Examining ontario's universal influenza immunization program with a multi-strain dynamic model, *Vaccine*, **32** (2014), 5098–5117.

19. M. A. Andrews and C. T. Bauch, Disease interventions can interfere with one another through disease-behaviour interactions, *PLoS Comput. Biol.*, **11** (2015), e1004291.

20. M. A. Andrews and C. T. Bauch, The impacts of simultaneous disease intervention decisions on epidemic outcomes, *J. Theor. Biol.*, **395** (2016), 1–10.

21. J. Medlock and A. Galvani, Optimizing influenza vaccine distribution, *Science*, **325** (2009), 1705–1708.

22. T. A. Reichert, N. Sugaya, D. S. Fedson, et al., The japanese experience with vaccinating schoolchildren against influenza, *N. Engl. J. Med.*, **344** (2001), 889–896.

23. A. Monto, F. M. Davenport and J. A. N. an T. Francis, Effect of vaccination of a school-age population upon the course of an A2/Hong Kong influenza epidemic, *Bull. W.H.O.*, **41** (1969), 537–542.

24. P. A. Piedra, M. J. Gaglani, C. A. Kozinetz, et al., Herd immunity in adults against influenza-related illnesses with use of the trivalent-live attenuated vaccine (CAIV-T) in children, *Vaccine*, **23** (2005), 1540–1548.

25. D. Weyecker, J. Edelsberg, M. E. Halloran, et al., Population-wide benefits of routine vaccination of children against influenza, *Vaccine*, **23** (2005), 1284–1293.

26. M. E. Halloran, I. M. Longini, D. M. Cowart, et al., Community interventions and the epidemic prevention potential, *Vaccine*, **20** (2002), 3254–3262.

27. I. M. Longini and M. E. Halloran, Strategy for distribution of influenza vaccine to high-risk groups and children, *Am. J. Epidem.*, **161** (2005), 303–306.

28. P. Beutels, Y. Vandendijck, L. Willem, et al., *Seasonal influenza vaccination: prioritizing children or other target groups? Part II: cost-effectiveness analysis*, Technical report, Belgian Health Care Knowledge Centre (KCE), 2013.

29. E. Vynnycky and W. Edmunds, Analyses of the 1957 (asian) influenza pandemic in the united kingdom and the impact of school closures, *Epidem. Infect.*, **136** (2008), 166–179.
30. R. J. Pitman, Estimating the clinical impact of introducing paediatric influenza vaccination in england and wales, *Vaccine*, **30** (2012), 1208–1224.
31. P. Poletti, M. Ajelli and S. Merler, The effect of risk perception on the 2009 H1N1 pandemic influenza dynamics, *PLOS One*, **6** (2011), e16460.
32. J. T. Wu, K. Leung, R. Perera, et al., Inferring influenza infection attack rate from seroprevalence data, *PLOS Pathogens*, **10** (2014), e1004054.
33. N. Goeyvaerts, L. Willem, K. V. Kerckhove, et al., Estimating dynamic transmission model parameters for seasonal influenza by fitting to age and season-specific influenza-like illness incidence, *Epidemics*, **13** (2015), 1–9.
34. J. B. Axelson, R. Yaari, B. T. Grenfell, et al., Multiannual forecasting of seasonal influenza dynamics reveals climatic and evolutionary drivers, *Proc. Natl. Acad. Sci. USA*, **111** (2014), 9538–9542.
35. R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, 1992.
36. J. Mossong, N. Hens, M. Jit, et al., Social contacts and mixing patterns relevant to the spread of infectious diseases, *PLOS Med.*, **5** (2008), e74.
37. Ontario Ministry of Finance, Ontario Population Projections Update 2010-2036, 2011. Available from: `http://www.fin.gov.on.ca/en/economy/demographics/projections/projections2010-2036.pdf`.
38. Ontario Ministry of Finance, Ontario Population Projections Update 2012-2036, 2013. Available from: `http://www.fin.gov.on.ca/en/economy/demographics/projections/projections2012-2036.pdf`.
39. Ontario Ministry of Finance, Ontario Population by Age 2013-2041, 2015. Available from: `http://www.fin.gov.on.ca/en/economy/demographics/projections/table6.html`.
40. Statistics Canada, Focus on Geography Series, 2011 Census, 2012. Available from: `http://www12.statcan.gc.ca/census-recensement/2011/as-sa/fogs-spg/Facts-pr-eng.cfm?Lang=Eng&GK=PR&GC=35`.
41. E. Zagheni, F. Billari, P. Manfredi, et al., Using time-use data to parameterize models for the spread of close-contact infectious diseases, *Am. J. Epidem.*, **168** (2008), 1082–1090.
42. Government of Canada Publications, FluWatch, 2015. Available from: `http://publications.gc.ca/site/eng/9.507424/publication.html`.
43. J. Dushoff, J. B. Plotkin, S. A. Levin, et al., Dynamical resonance can account for seasonality of influenza epidemics, *PNAS*, **101** (2004), 16915–16916.
44. C. Fuhrmann, The effects of weather and climate on the seasonality of influenza: What we know and what we need to know, *Geog. Compass*, **4** (2010), 718–730.
45. J. Shaman and M. Kohn, Absolute humidity modulates influenza survival, transmission, and seasonality, *Proc. Natl. Acad. Sci. USA*, **106** (2009), 3243–3248.

46. Ontario Ministry of Health and Long-Term Care, Universal Influenza Immunization Program, 2015. Available from: http://www.health.gov.on.ca/en/pro/programs/publichealth/flu/uiip/.

47. J. Truscott, C. Fraser, W. Hinsley, et al., Quantifying the transmissibility of human influenza and its seasonal variation in temperate regions, *PLOS Currents 1*.

48. E. Vynnycky, R. Pitman, R. Siddiqui, et al., Estimating the impact if childhood influenza vaccination programmes in england and wales, *Vaccine*, **26** (2008), 5321–5330.

49. L. F. White, J. Wallinga, L. Finelli, et al., Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the USA, *Influenza and Other Respiratory Viruses*, **3** (2009), 267–276.

50. C. Simpson, L. Ritchie, C. Robertson, et al., Effectiveness of H1N1 vaccine for the prevention of pandemic influenza in scotland, UK: A retrospective cohort study, *Lancet Infect. Dis.*, **12** (2012), 696–702.

51. K. Widgren, M. Magnusson, P. Hagstam, et al., Prevailing effectiveness of the 2009 influenza H1N1 pdm09 vaccine during the 2010/11 season in sweden, *Eurosurveillance*, **18** (2013), 20447.

52. J. Breteler, J. Tam, M. Jit, et al., Efficacy and effectiveness of seasonal and pandemic A (H1N1) 2009 influenza vaccines in low and middle income countries: A systematic review and meta-analysis, *Vaccine*, **31** (2013), 5168–5177.

53. J. Truscott, C. Fraser, S. Cauchemez, et al., Essential epidemiological mechanisms underpinning the transmission dynamics of seasonal influenza, *J. Royal Soc. Interface*, **9** (2011), 304–312.

54. S. Blower and H. Dowlatabadi, Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example, *Int. Statist. Rev.*, **62** (1994), 229–243.

55. M. Campitelli, M. Inoue, A. Calzavara, et al., Low rates of influenza immunization in young children under ontario's universal influenza immunization program, *Pediatrics*, 1421–1430.

56. J. J. Cannell, M. Zasloff, C. F. Garland, et al., On the epidemiology of influenza, *Virol. J.*, **5** (2008), 29.

57. S. Cauchemez, A. Valleron, P. Boelle, et al., Estimating the impact of school closure on influenza transmission from sentinel data, *Nature*, **452** (2008), 750.

58. G. Chowell, C. Ammon, N. Hengartner, et al., Estimation of the reproductive number of the Spanish flu epidemic in Geneva, Switzerland, *Vaccine*, **24** (2006), 6747–6750.

59. N. Cox and C. Bender, The molecular epidemiology of influenza viruses, *Sem. Virol.*, **6** (1995), 359–370.

60. A. C. Hayward, E. B. Fragaszy, A. Bermingham, Comparative community burden and severity of seasonal and pandemic influenza: Results of the Flu Watch cohort study, *Lancet Resp. Med.*, **2** (2014), 445–454.

61. S. Jiménez-Jorge, S. de Mateo, C. Delgado-Sanz, et al., Effectiveness of influenza vaccine against laboratory-confirmed influenza, in the late 2011-2012 season in Spain, among the population targeted for vaccination, *BMC Infect. Dis.*, **13** (2013), 441.

62. J. C. Kwong, H. Ge, L. C. Rosella, et al., School-based influenza vaccine delivery, vaccination rates, and healthcare use in the context of a universal influenza immunization program: An ecological study, *Vaccine*, **28** (2010), 2722–2729.

63. J. C. Kwong, T. A. Stukel, J. Lim, et al., The effect of universal influenza immunization on mortality and health care use, *PLOS Med.*, **5** (2008), 1440–1452.

64. K. Moran, S. Maaten, A. Guttmann, et al., Influenza vaccination rates in ontario children: Implications for universal childhood vaccination policy, *Vaccine*, **27** (2009), 2350–2355.

65. A. S. S. Rao, M. H. Chen, P. Ba'Z, et al., Cohort effects in dynamic models and their impact on vaccination programmes: an example from hepatitis a, *BMC Infect. Dis.*, **6** (2006), 174.

66. J. Shaman, C. Jeon, E. Giovannucci, et al., Shortcomings of vitamin d-based model simulations of seasonal influenza, *PLOS One*, **6** (2011), e20743.

67. J. Tamerius, C. Viboud, J. Shaman, et al., Impact of school cycles and environmental forcing on the timing of pandemic influenza activity in mexican states, may-december 2009, *PLOS Comput. Biol.*, **11** (2015), e1004337.

68. D. He, J. Dushoff, T. Day, et al., Mechanistic modelling of the three waves of the 1918 influenza pandemic, *Theor. Ecol.*, **4** (2011), 283–288.

AIMS Press