# Title: Pandemic Influenza Mitigation Policy: Local Population Demographics Matter

# Abstract

The 2009 A (H1N1) influenza pandemic demonstrated that society is inadequately prepared to prevent and mitigate the spread of the disease. In this study, a network agent-based model was used to simulate influenza pandemic to determine if demographics play a key role in the application of mitigation resources. The output from the simulation was analyzed with robust and conservative statistical methods, and this study found that demographics play a critical role in the application of scarce pandemic influenza mitigation resources.

The 2009 A H1N1 pandemic demonstrated that society is not well prepared to prevent or mitigate the spread of emerging and reemerging strains of pandemic influenza. Despite the high profile vaccination campaigns in many countries during the 2009 H1N1 pandemic, high levels of vaccination were not achieved [1]. Arguably worse, the 2009 A H1N1 influenza pandemic demonstrated that the global production levels of the A(H1N1)pdm09 vaccine fell well short of the World Health Organization’s vaccine production estimate [2]. Even if the pandemic influenza vaccine can be rapidly produced in massive quantities, the vaccine must be highly effective to prevent the spread of the disease through the population. Unfortunately, the effectiveness of seasonal and pandemic influenza vaccines is highly variable and may not provide adequate protection across all age groups [3]. For these reasons, society cannot solely rely upon pandemic influenza vaccines as a mitigation strategy to protect the public’s health. Until a highly effective and reliable pandemic influenza vaccine is developed, the scientific community must identify cost effective and realistic alternative solutions to prevent the spread of pandemic influenza strains, identify methods to determine which influenza mitigation policies are the most effective, and determine how and where to best deploy limited influenza mitigation resources (e.g., personal protective equipment, antiviral treatment, and prophylaxis)

There have been many studies that have evaluated pandemic influenza mitigation policies (e.g., adult social distancing, child social distancing, quarantines, school closures, and public transportation bans) and mitigation treatments or technologies (e.g., vaccination, household member antiviral prophylaxis, and antiviral treatment) to determine which ones are optimal [4,5,6,7,8,9]. These studies have used a variety of methods to evaluate influenza prevention and response policies, and have predominantly relied upon agent based models, system dynamics models, and epidemiological studies. Most of the agent based and systems dynamics models created after the 2009 H1N1 pandemic used the observed H1N1 attack rate of 1.4 that was collected from retrospective epidemiologic studies [1, 6, 8, 9,10,11,12]. Less frequently, researchers have varied *R0* in pandemic influenza models to determine what effect mitigation strategies have across varying levels of the basic reproductive number. Varying *R0* in pandemic influenza, across all types of models, has demonstrated that effective pandemic mitigation strategies vary as *R0* increases. Specifically, some mitigation strategies work relatively well with low basic reproduction rates (1.2-1.6) and mitigation strategies typically do not work well with higher reproductive rates (*R0* >1.6). For these reasons, understanding the relationship between the disease transmission rate, the age-structure of pandemic influenza cases, and the available mitigation strategies is imperative to creating and implementing effective pandemic influenza mitigation policy.

Although understanding pandemic influenza disease transmission dynamics is important to selecting the best mitigation strategies, the disease’s transmission dynamics are typically not well understood or known until the initial outbreak has been investigated in detail. For example, in the 2009 H1N1 pandemic the first published report indicated three different basic reproductive rates: epidemiologic analyses indicated a *R0* between 1.4 and 1.6 in Mexico and a median *R0* of 1.22 in laboratory confirmed cases [13]. Furthermore, the 2009 type A H1N1 influenza strain probably emerged in September of 2008 and biosurveillance was not able detect the emergence of the disease until April of 2009 in Mexico [14]. Also, previous studies have argued that if severity and the strain’s basic reproductive rate are not highly correlated, then mitigation strategies that reduce mortality may be different than mitigation strategies that reduce prevalence [15]. Since the effectiveness of pandemic influenza mitigation strategies is partially dependent on the strain’s basic reproductive number and strain severity, and disease transmission rates are ambiguous during the initial phase of a pandemic influenza outbreak, using epidemic models to determine which mitigation strategies are most prudent, with an unknown *R0*, has insufficient value to the people who establish influenza control policy and distribute pandemic influenza mitigation resources.

Influenza pandemics are composed of multiple slightly desynchronized local epidemics, and in the United States the demographic data of these local epidemics are easily obtained through publicly available sources [16]. One way to potentially increase the value of pandemic influenza modeling is to incorporate the local population’s demographics into the analysis. Population demographics can be used to infer more specific population characteristics (e.g., family size and structure) and can be used to build more predictive pandemic influenza models. For some time, researchers have known that households with more children have higher incidence of seasonal influenza infections and household structure can be inferred from the demographic composition of a geographic area [17,18,19, 20]. In addition to inferring household contact structure from demographic data, the complexity of contact networks can be inferred from these same data. Glass et al.[21] found that the contact networks vary on location and over time, and, more importantly, that not including demographics into analyses over highly heterogeneous contact networks results in masking critical contact structure features and subsequently causes the selection of inefficacious pandemic influenza mitigation strategies. Previous studies have also found that household crowding, social economic status, race, work leave policies, and transportation types used by the population is correlated with the incidence of influenza like illnesses [22]. Therefore, incorporating the target population’s demographic characteristics into pandemic influenza models may yield more accurate predictions of the spread and burden of the disease. For these reasons, pandemic influenza models should consider the demographic characteristics of the population and the desired outcomes of the policy before ranking mitigation strategies.

This study uses an existing networked agent-based pandemic influenza model developed by National Infrastructure Simulation and Analysis Center (NISAC) to determine: (1) if differences in demographic characteristics communities results in consistent changes in the recommended mitigation strategies during an influenza pandemic; and, (2) whether changes in demographics have consistent or equal effects across all pandemic influenza mitigation strategies.

# Method

*Network agent-based model summary*

The Loki-Infect 3 is a social network agent-based model that was developed by NISAC to study epidemics within local populations and to determine epidemic mitigation strategies [23]. Loki-Infect 3 simulates disease propagation and mitigation strategies using networks of agents that represent individuals of different age classes. The agents are linked to each other and within and among social groups (e.g., households, neighborhoods, school classrooms, businesses) to form contact networks reflective of a structured community where individuals’ social networks overlap (see Supplement A for detailed contact network parameters). Behavioral rules for individuals and network topologies are parameterized to model the spread of influenza under multiple conditions, and the behavioral rules are specified to reflect the target populations’ available mitigation strategies during a simulation.

Simulations begin by creating a community (composed of 10,000 individuals in this study) and then seeding the community with an initial group of randomly selected infected adults (cases). Cases may infect susceptible individuals, depending on their contacts’ age-related susceptibility. In this study, school closure, child social distancing, adult social distancing, household quarantine, antiviral treatment, household member antiviral prophylaxis, adult *and* child social distancing, school closure *and* adult social distancing, and school closure *and* child social distancing were the mitigation strategies simulated with a 90% compliance rate (see Supplement B for detailed definitions of the mitigation strategies).

Mitigation strategies are applied at predefined number of incident cases (10, 30,

100) and mitigation strategies cease within a seven-day period if cases fall below the predefined threshold of (0 or 3). The immediate implementation includes distribution and administration of antivirals, which is a highly optimistic best-case scenario. In the simulation, if the number of cases rises above the implementation threshold, then mitigation strategies are reapplied in additional cycles anytime the number of cases rises above the threshold until there are no cases in the community. When no cases remain in the community the simulation ends.

Inputs for this study were selected to correspond with previous published studies. Population characteristics and contact network structure were generated from model input parameters and were not obtained from the U.S. Census or other surveys for a specific location. Contact network structure emerges from the model’s network generation processes and is parameterized with distributions of scalar values representing the communities’ demographics, assortativity in the network, and stipulated contact patterns. In this study, the community’s demographic characteristics (i.e., sick at home, babysitting, children per household, teens per household, adults per household, household size, and classroom size; Supplement C) were tested to determine: (1) if differences in demographic characteristics communities results in consistent changes in the recommended mitigation strategies during an influenza pandemic; and, (2) whether changes in demographics have consistent or equal effects across all pandemic influenza mitigation strategies. These parameters were selected because of their association with demographics in multiple previous studies [20,24,25,26]. Specifically, the sick at home parameter is used as a proxy for liberal leave polices in schools and workplace; babysitting was used as proxy for workplace policies that enable parents to stay at home with sick children; family size and age structure varies across communities and family’s with larger sizes have higher incidence of infectious diseases; and, classroom size can be inferred from social economic status.

*Model output analysis*

All output analyses were performed with R version 2.13.1 (R Foundation for Statistical Computing, Vienna, Austria). Logistic regression and random forests were performed to assess and rank pandemic influenza mitigation strategies within a simulated population to determine if differences in population demographics affected the effectiveness of multiple pandemic influenza mitigation strategies. Logistic regression and random forest statistical methods are effective at classifying pandemic influenza mitigation strategies while reducing bias and over-fitting. The random forest statistical analysis was chosen for this particular analysis for its high predictive accuracy for many types of data, unique permutation based variable importance ranking methods [27]. Recent work in computational biology has seen an increased use of random forest, owing to its unique advantage in dealing with small sample size, high- dimensional feature space, and complex data structures [28].

The primary concern when these statistical models were fitted was determining a penalty term *λ* for the regression coefficients to penalize the residual sum of squares, and the choice for a given *λ* was performed using cross-validation methods. To address these concerns, the least absolute shrinkage and selection operator (LASSO) and elastic net regression models were used to prevent over-fitting. These methods are a shrinkage method to retain a subset of the predictors, discarding predictors that are attributing redundant information in the predictive power of the logistic regression. In additional to logistic regression, an ensemble learning method random forest, for classification and ranking, of the different pandemic influenza mitigation strategies, was used.

*Validation: reliability of classifying pandemic influenza mitigation strategies*

To assess the accuracy and credibility of the statistical model’s ability to correctly categorize different pandemic influenza mitigation strategies, based on changes in the population demographics, the original data set had to be split into training and test sets. Specifically, test and training sets were created by splitting the original data by taking into account the levels of the response variable by balancing the class distributions within the splits. Then, the predictive power of the statistical model was compared using both training and test sets from the original data set comparing both logistic regression and random forest results to validate the models.

*Logistic regression*

Of the original community, 150 observations were randomly selected and were split into a training and test set (80% and 20%, respectively). The predictor variables used were: stay at home, adults, teens, children, and babysitting. The responses used were the pandemic influenza mitigation strategies (school closure, child social distancing, adult social distancing, household quarantine, antiviral treatment, household member antiviral prophylaxis, adult *and* child social distancing, school closure *and* adult social distancing, and school closure *and* child social distancing). All main effects and two-way effects for the predictor variables were used to determine an appropriate penalty term (*λ*) for the model. The predictor variables were standardized to uniform their contribution to the analysis, this allowed for a balanced interpretation of the model coefficients in terms of their importance rank. Fig 1 illustrates scatter plots of before and after standardized transformations of the babysitting factor for both training and test set. A leave-one-out cross validation approach was used on 122 sub-samples of the training set to determine the penalty term *λ* for each model. For the LASSO model a *λ* = 0.01 was selected and for the elastic net regularization method a *λ* = 0.02 at a *α* = 0.5 statistical significance level where chosen based on the correct classification of the different pandemic influenza mitigation strategies (see Supplement D for the data dictionary).

Fig 2 trace plot depicts nonzero model coefficients as a function of the regularization parameter *log*(*λ*) for the specific response level *P* > *T*. Because there are 15 predictors (5 main and 10 interaction effects) and a linear model, there are 15 curves represented in the graph. The left panel is the LASSO path while the right is for the elastic net path with *α* = 0.5. The extreme left of each path indicates that there are more non-zero coefficients for elastic with a much smaller magnitude than the LASSO path. This is due to the averaging effect of the elastic net penalty, which can be thought of as a combination of ridge and lasso penalties [29]:

A leave-one-out cross-validation method was used to examine the ability of the model to correctly classify the correct pandemic influenza mitigation strategies from the given data and subset selection. A random test set was created using 70% of the original data balancing for the class distributions of the response variable for the test and training sets. Fig 2 portrays the regularization paths for both the LASSO and elastic net methods for the response P>T and shows how the coefficients of each input variable changes as λ changes.

*Variable importance ranking*

Variable importance plots are used as method to assess the predictive accuracy for each variable in the input space using a permutation based approach. A random forest uses a portion of the training set for the variable subset selection process. The remaining observations are termed as out-of-bag samples. These out-of-bag (OOB) samples can then be used to validate classification accuracy for the tree. For a given tree in the forest, the OOB samples are used to record the predictive accuracy. Then, random permutations for a variable in the input space from the OOB samples are then assessed for predictive accuracy. The decrease in predictive accuracy is recorded for all trees and averaged, which become our mean decrease in accuracy. The idea is that if the variable is not important for classification, then changing the value of that variable will not degrade the predictive power of the model. The permutation method is robust in that it takes into account both the individual and interactions for each variable in computing importance. Random forests have a distinct advantage over logistic regression modeling approaches in that they can cope with complex interactions and highly correlated variables. The mean decrease accuracy is mathematically represented by calculating the permutation based variance importance (PBVI) for variable *Xi* as the following [30].

Where  is the error rate of the variable ,  is the error

rate on the permuted variable , and 𝑡 is an individual tree in the forest of 𝑇 trees.

*K-fold variable importance (KFVI) ranking*

Traditional permutation based methods for assessing variable importance give random forests a distinct advantage over other competing methods. These variable importance methods have shown to be biased towards correlated predictor variables, leading to conditional variable importance ranking methods [31]. In addition to conducting the traditional permutation method for assessing variable importance, we purpose a new approach called k-fold variable importance (KFVI) ranking for ranking correlated predictor variables. This approach compares the predictive power of the full model and a subset of the input space to rank variable importance. An added bonus to conducting the variable importance ranking in this manner allows us to quantify the loss in prediction for each input variable (including its interaction effect) as it applies to each category in the output space. This creates a detailed description of variable importance ranking by producing ranking for each category in the output space as opposed to the traditional permutation method. Additionally, further variable importance sampling can potentially be conducted on groups of input variables using the KFVI method for separating the loss of predictive power for main and interaction effects for each pandemic influenza mitigation strategy.

# Results

*Logistic regression (GLM)*

Table 1 illustrates predictions for the fitted models based on the following penalty terms. The LASSO model was able to classify the four different pandemic influenza mitigation strategies with a mean accuracy of 74% (Table 1). The elastic net logistic regression had a mean accuracy of 79%. The results indicate that the reduced model parameters in the LASSO model have a relatively small effect on the predictive power of the model but the elastic net was consistently slightly better than the LASSO algorithm for all models tested. The predictive power for successfully selecting the CS>SA response was consistently the lowest among all four categories throughout the analysis.

**Table 1.** Predictions for LASSO and elastic net fitted models.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| **LASSO** | **CS>P** | **CS>S** | **P>CS** | **P>T** |
| CS>P | 0.75 | 0.25 | 0 | 0 |
| CS>SA | 0.4 | 0.6 | 0 | 0 |
| P>CS | 0 | 0 | 0.8 | 0.2 |
| P>T | 0 | 0 | 0.2 | 0.8 |
| **Elastic Net** |  |  |  |  |
| CS>P | 0.7 | 0.2 | 0.1 | 0 |
| CS>SA | 0.25 | 0.75 | 0 | 0 |
| P>CS | 0 | 0 | 0.8 | 0.2 |
| P>T | 0 | 0 | 0.11 | 0.89 |

Each model was fit using a multinomial model where each class is represented by a linear model. The interpretation of the importance of the variables in LASSO models is difficult since the algorithm tends to select one variable from a group of highly correlated variables and ignores the others. For instance, if two predictors are perfectly collinear then the shrinkage method will select one predictor at random to retain in the model. The elastic net algorithm is less susceptible to dropping correlated variables than LASSO and in general is more reliable in its ability to interpret the importance of coefficients retained in the model. The coefficients for the LASSO and elastic net models can be ranked by importance due to the standardized methods employed. The averaging effect of the elastic net penalty generally retains more non-zero coefficients than the LASSO algorithm and results using the elastic net penalty term are presented below (Table 2).

**Table 2.** Coefficients for multinomial logistic regression model (elastic net).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameters** | **CS>P** | **CS>SA** | **P>CS** | **P>T** |
|  |  |  |  |  |
| (Intercept) | 0.54 | -1.16 | 0.7 | -0.08 |
|  |  |  |  |  |
| Adults | -0.02 | 0 | 0 | 0.28 |
| Adults & Children | 0 | -0.16 | 0 | 0 |
| Adults & Teens | -0.45 | -0.08 | 0.08 | 0.45 |
| Babysitting | 0.23 | 0 | 0 | -0.26 |
| Babysitting & Adults | 0 | -0.16 | 0.1 | 0 |
| Babysitting & Children | 0 | 0 | 0.07 | 0 |
| Babysitting & Teens | 0.06 | 0 | 0 | 0 |
| Children | -0.28 | -0.09 | 0.09 | 0.28 |
| Stay Home | 0.17 | 1.16 | -0.17 | -1.16 |
| Stay Home & Adults | 0 | 0.24 | 0 | -0.39 |
| Stay Home & Babysitting | 0 | 0.87 | -0.14 | -0.68 |
| Stay Home & Children | 0.18 | 0 | 0 | -0.44 |
| Stay Home Teens | 0.55 | 0.17 | 0 | -0.72 |
| Teens | -0.13 | -0.06 | 0 | 0.59 |
| Teens & Children | -0.12 | -0.86 | 0.43 | 0.54 |

*Permutation based variable importance ranking*

Fig 3 illustrates the mean decrease in accuracy for each of the six factors calculated using permutation based variable importance. The variables are ordered

top-to-bottom (y-axis) from most to least important in classifying the pandemic mitigation strategies: CS>P, CS>SA, P>CS, and P>T. By ranking the population demographic characteristics as described in equation 2 (above), Stay at Home is ranked the most important variable in classifying the four different combinations of pandemic mitigation strategies, which is consistent with the findings from the logistic regression model.

*K-fold variable importance (KFVI) ranking*

Using the KFVI method described previously in the methods section, the following rankings were obtained (Figs 4 and 5) for classifying the four pandemic mitigation methods: CS>P, CS>SA, P>CS, P>T. Figs 4 and 5 illustrate that the removal of the Stay at Home variable has the most profound effect on classifying all four pandemic mitigation methods, while factors Children and Teens rank second for predicting the four pandemic mitigation methods. While the results are consistent with both the permutation based variable importance for the random forest model and the logistic regression, the KFVI ranking method allows for more detailed quantification of the ranking of the population demographics that attribute to prediction of the different pandemic mitigation strategies.

# Discussion

This purpose of the study was to determine: (1) if differences in demographic characteristics results in consistent changes in the recommended mitigation strategies during an influenza pandemic; and, (2) whether changes in demographics have consistent or equal effects across all pandemic influenza mitigation strategies. With the use of conservative statistical methods this study found that: (1) differences in demographic characteristics result in different ranking of pandemic influenza mitigation strategies; and, (2) changes in demographics have unequal effects across pandemic influenza mitigation strategies. As indicated by the logistic regression model, the cost effective (but politically difficult) pandemic influenza mitigation strategy combination of *school closures and child social distancing* (CS) and *school closures and adult social distancing* (SA) had the least amount of predictive power. In contrast, the costly and logistically difficult pandemic influenza mitigation strategy combination of *household member antiviral prophylaxis* (P) and *antiviral treatment* (T) had the most predictive power.

The multinomial regression model, permutation based variable importance ranking, and KFVI methods all had similar results. These methods all indicated that changes in demographics resulted in differing effectiveness of pandemic influenza mitigation strategies and that changes in the community’s demographics resulted in unequal shifts across the effectiveness of all the available influenza mitigation strategies, as illustrated in Table 2. Fig 3 indicates that three of the five most important variables are related to the community’s age structure and two of the five most important variables are related to workplace leave policies. Figs 4 and 5 show the net predictive loss of each influenza mitigation strategy when different demographic characteristics are removed from the model. Under every pandemic influenza mitigation strategy used in the model, the rank order of influenza mitigation strategies changes and the degree of the predictive loss for each demographic characteristic fluctuates. Simply, as the demographics of the community changes, so does the rank order of the most effective pandemic influenza strategies.

Instead of using a one-size fits all approach to mitigating pandemic influenza, the findings of this study imply that pandemic influenza mitigation strategies should be tailored to match the local demographics of the community. Similar to this study, Yang et al. [32] found that children and teens play a critical roll in mitigation of the A H1N1 strain of influenza, likely due to the high contact rate between cases and susceptible individuals. If vaccination is highly effective, then rapid dissemination of the vaccine is imperative to establishing herd immunity before the disease spreads widely through the population [33]. Brown et al. [34] opined that tailored combinations of influenza mitigation policies should be applied to specific age cohorts. Not only should mitigation strategies be tailored to age groups, mitigation strategies should be tailored to easily obtainable local demographics data. As Moss et al. [35] first discussed, the challenge in implementing real and practical pandemic influenza mitigation strategies is enabling local communities to act creatively and independently of the larger public health system. While using local demographics to tailor pandemic influenza mitigation strategy at the national level is likely untenable, local communities typically have less difficulty implementing potentially controversial policy and should be able to apply specific mitigation strategies to best meet the needs of their people. Regardless of which pandemic influenza mitigation strategies are used and the models use to test their likely performance, early detection of influenza outbreaks is imperative and global biosurveillance systems must be improved [36].

This study had several limitations. In this study a simulated population was used to test whether or not demographics matter in the selection of appropriate mitigation strategies. Optimally, actual demographic data and attack rates could be obtained from communities and then could be used to more accurately simulate and predict the effectiveness of pandemic influenza mitigation strategies. The 90% compliance rate for agents in the models was likely much higher than most community compliance rates, and the model assumes that the community does not have any outside contact with neighboring communities.

Future research and agent-based models should focus on incorporating more complexity into the models. Specifically, incorporating actual community demographics data will increase the validity of models. Geospatial models that incorporate the exchange of infected and susceptible people between neighboring communities can improve the predictive power. Lastly, detailed reference data on age specific attack rates from serologic tests and data collected on the specific application of pandemic influenza mitigation strategies to the community would enable modelers to better calibrate basic reproductive rates while providing the means to externally validate pandemic influenza models.

**References**  
  
1. Andradóttir S, Chiu W, Goldsman D, Lee ML, Tsui KL, Sander B, et al. Reactive strategies for containing developing outbreaks of pandemic influenza. BMC Public Health. 2011;11(Suppl 1):S1. doi:10.1186/1471-2458-11-S1-S1  
  
2. Osterholm MT, Kelley NS, Manske JM, Ballering KS, Leighton TR, Moore KA. The compelling need for game-changing influenza vaccines. Minneapolis (MN):University of Minnesota Center for Infectious Disease Research and Policy; 2012 Oct. Sponsored in part by the Alfred P. Sloan Foundation.  
  
3. Osterholm MT, Kelley NS, Sommer A, Belongia EA. Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis. Lancet Infect Dis. 2012;12(1):36-44. doi: <http://dx.doi.org/10.1016/S1473-3099(11)70295-X>  
  
4. Barrett C, Bisset K, Leidig J, Marathe A, Marathe M. Economic and social impact of influenza mitigation strategies by demographic class. Epidemics. 2011;3(1):19-31. doi: [10.1016/j.epidem.2010.11.002](http://dx.doi.org/10.1016%2Fj.epidem.2010.11.002)  
  
5. Germann TC, Kadau K, Longini IM, Macken CA. Mitigation strategies for pandemic influenza in the United States. Proceedings of the National Academy of Sciences. 2006;103(15):5935-5940.  
   
6. Halder N, Kelso JK, Milne GJ. Cost-effective strategies for mitigating a future influenza pandemic with H1N1 2009 characteristics. PLoS One. 2011;6(7):1-10. doi: 10.1371/journal.pone.0022087  
  
7. Matrajt L, Halloran ME, Longini Jr IM. Optimal Vaccine Allocation for the Early Mitigation of Pandemic Influenza. PLoS Comput Biol. 2013;9(3):e10029. doi: 10.1371/journal.pcbi.1002964  
  
8. Perlroth DJ, Glass RJ, Davey VJ, Cannon D, Garber AM, Owens DK. Health outcomes and costs of community mitigation strategies for an influenza pandemic in the United States. Clin Infect Dis. 2010;50(2):165- 174. doi: 10.1086/649867  
  
9. Wu JT, Cowling BJ, Lau EH, Ip DK, Ho LM, Tsang T, et al. School closure and mitigation of pandemic (H1N1) 2009, Hong Kong. Emerg Infect Dis. 2010;16(3):538.  
  
10. Black AJ, House T, Keeling MJ, Ross JV. Epidemiological consequences of household-based antiviral prophylaxis for pandemic influenza. J R Soc Interface. 2013;10(81):1-10. doi: 10.1098/rsif.2012.1019.  
  
11. Kumar S, Grefenstette JJ, Galloway D, Albert SM, Burke DS. Policies to Reduce Influenza in the Workplace: Impact Assessments Using an Agent-Based Model. Am J Public Health. 2013;103:1406-1411. doi: 10.2105/AJPH.2013.301269  
12. Fumanelli L, Ajelli M, Manfredi P, Vespignani A, Merler S. Inferring the structure of social contacts from demographic data in the analysis of infectious diseases spread. PLoS Comput Biol. 2012;8(9):1-10. doi: 10.1371/journal.pcbi.1002673  
  
13. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. Science. 2009;324(5934):1557-1561. doi: 10.1126/science.1176062  
  
14. Check, H. E. The turbulent history of the A(H1N1) virus. Nature. 2009;459(7243):14– 15. doi:10.1038/459014a  
  
15. Araz OM, Galvani A, Meyers LA. Geographic prioritization of distributing pandemic influenza vaccines. Health Care Manag Sci. 2012;15(3):175-187. doi: 10.1007/s10729-012-9199-6  
  
16. Ferguson NM, Cummings DA, Fraser C, Cajka JC, Cooley PC, Burke, DS. Strategies for mitigating an influenza pandemic. Nature. 2006;442(7101):448-452. doi:10.1038/nature04795  
  
17. Glezen PW, Couch RB, MacLean RA, Payne A, Baird J, Vallbona C, et al. Interpandemic Influenza in the Houston Area. N Engl J Med. 1978 Mar;298:587-592.  
  
18. Glass RJ, Glass LM, Beyeler WE, Min HJ. Targeted social distancing design for pandemic influenza. Emerg Infect Dis. 2006;12(11):1671-1681.  
  
19. Philip RN, Bell JA, Davis DJ, Beem MO, Beigelman PM, Engler JL, et al. Epidemiologic studies on influenza in familial and general population groups, 1951–1956, characteristics of occurrence. Am J Hyg. 1961;73(2):123-137.  
  
20. Taber LH, Paredes A, Glezen WP, Couch RB. Infection with influenza A/Victoria virus in Houston families, 1976. J Hyg (Lond). 1981;86(3):303-313.  
  
21. Glass RJ, Glass LM, Beyeler WE. Local mitigation strategies for pandemic influenza. . Sandia National Laboratories Report. Alburquerque (NM):Department of Homeland Security by the National Infrastructure Simulation and Analysis Center; 2005. 2005;SAND2005-7955J  
  
22. Kumar S, Quinn SC, Kim KH, Daniel LH, Freimuth VS. The impact of workplace policies and other social factors on self-reported influenza-like illness incidence during the 2009 H1N1 pandemic. Am J Public Health. 2012;102(1):134-140. doi: 10.2105/AJPH.2011.300307  
  
23. Davey VJ, Glass RJ, Min HJ, Beyeler WE, Glass LM. Effective, robust design of community mitigation for pandemic influenza: a systematic examination of proposed US guidance. PLoS One. 2008;3(7):e2606. doi: 10.1371/journal.pone.0002606  
  
24. Ehrenberg RG, Brewer DJ, Gamoran A, Willms  JD. Class size and student achievement. Psychol Sci Public Interest. 2001;2(1):1-30.  
  
25. Rousculp MD, Johnston SS, Palmer LA, Chu BC, Mahadevia PJ, Nichol KL. Attending work while sick: Implication of flexible sick leave policies. J Occup Environ Med. 2010;52(10):1009- 1013. doi: 10.1097/JOM.0b013e3181f43844.  
  
26. Stroud P, Del Valle S, Sydoriak S, Riese J, Mniszewski S. Spatial dynamics of pandemic influenza in a massive artificial society. J Artif Soc Soc Simul. 2007;10(4):9.

27. Breiman L. Random forests. Mach Learn. 2001;45(1):5-32. doi: 10.1023/A:1010933404324  
  
28. Chen X, Ishwaran H. Random forests for genomic data analysis. Genomics. 2012;99(6):323-329. doi:10.1016/j.ygeno.2012.04.003  
  
29. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 1st ed. New York: Springer; 2001.  
  
30. Breiman L,  Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. 1st ed. Monterey: Wadsworth & Brooks; 1984.  
  
31. Strobl C,  Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. BMC Bioinformatics. 2008;9(1):307. doi:10.1186/1471-2105-9-307  
  
32. Yang Y, Sugimoto JD, Halloran ME, Basta NE, Chao DL, Matrajt L, et al. The transmissibility and control of pandemic influenza A (H1N1) virus. Science. 2009;326(5953):729-733. doi: 10.1126/science.1177373  
  
33. Gojovic MZ, Sander B, Fisman D, Krahn MD, Bauch, CT. Modelling mitigation strategies for pandemic (H1N1) 2009. CMAJ. 2009;181(10):673-680. doi: 10.1503/cmaj.091641  
  
34. Brown ST, Tai JH, Bailey RR, Cooley PC, Wheaton WD, Potter MA, et al. Would school closure for the 2009 H1N1 influenza epidemic have been worth the cost?: a computational simulation of Pennsylvania. BMC Public Health. 2011;11(1):353. doi: [10.1186/1471-2458-11-353](http://dx.doi.org/10.1186%2F1471-2458-11-353)  
  
35. Moss R, McCaw JM, McVernon J. Diagnosis and antiviral intervention strategies for mitigating an influenza epidemic. PLoS One. 2011;6(2):e14505. doi: 10.1371/journal.pone.0014505  
  
36. Lipsitch M, Riley S, Cauchemez S, Ghani AC, Ferguson NM. Managing and reducing uncertainty in an emerging influenza pandemic. N Engl J Med. 2009;361(2):112-115. doi: 10.1056/NEJMp0904380.