

# FluTE, a Publicly Available Stochastic Influenza Epidemic Simulation Model

Dennis L. Chao<sup>1\*</sup>, M. Elizabeth Halloran<sup>1,2</sup>, Valerie J. Obenchain<sup>1</sup>, Ira M. Longini, Jr.<sup>1</sup>

**1** Center for Statistics and Quantitative Infectious Diseases/Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America, **2** Department of Biostatistics, School of Public Health/University of Washington, Seattle, Washington, United States of America

## Abstract

Mathematical and computer models of epidemics have contributed to our understanding of the spread of infectious disease and the measures needed to contain or mitigate them. To help prepare for future influenza seasonal epidemics or pandemics, we developed a new stochastic model of the spread of influenza across a large population. Individuals in this model have realistic social contact networks, and transmission and infections are based on the current state of knowledge of the natural history of influenza. The model has been calibrated so that outcomes are consistent with the 1957/1958 Asian A(H2N2) and 2009 pandemic A(H1N1) influenza viruses. We present examples of how this model can be used to study the dynamics of influenza epidemics in the United States and simulate how to mitigate or delay them using pharmaceutical interventions and social distancing measures. Computer simulation models play an essential role in informing public policy and evaluating pandemic preparedness plans. We have made the source code of this model publicly available to encourage its use and further development.

**Citation:** Chao DL, Halloran ME, Obenchain VJ, Longini IM, Jr. (2010) FluTE, a Publicly Available Stochastic Influenza Epidemic Simulation Model. *PLoS Comput Biol* 6(1): e1000656. doi:10.1371/journal.pcbi.1000656

**Editor:** Angela R. McLean, University of Oxford, United Kingdom

**Received:** July 17, 2009; **Accepted:** December 22, 2009; **Published:** January 29, 2010

**Copyright:** © 2010 Chao et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the National Institute of General Medical Sciences MIDAS grant U01-GM070749, the National Institute of Allergy and Infectious Diseases grant R01-AI32042, and the Los Angeles County Department of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

\* E-mail: dchao@fhcrc.org

## Introduction

Mathematical and computer models of epidemics have contributed to our understanding of the spread of infectious disease and the measures needed to contain or mitigate them [1–9]. Detailed computer simulations will play an important role in evaluating containment and mitigation strategies for future epidemics [8]. Although many simulation models have been described in the literature, few are publicly available. Releasing the source code of models would allow others to evaluate the quality of the simulation, replicate results, and alter and improve the model.

We have released the source code for a new stochastic model of influenza epidemics, FluTE. FluTE is an individual-based model capable of simulating the spread of influenza across major metropolitan areas or the continental United States. The model's structure is based on previously published work [3,6], but FluTE incorporates a more sophisticated natural history of influenza, more realistic intervention strategies, and can run on a personal computer. Here, we describe the new model and illustrate how it can be used to study the dynamics of an epidemic and to investigate the population-level effects of interventions.

## Model

FluTE is an individual-based simulation model of influenza epidemics. In this section, we describe the model's community structure, natural history of influenza, and simulated interventions. Briefly, all individuals in the model are members of social mixing groups, within which influenza is transmitted by random mixing.

The model can simulate several intervention strategies, and these can either change the transmission characteristics of influenza (e.g., vaccination) or change the contact probabilities between individuals (e.g., social distancing). Interventions can occur before the epidemic or in response to an ongoing epidemic.

## Community structure and social contacts

The simulation creates synthetic populations based on typical American communities. The population is divided into census tracts, and each tract is subdivided into communities of 500–3000 individuals based on earlier models [6,10]. Each community is populated by randomly generated households of size 1–7 using the US-wide family size distribution from the 2000 Census (Table 1). The household is the closest social mixing group, within which contacts between individuals occur most frequently and thus influenza is transmitted most often. The population is organized as a hierarchy of increasingly large but less intimate mixing groups, from the household cluster (sets of four socially close households), neighborhoods (1/4 of a community), and the community. Although the model results are not sensitive to the exact size of these groups, including such groups creates a realistic contact network for disease transmission [11]. At night, everyone can make contact with other individuals in their families, household clusters, home neighborhoods, and home communities. In the daytime, individuals might interact with additional groups. During the day, most children attend school or a playgroup, where there is a relatively high probability of transmission. Preschool-age children usually belong to either a playgroup of four children or a neighborhood preschool, which typically has 14 students. Each

## Author Summary

Computer simulations can provide valuable information to communities preparing for epidemics. These simulations can be used to investigate the effectiveness of various intervention strategies in reducing or delaying the peak of an epidemic. We have made a detailed influenza epidemic simulator for the United States publicly available so that others may use the software to inform public policy or adapt it to suit their needs.

community has mixing groups that represent two elementary schools, one middle school, and one high school, which typically have 79, 128, and 155 students, respectively.

Most working-age adults (about 72% of 19–64 year-olds) are employed. Employment rates are determined on a tract-by-tract basis using data from the US Census 2000's Summary File 3, table PCT35. Employed individuals often work outside of their home communities. Each employed individual is assigned to work in a destination census tract based on commuting data taken from Part 3 of the Census Transportation Planning Package (<http://www.fhwa.dot.gov/ctpp/dataproduct.htm>), which provides information on the home and destination census tracts of workers in the United States. We eliminated commutes over 100 miles from the data as in [6] because many of these trips represent sporadic long-distance travel rather than daily commutes. Working individuals are assigned to communities and neighborhoods within their destination tracts to simulate casual community contacts during the day, and a work group of about 20 people to represent their close contacts at the workplace. Unemployed individuals remain in their home communities and do not have close daytime contacts except with members of their households who are not employed or enrolled in school.

Individuals can engage in short-term, long-distance domestic travel to represent vacations and other trips. Travel in our model is based on the implementation in [6], which uses data from the 1995 American Travel Survey data available from the U. S. Department of Transportation, Bureau of Transportation Statistics ([http://www.bts.gov/publications/national\\_transportation\\_statistics/](http://www.bts.gov/publications/national_transportation_statistics/)). Each day, an individual has a fixed probability of starting a trip based on an age-specific probability of traveling: 0.0023 for 0–4 year olds, 0.0023 for 5–18, 0.0050 for 19–29, 0.0053 for 30–64, and 0.0028 for 65 and older. The traveler will stay at the destination for 0–11 nights, with 23.9% of trips lasting for a single day (and no nights), 50.2% including 1–3 nights away, 18.5% including 4–7 nights away, and 7.4% for 8–11 nights. We do not include differences in travel frequency or duration during different times of the year (e.g., summer and holiday trips). The

**Table 1.** Frequency of household sizes.

Frequency	Family size
33%	single adult
34%	two people (two adults or a parent and child)
13%	two adults, one child
10%	two adults, two children
7%	two adults, three children
2%	two adults, four children
1%	two adults, five children

Data from [6].

doi:10.1371/journal.pcbi.1000656.t001

destination is a randomly selected census tract, in which a random community, neighborhood, and workplace (if the traveler is between 19 and 64 years old) are assigned to be the traveler's mixing groups. A random member of this community is assigned to be the traveler's contact person, and at night the traveler will behave as if he/she belongs to the contact's household, household cluster, and neighborhood. The traveler may withdraw to this household if ill. The exact implementation of short-term, long-distance travel is not important, but some long-distance travel is required in large populations for the epidemic to spread in a realistic manner. For simulations of smaller regions, such as a single county, there is no need to include long-distance travel.

New infected individuals are introduced to a simulation by infecting randomly selected people. This epidemic seeding process can occur once at the beginning of a simulation or daily. In addition, one can simulate an epidemic that is seeded from international travelers. In this scenario, randomly selected individuals in the counties with one of the United States' 15 busiest international airports are infected each day, proportional to the daily traffic of these airports (see Table 2).

## Influenza natural history and transmission

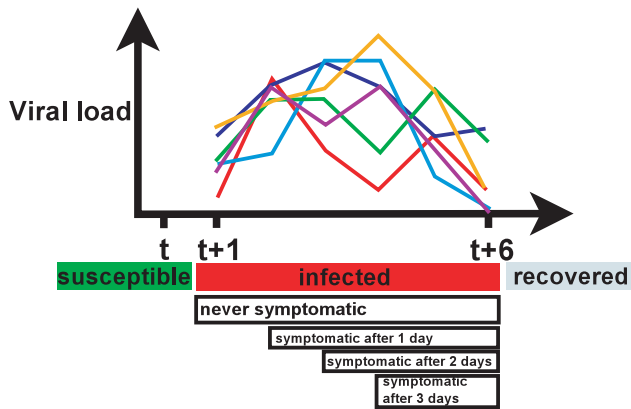
The current modeling of the natural history of influenza is as follows: An individual is infectious for six days starting the day after becoming infected. The individual's infectiousness is proportional to the log of the daily viral titers taken from a randomly chosen one of the six experimentally infected patients described in [12,13] (Figure 1). An individual is asymptomatic during the incubation period, which lasts from one, two, or three days (with 30%, 50%, and 20% probabilities, respectively). After incubation, the individual has a 67% chance of becoming symptomatic [14,15]. Symptomatic individuals are twice as infectious as asymptomatic people and may withdraw to the home after 0 to 2 days [16] (with probabilities summarized in Table 3). People who withdraw interact only with their households. Six days after infection, an individual recovers and is no longer susceptible.

**Table 2.** International traffic to the 15 US airports built into FluTE.

Airport	City	Passengers/year
JFK	New York, NY	21,842,544
LAX	Los Angeles, CA	17,019,166
MIA	Miami, FL	15,509,279
ORD	Chicago, IL	11,375,367
EWR	Newark, NJ	10,812,993
ATL	Atlanta, GA	9,166,055
SFO	San Francisco, CA	8,648,219
IAH	Houston, TX	7,627,942
IAD	Washington, DC	5,893,142
DFW	Dallas/Ft. Worth, TX	4,872,207
DTW	Detroit, MI	3,887,481
PHL	Philadelphia, PA	3,734,127
BOS	Boston, MA	3,673,748
FLL	Fort Lauderdale, FL	3,062,384
SEA	Seattle, WA	2,766,576

Data from [45].

doi:10.1371/journal.pcbi.1000656.t002



**Figure 1.** The natural history of influenza of simulated individuals in FLUTE. When a susceptible individual is infected (at time  $t$ ), that person will be infectious for six days with infectiousness proportional to his or her viral load. The six possible viral load trajectories are plotted. Most individuals become symptomatic, which occurs after a 1, 2, or 3 day incubation period. Symptomatic individuals are twice as infectious as asymptomatic individuals (i.e., infectiousness is proportional to twice the viral load). Individuals recover six days after infection and are immune.

doi:10.1371/journal.pcbi.1000656.g001

The simulation runs in discrete time, with two time steps per simulated day to represent daytime and nighttime social interactions. The contact probability of two individuals in the same mixing group is the probability that they will have sufficient contact for transmission during a time step. Contact probabilities of individuals within families were tuned so that the simulated household secondary attack rates match estimates from [17] (Table 4). Contact probabilities within other mixing groups were tuned so that the final age-specific illness attack rates were similar to past influenza pandemics (Table 5), particularly Asian A (H2N2) and 2009 novel influenza A(H1N1) influenza, and the percentage of transmissions that can be attributed to each mixing group matched those in [6,18–20], although these values depend on the transmissibility ( $R_0$ ) of the disease (Table 6). These contact probabilities are in general agreement with other simulation models [8] and with a recent study of physical contacts between individuals [21]. Contact probabilities for all types of mixing groups are summarized in Table 7.

Transmission probabilities in the simulation are adjusted by multiplying all contact probabilities by a scalar,  $P_{trans}$ , to obtain the desired  $R_0$ , the basic reproductive number, which is defined as the average number of secondary infections from a typical infected individual in a fully susceptible population [22]. To derive the relationship between  $R_0$  and  $P_{trans}$ , we infected a single randomly selected person in an otherwise fully susceptible 2000-

**Table 3.** Probabilities that an individual will withdraw to the home 0, 1, or 2 days after becoming symptomatic.

Age group	0 days	1 day	2 days
Preschool-age children	0.304	0.575	0.324
School-age children	0.203	0.498	0.375
Adults	0.100	0.333	0.167

Data from [16].

doi:10.1371/journal.pcbi.1000656.t003

**Table 4.** Estimates of secondary household attack rates from [17] and illness attack rates using FluTE, stratified by the ages of the index and secondary cases.

		Exposed			
		Addy 1991		simulated ( $R_0 = 1.30$ )	
		child	adult	child	adult
Infectious	child	29.0%	14.2%	28.6%	13.5%
	adult	10.3%	15.6%	9.3%	16.2%

doi:10.1371/journal.pcbi.1000656.t004

person community with a 74% working-age adult employment rate and counted the number of individuals that person infected, repeating this procedure 1,000 times for several values of  $P_{trans}$ . The relationship between  $P_{trans}$  the average number of secondary cases was approximately linear for a biologically plausible range of values:  $4.768P_{trans} + 0.040$  (Figure 2). However, the average number of secondary cases was higher when the index case was a child because children tend to infect more individuals (and become infected more often) than adults. Therefore, in a procedure borrowed from [6], we measured the age distribution of secondary cases when the index case was randomly selected and used this distribution to weight the contribution from the various age groups to the  $R_0$  calculation to define  $R_0 = 5.592P_{trans} + 0.068$ . The definition of  $R_0$  applies to a population with no pre-existing immunity, an assumption that may be violated for seasonal influenza. One can use the model to simulate seasonal influenza epidemics by substituting  $R_0$  with the desired  $R$ , the average number of people a typical infected case infects in a population with pre-existing immunity.

The simulated case generation time, or the time between infection of an individual and the transmission to susceptibles, was 3.4 days for a wide range of  $R_0$  in a fully susceptible population (Figure 2B). This is consistent with other estimates for seasonal and pandemic influenza [20,23].

### Simulated interventions

The primary pharmaceutical intervention is vaccination. Vaccinated individuals in the simulation have a reduced

**Table 5.** Age-specific influenza illness attack rates in past influenza epidemics (from [46]) and in a simulation of metropolitan Seattle.

Age group	Asian A (H2N2)	Hong Kong A (H3N2)	Age group	simulated ( $R_0 = 1.6$ )
	1957–8	1968–9		
Pre-school children	35%	34%	0–4 years	38%
School-age children	55%	35%	5–18 years	53%
Young adults	25%	35%	19–29 years	26%
Middle adults	20%	32%	30–64 years	28%
Old adults	14%	31%	≥ 65 years	23%
Overall	31%	34%		33%

doi:10.1371/journal.pcbi.1000656.t005

**Table 6.** Major sources of influenza transmission in simulations of metropolitan Seattle.

Mixing group	Fraction of transmissions		
	$R_0 = 1.2$	$R_0 = 1.6$	$R_0 = 2.0$
household	32%	31%	29%
schools/daycares	30%	24%	21%
workplace	10%	13%	15%
neighborhood/ community	18%	21%	23%

doi:10.1371/journal.pcbi.1000656.t006

probability of becoming infected ( $VE_S$ ), of becoming ill given infection ( $VE_P$ ), and of transmitting infection ( $VE_I$ ) [24]. In the model, these efficacy parameters are implemented by multiplying the transmission probability per time step by  $(1 - VE_S)$  if the susceptible individual is vaccinated and by  $(1 - VE_I)$  if the infectious individual is vaccinated. The probability of vaccinated individuals becoming symptomatic (ill) after they are infected is the baseline probability (67%) multiplied by  $(1 - VE_P)$ .

Vaccines do not reach full efficacy immediately – their protective effects may gradually increase over several weeks. The default behavior in the model is that the vaccine takes two weeks to reach maximum efficacy, with the efficacy increasing exponentially starting the day after the vaccination. Because of the delay in reaching maximum efficacy, it may be necessary to vaccinate the population early. In the simulation, vaccines can be administered at least four weeks before the epidemic (i.e., pre-vaccination), during the epidemic (reactive), or one dose can be administered at least three weeks before the epidemic and the boost can be administered reactively (prime-boost).

Antiviral agents (neuraminidase inhibitors) can be used for treatment of cases and for prophylaxis of susceptibles. A single course of antiviral agents is enough for 10 days of prophylaxis or 5 days of treatment. In the model, 5% of individuals taking antiviral agents prophylactically stop after 2 days and 5% taking them for treatment stop after 1 day [19]. As with vaccines, individuals taking antiviral agents can have reduced susceptibility ( $AVE_S$ ),

probability of becoming ill given infection ( $AVE_P$ ), and transmitting infection ( $AVE_I$ ). However, unlike vaccines, the protective effects of the antiviral agents last only as long as they are being taken (5 to 10 days). When a case is ascertained, the individual is treated with antiviral agents, and that individual's household members will also each be given a course if household targeted antiviral prophylaxis (HHTAP) is in effect.

Several non-pharmaceutical interventions can be simulated in the model. *School closures* are simulated by eliminating school group contacts (including preschools and daycares but not playgroups) for those enrolled in school, but adding daytime contacts with other household members not in school or at work and doubling their daytime neighborhood and community contact probabilities to account for their non-school activities. Schools can be closed when cases are ascertained in communities or in the schools, and they can be closed for a fixed number of days or for the duration of the simulation.

During an epidemic, individuals may be requested to stay at home if they become ill. When simulating *isolation* of cases, individuals withdraw to the home one day after becoming symptomatic (with a certain probability to represent the compliance probability). This will eliminate any daytime social contacts that they have other than with household members who are not working or at school. We simulate a *liberal leave policy* in a similar manner: employed individuals withdraw to the home with a pre-set compliance probability for one week one day after becoming symptomatic.

During an epidemic, those living with symptomatic individuals may be requested to stay home [25]. In simulations of *household quarantine*, family members of symptomatic individuals will independently decide (based on a compliance probability) whether to obey quarantine for 7 days one day after the first individual becomes symptomatic. Individuals electing to quarantine themselves withdraw to the household and interact only with household members. If other family members become ill during quarantine, household members independently decide whether to obey quarantine for 7 days one day after each individual becomes symptomatic.

### Implementation of the stochastic model

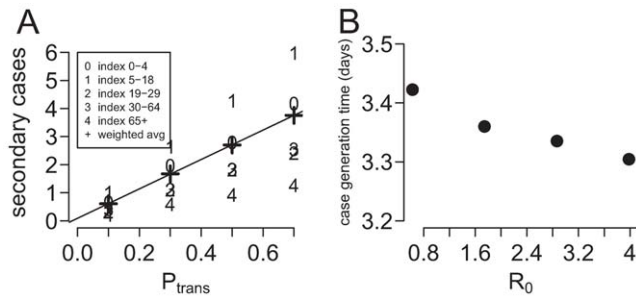
FluTE is written in C/C++ and is released under the GNU General Public License (GPLv3, see <http://www.gnu.org/licenses/gpl.html>). The source code is available at <http://www.csquid.org/software>, <https://www.epimodels.org/midas/flute.do>,

**Table 7.** Person-to-person contact probabilities for all social mixing groups in FluTE.

	Exposed				
	child 0–4	child 5–18	adult 19–29	adult 30–64	adult 65+
Family, infectious is child	0.8	0.8	0.35	0.35	0.35
Family, infectious is adult	0.25	0.25	0.4	0.4	0.4
Household cluster, infectious is child	0.08	0.08	0.035	0.035	0.035
Household cluster, infectious is adult	0.025	0.025	0.04	0.04	0.04
Neighborhood	0.0000435	0.0001305	0.000348	0.000348	0.000696
Community	0.0000109	0.0000326	0.000087	0.000087	0.000174
Workplace			0.05	0.05	
Playgroup	0.28				
Daycare	0.12				
Elementary school		0.0348			
Middle school		0.03			
High school		0.0252			

doi:10.1371/journal.pcbi.1000656.t007





**Figure 2. Influenza transmission properties in the simulation.** (A) Observed secondary cases vs  $P_{trans}$  by the age of the index case and the weighted average. (B) Average case generation time vs  $R_0$ .

doi:10.1371/journal.pcbi.1000656.g002

and the Models of Infectious Disease Agent Study (MIDAS) repository [26]. The software includes two source code files that are also freely distributable but may come with different licenses because they were written by others: one for the pseudorandom number generator (SIMD oriented Fast Mersenne Twister (SFMT) pseudorandom number generator [27]) and one to generate binomially distributed random numbers (from Numerical Recipes in C [28]). Version 1.11 of FluTE was used to produce the results in this manuscript.

A configuration file is used to specify the population to use for the simulation, the parameters for starting the epidemic, the transmissibility of the infectious agent, and the desired intervention strategies. The configuration file is text-based and can be typed in by a user or generated with a script. The simulation outputs results to text files, which can be easily parsed for plotting or statistical analysis.

A parallelized version of the code supports simulations of large populations (up to the entire continental United States). This version of the program assigns the populations of different counties to different processors, and OpenMPI is used to update the status of individuals who travel between communities that are located on different processors and to update the global status of the epidemic and the interventions (e.g., the total number of vaccines used). The simulation uses approximately 80 megabytes of memory per million simulated individuals.

The simulation was written with several competing goals: to explicitly represent each individual in the population, to conserve memory, to run quickly, and to be (relatively) easy to read and modify. Each simulated individual is represented by a C structure that includes unique identifiers for the person and for each of the social mixing groups to which that person belongs, the age of the individual, the person's infection and vaccination status and dates, and other attributes. For each infected individual, the simulation identifies all susceptible individuals in that person's community who share a common mixing group, the infectiousness of the infected individual, the susceptibility of the susceptible, and the probability that transmission takes place for every time step. Although comparing each individual with every other within a community results in the number of comparisons increasing with the square of the number of individuals, community sizes are always smaller than 3,000 residents. Therefore, the number of comparisons made between individuals scales approximately linearly with the number of individuals in the simulation. More sophisticated algorithms could improve the simulation's performance, but may do so at the expense of the code's flexibility and readability.

The running time depends on the number of individuals infected during the course of a simulation. Simulating an epidemic in a population of 10 million people can take up to two hours (on a single processor on an Intel Core2 Duo T9400), but it may take only seconds if the virus is not highly transmissible (low  $R_0$ ) or if there are effective interventions (e.g., high vaccination rates). On a cluster of 32 processors, simulating an epidemic covering the continental United States (population of 280 million) takes about 6 hours (192 hours of total CPU time).

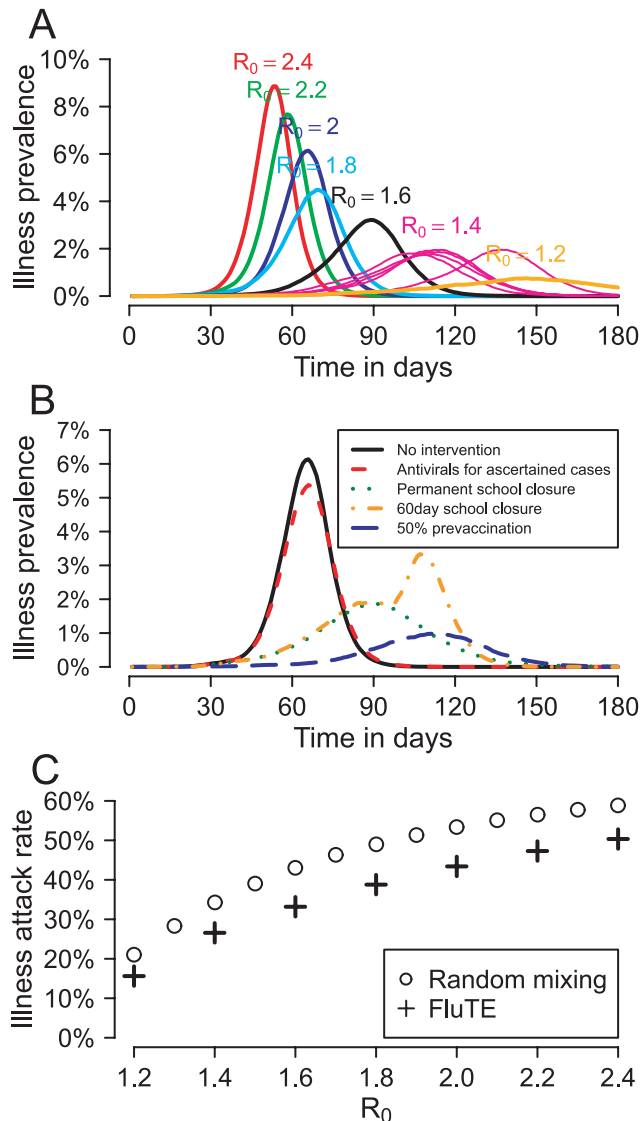
## Results

We illustrate the use of the model by simulating epidemics in metropolitan Seattle, a major metropolitan area with a population of approximately 560,000 according to the US 2000 Census. We ran simulations with different values of  $R_0$ , starting with ten infected individuals chosen at random, and found that the epidemic could peak as early as 45 days after the start if  $R_0$  is high ( $R_0=2.4$ ) (Figure 3A). Pre-vaccination (with vaccine efficacies of  $VE_S=40\%$ ,  $VE_P=67\%$ ,  $VE_I=40\%$ , which correspond to a well-matched seasonal influenza vaccine [29]) is likely to both lower and delay the epidemic peak (Figure 3B). Use of antivirals alone ( $AVE_S=30\%$ ,  $AVE_P=60\%$ , and  $AVE_I=62\%$  [11]) did not greatly reduce the epidemic peak, but they could reduce illness and mortality in an epidemic. Non-pharmaceutical interventions could be quite effective, but the epidemic may spike immediately upon ending the intervention (compare permanent school closure with school closure for 60 days in Figure 3B).

The illness attack rates in the simulation are lower than those in a SIR model with random mixing (where  $1-AR=e^{-R_0AR}$  [30], where AR is the infection attack rate, and the illness attack rate is  $0.67 \times AR$ ) (Figure 3C). As observed in earlier studies, models with community structure have lower attack rates than those with random mixing [31–33].

Simulated epidemics struck school-age children earlier than adults, which had been observed in earlier studies [6,34]. Therefore, we predict that early in an epidemic, the proportion of cases who are school-age children will be higher than later in the epidemic (Figure 4). This phenomenon might affect the accuracy of  $R_0$  estimates in unfolding epidemics. For example, most confirmed cases in the recent novel influenza A(H1N1) outbreaks in the United States have been school-age children [35] and several early estimates of  $R_0$  have been above 2 [36,37]. In our model, we observed that infected children generate more secondary cases than infected adults (Figure 2A). For example, infected school-age children would transmit to an average of 2.53 other individuals in a simulated epidemic with  $R_0=1.6$ . Therefore, estimates of  $R_0$  could be high early in an epidemic when a disproportionate number of infections are in children.

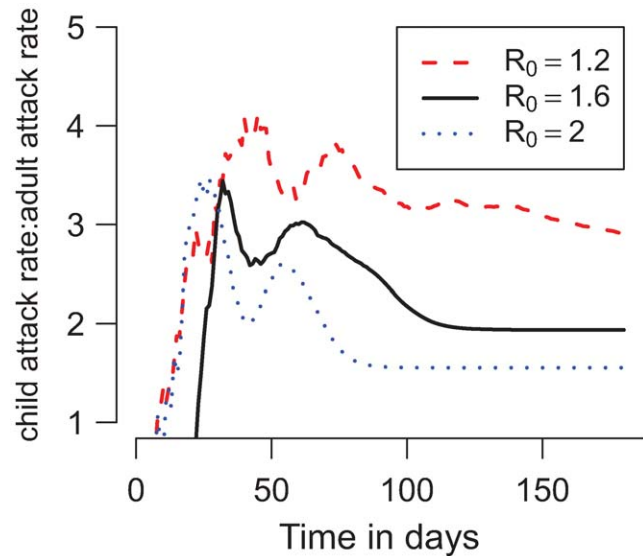
One can simulate the population of the entire continental US using the parallel version of FluTE (mpiflute). The continental US had 280 million people in 64735 census tracts in 2000, based on the US 2000 Census. In our simulations, we found that the final illness attack rates for the US to be nearly identical to those of metropolitan Seattle, but the epidemic peak for a given  $R_0$  is later for the United States (e.g., 94 vs 65 days for  $R_0=1.6$ ) (Figure 5). Therefore, simulations of a sufficiently large metropolitan area may be adequate for determining the effect of a strategy on the national level on final illness attack rates, but the nation-wide peak of the epidemic may be later than in the major metropolitan areas because of the time it takes the epidemic to reach outlying areas.



**Figure 3. Illness attack rates and daily prevalence of influenza in simulations of metropolitan Seattle.** (A) Daily prevalence of symptomatic influenza in simulations of metropolitan Seattle for various  $R_0$  and (B) for  $R_0=2.0$  with various interventions. The interventions, which begin 30 days after the first case is detected, are: giving a course of antiviral agents to ascertained cases, closing schools either permanently or for 60 days, and pre-vaccination of 50% of the population with a well-matched seasonal influenza vaccine. (C) Final illness attack rates (180 days) vs  $R_0$  for FluTE (simulating metropolitan Seattle) and a model with random mixing. Results for all panels are from one run of metropolitan Seattle for each  $R_0$  or intervention strategy except for the simulation for  $R_0=1.4$  in panel (A), which was run 5 times with different random number seeds and plotted to show stochastic variability. doi:10.1371/journal.pcbi.1000656.g003

## Discussion

We have described a new publicly available influenza epidemic simulator, FluTE. It explicitly represents every individual in the simulation, so simulated epidemics can be studied in detail, even tracing individual transmission events. We illustrated the use of FluTE with examples in which we explored the effect of various intervention strategies on influenza epidemics in the United States and showed how transmissibility can be over-estimated early in an epidemic.

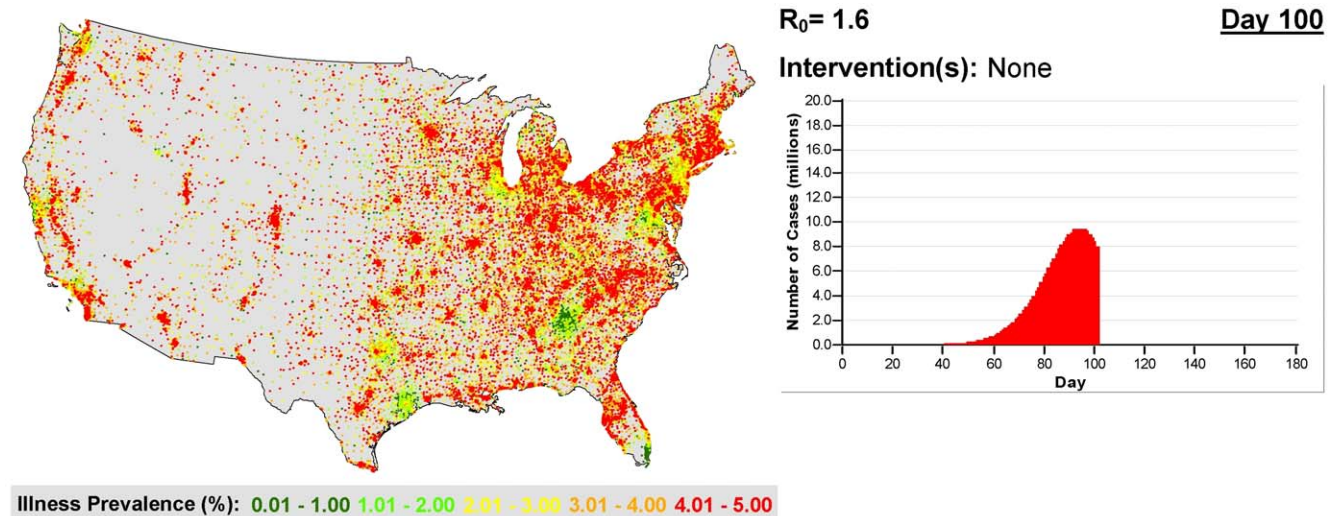


**Figure 4. The ratio of cumulative illness attack rates between school-age children (ages 5–18) and adults (ages 19–64) over time in simulated epidemics.** Results plotted are from one simulation of metropolitan Seattle for each value of  $R_0$ . doi:10.1371/journal.pcbi.1000656.g004

The simulation was written so that one can easily set the transmissibility, vaccination policies (e.g., fraction of the population to vaccinate), and other reactive strategies (e.g., school closures). These settings can be used to investigate questions such as: 1) What fraction of the population will become infected or ill? 2) How much vaccine coverage is required to mitigate an epidemic with a given  $R_0$ ? 3) What segment of the population should be vaccinated to reduce overall illness attack rates the most? 4) How long can one wait before reacting to an epidemic? and 5) What range of  $R_0$  can be managed by a particular pandemic strategy? We have used FluTE to investigate some of these questions by simulating vaccinating children against seasonal and pandemic influenza [38] and pandemic mitigation [20].

The model was calibrated to simulate epidemics of a virus similar to 1957/1958 Asian A(H2N2) and 2009 pandemic A(H1N1). We attempted to model realistic pharmaceutical and non-pharmaceutical interventions, but their effects on an epidemic have not been well quantified. The model's results are plausible and likely to be qualitatively correct, but there is insufficient data to calibrate it to produce quantitatively accurate results for the various possible disease parameters and mitigation strategies. Although the model generates realistic population-level results, the spatial dynamics of the epidemics it produces should be used for illustrative purposes only. When using the model to evaluate mitigation strategies, it is important to consider one's goals. For example, using antiviral agents to treat cases does not greatly reduce the final illness attack rate in the simulation, but it could greatly reduce mortality. The model does not directly evaluate the cost of interventions, but the numbers of cases in a simulated epidemic can be linked to cost and healthcare utilization data [39].

Differential equation models are the most popular approach to disease modeling. The simplest of these (such as the SIR model [40]) can be used to study epidemics analytically, and more complex versions have been used to model the dynamics of epidemics on a global scale [41,42]. However, if one wants to include a complicated natural history of disease or detailed



**Figure 5. The prevalence of influenza in a single simulation of the United States 100 days after the start of an influenza epidemic with  $R_0 = 1.6$ .** The color of each dot corresponds to the illness prevalence in a census tract. Image created using ArcGIS (Environmental Systems Research Institute, Inc.)  
doi:10.1371/journal.pcbi.1000656.g005

intervention strategies, individual-based models, such as FluTE, may be more suitable.

The current software supports a limited set of configuration options and is intended for batch runs using a scripting language. Using the model for scenarios not supported by the existing code, such as testing a novel intervention strategy or altering the contact parameters for a different attack rate pattern, would require modification of the source code, which we have released so that others can make such changes if needed. We decided to adopt the GNU General Public License (GPL), so that the source code of derivative works must be released. We believe this will facilitate the sharing of improvements. The availability of source code allows others to adapt the model to simulate outbreaks of other airborne infectious diseases such as smallpox [3,43,44] or to simulate other regions of the world with different social structures [3].

In the future, we would like to make our model more accessible to non-programmers. This may involve developing a user interface or adding new parameters to the configuration file. We would also like to include intervention strategies that best reflect government pandemic mitigation plans. Achieving these goals would depend upon close collaboration with public health officials to better

understand their needs and to carefully simulate existing pandemic mitigation plans and capacities. Although we have calibrated our model to the best available data, more detailed and reliable information on the natural history of influenza, influenza transmission, human behavior in response to infection, and vaccine efficacy is needed. Sensitivity analyses of similar epidemic models have shown that results are robust to uncertainty in many parameters [3,5,6,11]. However, more accurate model inputs would improve the quantitative predictions. Well-designed studies are needed to acquire these data.

## Acknowledgments

We thank Brandon Dean for helpful discussions and Jon Sugimoto for producing the image in Figure 5.

## Author Contributions

Conceived and designed the experiments: DLC MEH IMLJ. Performed the experiments: DLC. Analyzed the data: DLC. Contributed reagents/materials/analysis tools: DLC VJO. Wrote the paper: DLC MEH VJO IMLJ.

## References

- Ross R (1916) Application of the theory of probabilities to the study of a *priori* pathometry. part I. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character 92: 204–230.
- Elveback L, Varma A (1965) Simulation of mathematical models for public health problems. Public Health Rep 80: 1067–76.
- Halloran ME, Longini IM Jr, Nizam A, Yang Y (2002) Containing bioterrorist smallpox. Science 298: 1428–32.
- Eubank S, Guclu H, Kumar VSA, Marathe MV, Srinivasan A, et al. (2004) Modelling disease outbreaks in realistic urban social networks. Nature 429: 180–4.
- Ferguson NM, Cummings DAT, Fraser C, Cajka JC, Cooley PC, et al. (2006) Strategies for mitigating an influenza pandemic. Nature 442: 448–52.
- Germann TC, Kadau K, Longini IM Jr, Macken CA (2006) Mitigation strategies for pandemic influenza in the United States. Proc Natl Acad Sci U S A 103: 5935–40.
- Hatchett RJ, Mecher CE, Lipsitch M (2007) Public health interventions and epidemic intensity during the 1918 influenza pandemic. Proc Natl Acad Sci U S A 104: 7582–7.
- Halloran ME, Ferguson NM, Eubank S, Longini IM Jr, Cummings DA, et al. (2008) Modeling targeted layered containment of an influenza pandemic in the United States. Proc Natl Acad Sci U S A 105: 4639–44.
- Coburn BJ, Wagner BG, Blower S (2009) Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1). BMC Med 7: 30.
- Halloran ME, Longini IM, Cowart DM, Nizam A (2002) Community interventions and the epidemic prevention potential. Vaccine 20(27–28): 3254–62.
- Longini IM Jr, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, et al. (2005) Containing pandemic influenza at the source. Science 309: 1083–7.
- Murphy BR, Rennels MB, Douglas RG Jr, Betts RF, Couch RB, et al. (1980) Evaluation of influenza A/Hong Kong/123/77 (H1N1) *ts-1A2* and cold-adapted recombinant viruses in seronegative adult volunteers. Infect Immun 29: 348–55.
- Baccam P, Beauchemin C, Macken CA, Hayden FG, Perelson AS (2006) Kinetics of influenza A virus infection in humans. J Virol 80: 7590–9.
- Halloran ME, Hayden FG, Yang Y, Longini IM Jr, Monto AS (2007) Antiviral effects on influenza viral transmission and pathogenicity: observations from household-based trials. Am J Epidemiol 165: 212–21.
- Carrat F, Vergu E, Ferguson NM, Lemaître M, Cauchemez S, et al. (2008) Time lines of infection and disease in human influenza: a review of volunteer challenge studies. Am J Epidemiol 167: 775–85.
- Elveback LR, Fox JP, Ackerman E, Langworthy A, Boyd M, et al. (1976) An influenza simulation model for immunization studies. Am J Epidemiol 103: 152–65.

17. Addy CL, Longini IM Jr, Haber M (1991) A generalized stochastic model for the analysis of infectious disease final size data. *Biometrics* 47: 961–74.
18. Jordan WS Jr (1961) The mechanism of spread of Asian influenza. *Am Rev Respir Dis* 83: 29–40.
19. Longini IM Jr, Halloran ME, Nizam A, Yang Y (2004) Containing pandemic influenza with antiviral agents. *Am J Epidemiol* 159: 623–33.
20. Yang Y, Sugimoto JD, Halloran ME, Basta NE, Chao DL, et al. (2009) The transmissibility and control of pandemic influenza A (H1N1) virus. *Science* 326: 729–733.
21. Mossong J, Hens N, Jit M, Beutels P, Auranen K, et al. (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med* 5: e74.
22. Diekmann O, Heesterbeek JA, Metz JA (1990) On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *J Math Biol* 28: 365–82.
23. Cowling BJ, Fang VJ, Riley S, Peiris MJS, Leung GM (2009) Estimation of the serial interval of influenza. *Epidemiology* 20: 344–7.
24. Halloran ME, Longini IM Jr, Struchiner CJ (1999) Design and interpretation of vaccine field studies. *Epidemiol Rev* 21: 73–88.
25. Centers for Disease Control (2007) Interim Pre-Pandemic Planning Guidance: Community Strategy for Pandemic Influenza Mitigation in the United States — Early, Targeted, Layered Use of Nonpharmaceutical Interventions. USA: Department of Health & Human Services.
26. Cooley PC, Roberts D, Bakalov VD, Bikmal S, Cantor S, et al. (2008) The model repository of the Models of Infectious Disease Agent Study. *IEEE Trans Inf Technol Biomed* 12: 513–22.
27. Saito M, Matsumoto M (2008) SIMD-oriented Fast Mersenne Twister: a 128-bit pseudorandom number generator. In: Monte Carlo and Quasi-Monte Carlo Methods 2006 Springer Berlin Heidelberg. pp 607–622. 10.1007/978-3-540-74496-2.
28. Press WH, Flannery BP, Teukolsky SA, Vetterling WT, eds (1995) *Numerical Recipes. The Art of Scientific Computing*. Cambridge: Cambridge Univ. Press.
29. Basta NE, Halloran ME, Matrajt L, Longini IM Jr (2008) Estimating influenza vaccine efficacy from challenge and community-based study data. *Am J Epidemiol* 168: 1343–52.
30. Longini IM Jr (1986) The discrete-time general epidemic model: a synthesis. *Mathematical Biosciences* 82: 19a–41.
31. Elveback L, Ackerman E, Gatewood L, Fox JP (1971) Stochastic two-agent epidemic simulation models for a community of families. *Am J Epidemiol* 93: 267–80.
32. Ball F, Mollison D, Scalia-Tomba G (1997) Epidemics with two levels of mixing. *The Annals of Applied Probability* 7: 46–89.
33. Becker NG, Utev S (1998) The effect of community structure on the immunity coverage required to prevent epidemics. *Math Biosci* 147: 23–39.
34. Wallinga J, Teunis P, Kretzschmar M (2006) Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am J Epidemiol* 164: 936–44.
35. Dawood FS, Jain S, Finelli L, Shaw MW, Lindstrom S, et al. (2009) Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N Engl J Med* 360: 2605–2615.
36. Boëlle PY, Bermillon P, Desenclos JC (2009) A preliminary estimation of the reproduction ratio for new influenza A(H1N1) from the outbreak in Mexico, March–April 2009. *Euro Surveill* 14: 19205.
37. Nishiura H, Castillo-Chavez C, Safan M, Chowell G (2009) Transmission potential of the new influenza A(H1N1) virus and its age-specificity in Japan. *Euro Surveill* 14: 19227.
38. Basta NE, Chao DL, Halloran ME, Matrajt L, Longini IM Jr (2009) Strategies for pandemic and seasonal influenza vaccination of schoolchildren in the United States. *Am J Epidemiol* 170: 679–86.
39. Sander B, Nizam A, Garrison LP Jr, Postma MJ, Halloran EM, et al. (2009) Economic evaluation of influenza pandemic mitigation strategies in the United States using a stochastic microsimulation transmission model. *Value in Health* 12: 226–233.
40. Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London Series A* 115: 700–721.
41. Colizza V, Barrat A, Barthelemy M, Valleron AJ, Vespignani A (2007) Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS Med* 4: e13.
42. Kaufman J, Edlund S, Douglas J (2009) Infectious disease modeling: Creating a community to respond to biological threats. *Statistical Communications in Infectious Diseases* 1.
43. Ferguson NM, Keeling MJ, Edmunds WJ, Gani R, Grenfell BT, et al. (2003) Planning for smallpox outbreaks. *Nature* 425: 681–5.
44. Longini IM Jr, Halloran ME, Nizam A, Yang Y, Xu S, et al. (2007) Containing a large bioterrorist smallpox attack: a computer simulation approach. *Int J Infect Dis* 11: 98–108.
45. U S Department of Transportation (2008) U. S. International Air Passenger and Freight Statistics. Office of the Assistant Secretary for Aviation and International Affairs.
46. Longini IM, Ackerman E, Elveback LR (1978) An optimization model for influenza A epidemics. *Mathematical Biosciences* 38: 141–157.