Original Article   **Open Access**

A Disease Outbreak Prediction Model Using Bayesian Inference: A Case of Influenza

Atefeh Sadat Mirarabshahi[1], Mehrdad Kargari[1*]

[1]Information Technology Department, Faculty of Industrial Engineering, Tarbiat Modares University, Tehran, Iran

Corresponding Author: Mehrdad Kargari, PhD, Assistant Professor, Information Technology Department, Faculty of Industrial Engineering, Tarbiat Modares University, Tehran, Iran. Tel/Fax: +98-2182884955, Email: m_kargari@modares.ac.ir

Abstract

Introduction: One major problem in analyzing epidemic data is the lack of data and high dependency among the available data, which is due to the fact that the epidemic process is not directly observable.

Methods: One method for epidemic data analysis to estimate the desired epidemic parameters, such as disease transmission rate and recovery rate, is data intensification. In this method, unknown quantities are considered as additional parameters of the model and are extracted using other parameters. The Markov Chain Monte Carlo algorithm is extensively used in this field.

Results: The current study presents a Bayesian statistical analysis of influenza outbreak data using Markov Chain Monte Carlo data intensification that is independent of probability approximation and provides a wider range of results than previous studies. A method for estimating the epidemic parameters has been presented in a way that the problem of uncertainty regarding the modeling of dynamic biological systems can be solved. The proposed method is then applied to fit an SIR-like flu transmission model to data from 19 years leading up to the seventh week of the 2017 incidence of influenza.

Conclusion: The proposed method showed an improvement in estimating the values of all the parameters considered in the study. The results of this study showed that the distributions are significant and the error ranges are real.

Keywords: Disease Outbreak, Metropolis-Hastings Algorithm, Influenza

Introduction

Understanding, modeling, and predicting epidemic outbreaks of infectious diseases bear great importance. Not only will the costs of vaccination and treatment be reduced, but also this may be the only way to control incurable diseases such as AIDS. Because of high disease transmission rates and their complications, infectious diseases such as influenza add to the need to investigate the spread of contagious diseases. Although the mechanism of epidemiological transmission of infectious diseases, such as fever, AIDS, and influenza, is complex, a better understanding of the problem can be achieved by using mathematical modeling.

Research in the epidemiology field has traditionally employed mathematical models to successfully reproduce the observed incidence and prevalence of diseases,[1] including influenza,[2] HIV,[3] smallpox,[4] and malaria[5] among others. These models are important for both improving our understanding of potentially novel disease strains, e.g., A/H1N1,[6] and

developing the actions for infectious disease outbreaks.

Although in 1760, Daniel Bernoulli formulated and solved a model for smallpox to evaluate the effectiveness of variolation of healthy people with the smallpox virus,[7] deterministic epidemiology modeling seems to actually have started in the 20th century. In 1906, Dietz formulated and analyzed a discrete time model in an attempt to understand the recurrence of measles epidemics.[8,9] Ross was interested in the incidence and control of malaria, so in 1911, he developed differential equation models for it as a host-vector disease.[10,11] Starting in 1975 and 1967, Bailey and Dietz,[12,13] published papers on epidemic models and obtained the epidemic threshold result that the density of susceptibility must exceed a critical value in order for an epidemic outbreak to occur. Mathematical epidemiology seems to have grown exponentially starting in the middle of the 20th century (the first edition in 1957 of Bailey's book[14] is an important landmark), so that a tremendous variety of models has now

been formulated, mathematically analyzed, and applied to infectious diseases. Literature reviews[15-18] show the rapid growth of epidemiology modeling. Recent models have considered aspects such as passive immunity,[19,20] gradual loss of vaccine and disease-acquired immunity, stages of infection,[21] vertical transmission, disease vectors, macro-parasitic loads, age structure, social and sexual mixing groups, spatial spread, vaccination, quarantine, and chemotherapy. Special models have been formulated for diseases such as measles, rubella, chickenpox, whooping cough, diphtheria, smallpox, malaria, onchocerciasis, filariasis, rabies, gonorrhea, herpes, syphilis, and HIV/AIDS.

Compartmental Models

Compartmental models estimate the number of people in different disease phases.[22] SIR, SIS, SIRS, SEIR, and MSIR are prevalent compartment models in which S, E, I, R, and M, respectively, represent the population susceptible to infection, exposed to infection, infected with the virus, recovered from infection, and having passive immunity.

The decision to select and include compartments in a model depends on the characteristics of the particular disease being modeled and the purpose of the model. The passively immune class M and the latent period class E are often omitted, because they are not crucial for the susceptible-infective interaction.

The threshold for many epidemiology models is the basic reproduction number R_0, defined as the number of cases generated on average over the course of the infectious period of a disease in an otherwise uninfected population.

For many deterministic epidemiology models, an infection can get started in a fully susceptible population if and only if $R_0 > 1$. Thus, the basic reproduction number R_0 is often considered as the threshold quantity that determines when an infection can invade and persist in a new host population.[20]

Bayesian Inference

Bayesian inference, in its basic mathematical meaning, starts with a global probability distribution for all relevant variables, observing the values of some of these variables, and quoting the conditional distribution of the remaining variables given the observations.[23] Bayesian methods are criticized for the necessity of the priors to be arbitrary or subjective in a pernicious or special way. In observational studies, however, the priors need be no more arbitrary than the largely arbitrary data models routinely slapped on data, and prior models can often be given a scientific foundation that is as form or more so than that of frequentist data models. Like any analysis element, prior models should be scrutinized critically (and rejected as warranted), just like frequentist models. When prior frequency data is absent or invalid, however, other sources of priors will enter, and must be judged critically. In the following, how normal log relative-risk priors can be transformed into 'informationally equivalent' prior frequency data will be shown. Elsewhere, how this transformation extends to non-normal priors and to other parameters will be shown.[24] It will be argued that this transformation should become a standard method for evaluating whether a prior is contextually reasonable, even if the translation is not used to compute posteriors.[25]

Markov Chain Monte Carlo Algorithms

A sequence of random variables X_1, X_2, … from a set is considered as a Markov chain if moving to the next state depends only on the present state and not on the previous states, meaning moving to X_{n+1} in a set of X_1, …, X_n depends only on X_n. The set in which the X_i take values is called the state space of the Markov chain.

A Markov chain has stationary transition probabilities if the conditional distribution of X_{n+1} given X_n does not depend on n. This is the main kind of Markov chain of interest in Markov Chain Monte Carlo (MCMC). The joint distribution of a Markov chain is determined by the marginal distribution of X_1, called the initial distribution, and the conditional distribution of X_{n+1} given X_n, called the transition probability distribution (Because of the assumption of stationary transition probabilities, this does not depend on n).

People introduced to Markov chains through a typical course on stochastic processes have usually only seen examples where the state space is finite or countable.[26] If the state space { X_1, …, X_n}, then the initial distribution can be associated with a vector $\lambda = (\lambda_1, …, \lambda_n)$ defined by equation (1):

$$P(X_1 = x_i) = \lambda_i, i = 1, …, n \ (1)$$

and the transition probabilities can be associated with a matrix P having elements p_{ij} defined by:

$$(X_{n+1} = x\_j \,|X_n = x_i) = p_{ij} \ i=1,…,n \ and \ j=1,…,n \ (2)$$

When the state space is countably infinite, we can think of an infinite vector and matrix. Most Markov chains of interest in MCMC, however, have uncountable state space, and thus, we cannot think of the initial distribution as a vector or the transition probability distribution as a matrix. They must be considered as an unconditional and a conditional probability distribution, respectively.

Methods
Study Population

Infectious disease surveillance systems are powerful tools for monitoring and understanding infectious disease dynamics; however, underreporting (due to both unreported and asymptomatic infections) and observation errors in these systems create challenges for delineating a complete picture of infectious disease epidemiology.

The Center for Disease Prevention and Control in the United States publishes positive flu test reports from clinical laboratories and public health labs on a weekly basis on its website, cdc.gov. This data includes the reported number of Type-A flu (including H3N2v, H1N1 and H3) and type-B flu (including BVIC and BYAM) cases per week. The data belonging to the 19 years leading up to the 33rd week of 2017 is accessible from the CDC website. The research population used included the number of type-A flu cases reported from the year 1997 to August 19th, 2017.

The Stochastic SIR Model

The proposed model in this paper is a stochastic SIR model in which the state of the population at time t is defined through a vector with non-negative integer elements s(t) = (S(t), I(t)) denoting the number of individuals in susceptible and infective states. The model is specified using the parameter vector $v = (S_0, a, R_e, m)$.

If each transition from one state to another is defined as an event, this model includes a total of 2 event types. The first is when a susceptible individual becomes infected and immediately then infective with the rate of λ; the second is when an infective is removed from the population with the rate of τ. These 2 types of event are shown with S(0) and I(0), in which S(0) is an unknown parameter and is to be simulated, but the I(0) values are known and I(0)>0. A limit for the occurrence frequency of each transition is considered. This frequency is shown as ω_j and $1 \leq \omega_j \leq \omega_{19}$, and $t_j \in (0,19)$ shows the time of each transition. First, vector v is selected uniformly from V; then s is simulated from the model with the parameters of v. The joint density of (v,s) is defined by
$$\beta(v,s)dvds \propto L(v,s)dvdt_1 \ldots dt_{\omega T} \tag{3}$$

Where

$$L(v,s) = exp\left(-(19 - t_{\omega_{19}}) \sum_{n \in R} \eta_n\left(v, s(t_{\omega_{19}})\right)\right)$$
$$\times \prod_{j=1}^{\omega_{19}} \eta_{\omega_j}\left(v, s(t_j)\right) exp\left(-(t_j - t_{j-1}) \sum_{n \in R} \eta_n\left(v, s(t_j)\right)\right)$$

Parameters of the model
Variable: Definition
S_0: Initial number of people susceptible to the disease
A: Fraction of infected people who are infectious
R_e: Effective reproduction number
M: migration rate

The parameter-estimation approach used in this study considers a density $\pi(v,s)$ which is proportional to $\beta(v,s)$. Thus, $\pi(v,s)$ represents the posterior joint density of v and s conditional upon the observations N(t), assuming a uniform prior for v.

Since no information is available on the output of the model, uninformative priors are used as the parameters.
U(0,1)
A search is performed for the parameters' feasible space through the Markov Chain Monte Carlo. Error distribution, φ_i, $i \in \{1, \ldots, K\}$, is assumed to be of normal distribution with the following attributes:
$N(\mu = 0.002, 0.0005^2)$

Thus, $L(\varphi_i)$ will be a normal probability function with fixed variance of σ^2.

Construction of the MCMC Sampler

Constructing a Markov chain for problems is not unique; here, we propose a Metropolis-Hastings algorithm in which a new state is generated based on the current state. To construct our Markov chain, the new state $\omega_{i+1} = (\alpha_{i+1}, s_{i+1})$ is obtained from (α_i, s_i) through the following steps:
α_{i+1} is selected from the density distribution (5).

$$\pi(\alpha|s_i) = \frac{L(\alpha, s_i)}{\sum_{b \in A} L(b, s_i)} \tag{5}$$

The realization s_{i+1} is obtained by applying the realization updating procedure M times to s_i to generate a sequence of states $\sigma_{i0}, \sigma_{i1}, \ldots, \sigma_{iM}$, where $s_i = \sigma_{i0}$ and $s_{i+1} = \sigma_{iM}$. Each σ_{ij} is obtained from $\sigma_{i(j-1)}$ by inserting, deleting, or translating an event of type I' in $\sigma_{i(j-1)}$ to produce a candidate state σ'. Mutations of each type are proposed with equal probability ($p1 = p2 = 1/3$). Depending on the kind of mutation proposed, the state σ' is accepted ($\sigma_{ij} = \sigma'$) with α_i replaced by α_{i+1}; α_i has already been updated in Step 1; otherwise, the change is rejected ($\sigma_{ij} = \sigma_{i(j-1)}$). M is selected to be large enough (5000 in this paper) to explore more space in the s-direction than in the α-direction.

Multi-seasonal Influenza Spread

In this paper, an SIR transmission model is used to describe the transmission of influenza. The major difference between the proposed model and the standard SIR model is that despite the SIR model, a fraction of infected people are infectious at the same time; others are undefeated and show no symptom of the virus. A migration rate is also assumed in the model. The standard SIR model, which was introduced by Kermack, is shown through relations (6) to (9).

$$\frac{dS}{dt} = -\beta IS \tag{6}$$
$$\frac{dI}{dt} = \beta IS - \alpha I \tag{7}$$
$$\frac{dR}{dt} = \alpha I \tag{8}$$

Thus, by considering the fraction α and the migration rate, the proposed model would be defined as relations (9) to (12):

$$\lambda = \beta(\alpha I + m) \tag{9}$$
$$\frac{dS}{dt} = -\lambda S \tag{10}$$
$$\frac{dI}{dt} = \lambda S - \tau I \tag{11}$$
$$\frac{dR}{dt} = \tau I \tag{12}$$

The variables and parameters used in the model above are defined as:
Variables used
S: Number of susceptible people
α: Fraction of infected people who are infectious
m: Migration rate
Parameters used
I: Number of infected people
R: Number of recovered people
λ: Influenza transmission rate in proposed model
β: Influenza transmission rate in SIR model
τ: Recovery rate
t: Time

In order to model the seasonal data of influenza epidemics in the United States, it is assumed that the transmission rate β decreases during summer holidays resulting in a reduction in the transmission rate. For the rest of the year, β is assumed to be large enough to keep the transmission stable. During this period, the effective reproduction number, R_e, is defined according to relation (13).

$$R_e = \frac{\beta\alpha}{\tau} S_0 \tag{13}$$

In this relation, S_0 is the number of susceptibles at the start of each transmission season.

In the proposed model, the total population is assumed to be constant and is the summation of susceptible, infected, and recovered people. S_0 is estimated as one of the model outputs, and I_0 is the number of infected people in the first week of the data series used. The remaining population belongs to R.

There is a lack of information on the history of the epidemic. Thus, to begin the simulation, arbitrary random distributions are assigned to the parameters of interest. The uniform distributions assigned to each model variable are:

$S_0 \sim U(0,1)$
$\alpha \sim U(0,0.4)$
$R_e \sim U(1,1.4)$
$m \sim U(0, 4e-6)$

The posterior distributions of each variable are obtained from 5000 samples using the Metropolis-Hastings algorithm. The steps of this algorithm are discussed in Methods.
The MCMC algorithm is used to generate estimates of the marginal parameter density $\pi_c(v)$ over $v = 0.001, 0.002, 0.1$ and the marginal initial-condition density $\pi_c(652)$. The first of these is estimated by averaging 5000 conditional densities $\pi_c(v|s_c)$ with M = 1000 mutations being proposed between the calculation of each density. The density $\pi_c(652)$ is estimated from a histogram formed from the set of all 5×10^6 realizations, σ_{ij} ($1\leq i\leq 5000$, $1\leq j \leq 1000$), generated in the simulation. Note that the joint density $\pi_c(v, 652)$ can be estimated as a histogram over a two-dimensional grid of cells indexed by $I(0)$ and the current value of v.

Simulation Convergence Study
To investigate the convergence of the Markov Chain Monte Carlo algorithm used for parameter estimation, trace diagrams and the probability density function for each parameter are plotted.

In the trace diagrams, the vertical axis represents the simulated values of the parameters, and the horizontal axis represents the number of repetitions; these simulated values are connected with a line. Due to the scattering and trends in the trace diagrams, one can comment on the convergence of the simulation results.

Results
After collecting the data and developing the model, the proposed algorithm was run and the results were recorded. The results of estimating the proposed model variables and the results from the filter-based simulation method (EAKF/PF)[27] are provided and compared in Table 1. The estimated mean and mode of the unknown parameters were all in the equal-tailed 95% confidence interval, and no significant difference was observed between the proposed model results and the results of Yang.[28] In this study, 5000 MCMC samples were used to estimate the parameters of interest. The estimated value of the basic reproduction number mean was 7.96, which is 1.09 greater than the Yang[28] estimate. Similarly, the estimation of the proposed method from the reproduction number of the symptoms period was 0.53, which was 0.164 greater than the estimate of Yang.[28] The significant difference in this case can be explained by the high skewness of the b parameter posterior density function, which resulted in a significant difference in mean and mode values.

Sensitivity Analysis
The sensitivity analysis of individual-based models of epidemic dynamics were used to evaluate the effect of disease parameters on public policy-related questions.[28]

For individual-based models used in the study of influenza-like epidemics, sensitivity analyses could focus on changes in interventions and response strategies such as in the studies by Nsoesie et al,[29] Halloran et al,[30] Germann et al,[31] and Brooks et al.[32] The current study aimed to explore the sensitivity of an individual-based epidemiology model to changes in the assumptions made regarding the characteristics of the disease. The sensitivity analysis of the results to decide on the behavior of the disease transmission is provided in Table 2.

To evaluate the fitness of the outbreak data with the proposed model, the Posterior Probability distribution was used. The samples from the model's basic parameters were taken as output, and a forward time model simulation was performed for the MCMC of each set of parameters. Then, different aspects of the observed data were compared with the simulation results in order to determine the data alignment of the previous and new results.

Discussion
The current study analyzed the incidence of influenza in the United States between 1997 and 2017. The stochastic

Table 1. Parameter Estimates and Confidence Intervals of 95% Compared With the Posterior of Yang and Lipsitch

Parameter	Posterior Mean	Posterior Mode	Confidence Interval	Yang and Lipsitch Estimates	Confidence Interval of Yang and Lipsitch
S_0	68.5	69.1	(64.6, 72.4)	68.9	(65.0, 71.6)
α	0.0882	0.0903	(0.0637, 0.1127)	0.0792	(0.0547, 0.0861)
R_e	2.97	3.04	(2.77, 3.14)	2.04	(1.84, 2.21)
m	1.78E-0.6	1.98E-0.6	(1.35E-0.7, 2.95E-0.6)	2.84E-0.6	(8.98E-0.7, 3.92E-0.6)

Table 2. Sensitivity Analysis of Model Parameter Estimates

Parameter	Posterior Mean	Lower Limit of Confidence Interval	Upper Limit of Confidence Interval
S_0	-0.4	-0.4	+0.8
α	+0.009	+0.019	+0.0266
R_e	+0.93	+0.84	+0.97
m	-1.06E	-7.63E	-0.97E

transmission model used in this study is similar to the model developed by Yang,[28] but some assumptions of this model are controversial, and there are grounds for more studies for model development.

In Figure 1, the simulation results of the proposed Monte Carlo model are compared with the SIR model. The number of susceptible people is higher in the Monte Carlo simulation model than in the SIR model, because in the proposed model, the spread of disease was prevented by using the information method, thus reducing the number of patients. By reducing the number of patients and raising public awareness about the spread of the disease and ways to prevent it, such as reducing contact with patients, there is less need for a vaccine in the community.

In this research, the main focus was on the original model, and the probabilistic approximation methods used by Yang[28] were avoided. The imputation method used to improve the time series data in this research has the advantages offlexibility, because they are applied in many areas, and the ability to consider and estimate unexpected quantities during an epidemic, such as transmission times.

In addition to estimating the parameters of the model, it is possible to estimate the quantities that are functions of the parameters, such as the propagation numbers and the corresponding uncertainty (In contrast, the maximum likelihood estimation method usually requires asymptomatic results to exceed the estimated points).

To determine the implementation convergence for the Markov Chain Monte Carlo algorithm so as to estimate the parameters, trace diagrams, histogram diagrams, and the probability density function for each parameter are illustrated. In the trace diagrams, the vertical axis represents the simulated parameters, and the horizontal axis represents

the iteration numbers; these simulated values are linked to each other. Based on the scattering and trend of the data in the trace diagrams, a conclusion on the simulation convergence is provided.
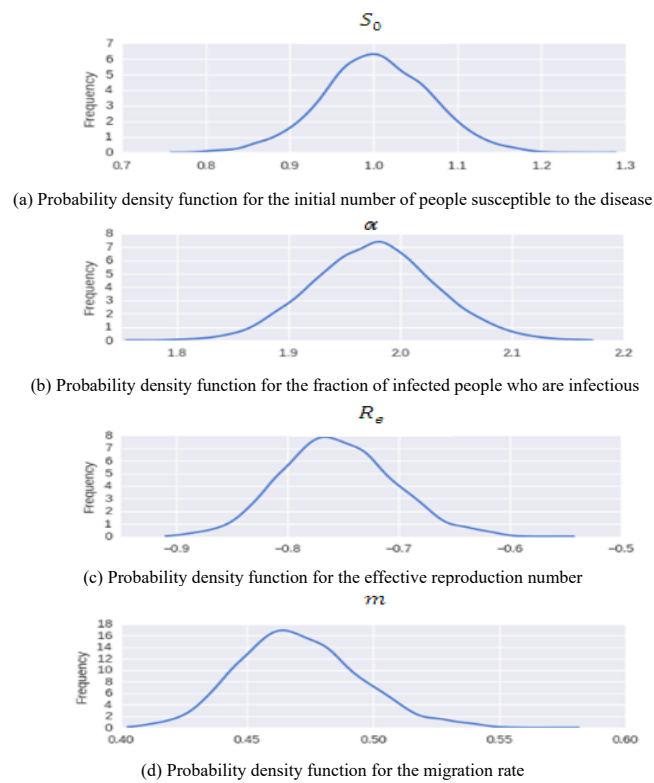
The histogram diagrams show the general shape of the posterior marginal distribution of the model parameter. The use of histograms and kernel density estimates of the posterior marginal distributions of the parameters is necessary to ensure that these empirical distributions have the expected behavior.

Figure 2 shows the scatter diagrams of samples derived by MCMC methods. In these diagrams, the horizontal axis represents the time, and the vertical axis represents the number of cases (diseases) observed. Figure 3 shows the trace diagrams regarding the (a) initial number of people susceptible to the disease, (b) fraction of infected people who are infectious, (c) effective reproduction number, and (d) migration rate based on 5000 MCMC samples. As stated above, the vertical axis represents the simulated values, and the horizontal axis represents the number of repetitions.

Another strategy for disease control is the optimal control strategy in the form of vaccination to control suspected

Figure 1. Comparison of the Number of Susceptible People (the Monte Carlo Simulation Model and the SIR Model).

(a) Sample values in each iteration for the initial number of people susceptible to the disease

(b) Sample values in each iteration for the fraction of infected people who are infectious

(c) Sample values in each iteration for effective reproduction number

(d) Sample values in each iteration for migration rate

Figure 2. Trace Diagrams of Simulated Model Parameters.

S_0

(a) Probability density function for the initial number of people susceptible to the disease

α

(b) Probability density function for the fraction of infected people who are infectious

R_e

(c) Probability density function for the effective reproduction number

m

(d) Probability density function for the migration rate

Figure 3. Posterior Density Functions Estimated From Model Parameters

Figure 4. Comparison of SIR and Monte Carlo Models With and Without Optimal Control.

populations and to increase the population of the recovered people.[33]

The optimal control theory is used, because it is a mature mathematical field with multiple applications in both science and engineering.

For example, in one paper, optimal control strategies of a time-delayed epidemic SIR model are introduced, in which control means the treatment of contagious hosts.[34] Here, optimal control is used to minimize the probability of the expansion of the contagious population and to maximize the total number of suspected and recovered populations. To this end, a control variable, which represents the optimal treatment for contagious hosts is first introduced, and then an optimal control system for the SIR epidemic model is given. Next, the base propagation rate is calculated, and the dynamic behavior of the controlled SIR epidemic model is examined. In addition, an optimal control for this problem has been developed and indicates that infection in a set of cells will be eliminated by this optimal control treatment. It also examines the optimal control and optimization of the system using optimal control techniques. It shows that the base reproductive number is less than one, so the disease is eliminated within the target population by using the control method. In addition, optimal control strategies reduce the number of influenza patients and increase the total number of suspected and improved populations.

As can be seen in Figure 4, the number of vaccinated individuals was reduced using optimal control, which was one of the goals in the current study. Reducing the number of vaccinated people and increasing the number of susceptible people means that there are fewer costs involved in vaccination preparation, because the disease can be prevented by increasing people's awareness; there is no need to vaccinate large numbers of people to prevent the spread of the disease. Moreover, as people's awareness increases, the relationship between susceptible and infected people decreases, resulting in a decrease in the number of patients, both non-carriers and those who are contagious.
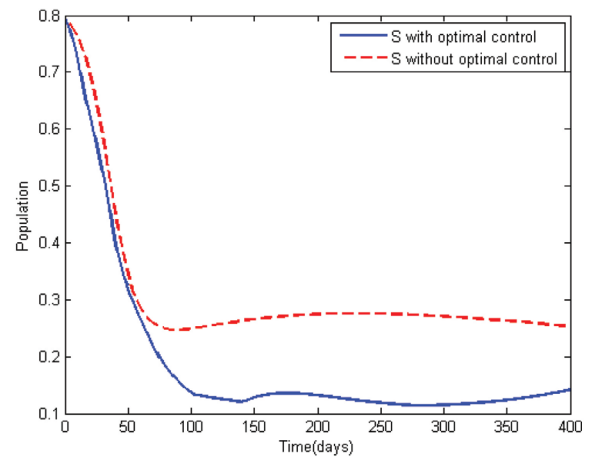
- Ability to estimate communication between parameters.
- Because the analysis performed in the current study was based on a transmission model, it is possible that by simulating the model, the ability to evaluate and examine other scenarios can be modeled.

Conclusion

Estimation of the basic parameters of the model shows that the transmission between the individuals in each group was dominant; this means that the vast majority of transmission events occur within each group, and the disease is spread within groups and households.

Estimated posterior mean R_e is close to 2.97 in this study. The value obtained is greater than the estimate of old models, which was 2.04, reminding us of the devastating nature of the flu. These values are fundamentally different from those obtained from simpler models that emphasized the importance of models that take the demographic structure into account.

According to the assumptions made for the model, control measures reduce the duration of a person's disease symptoms from an average of 7 days to 3 or 4 days, resulting in fewer infected people. If the pre-control and post-control measures for dissemination within larger groups or populations are defined, then estimates from the latter mean that (a) within-group epidemics before and after quarantine were beyond critical; (b) in the vaccinated population, the epidemic state shifts from super-critical to sub-critical, and (c) in larger populations, the epidemic will always be sub-critical. Despite this, the increase in quarantine time in the simulations shows an increase in the size of the outbreak. For example, considering quarantine times of 10, 20, and 40 days results in

approximately 12, 22, and 32 outbreaks on average. However, with no limitations in this study, the average outbreak was approximately 43 days, indicating that the epidemic was sub-critical.

As previously mentioned, the effective reproduction number shows the average number of secondary people defined as suffering from an infected person in a susceptible, sufficiently large population. In the current study, the practice of such an interpretation is directly problematic, because the population of each group, which is dominated by traps within them, is also small enough to create a quick saturation effect by engaging all available susceptible people.

The Markov Chain Data Intensification Methods used in this study have the following advantages:
(A) Flexibility. They have applicability in many areas.
(B) They have the ability to consider and estimate unobserved quantities during epidemics such as disease transmission times.
(C) In addition to estimating model parameters, it is possible to estimate quantities that are a function of parameters such as the corresponding duplication numbers and related uncertainties (In contrast, the maximum likelihood estimation method usually requires asymptotic results to yield results beyond the point estimates).
D) They have the possibility of estimating inter-parametric relationships.

Since the analysis performed in this study is based on a transfer model, it is possible to perform model simulations to evaluate the model and investigate other scenarios. This feature is unlike other methods that use a model based on the observed data.

Limitation
The major limitations we face in modeling issues are access to sufficient and appropriate data to ensure the model's data fitness, particularly in phenomena such as epidemics; we often encounter data malfunction or low data volume. Depending on its type, it is possible to simulate the data, but in cases such as community vaccination or prevention status, precise information from relevant institutions must be available if we want to achieve high accuracy of the model.

Authors' Contributions
ASM and MK contributed equally to this research.

Conflict of Interest Disclosures
The authors declare that they have no conflicts of interest.

Ethical Approval
This study was approved by Ethics Committee of Tarbiat Modares University.

References
1. Anderson RM, May RM. Infectious diseases of humans: Dynamics and control. Oxford: Oxford University Press; 1991.
2. Anderson RM. The role of mathematical models in the study of HIV transmission and the epidemiology of AIDS. J Acquir Immune Defic Syndr. 1988;1(3):241-256.
3. Ferguson NM, Keeling MJ, Edmunds WJ, et al. Planning for smallpox outbreaks. Nature. 2003;425(6959):681-685. doi:10.1038/nature02007.
4. Ngwa GA, Shu WS. A mathematical model for endemic malaria with variable human and mosquito populations. Math Comput Model. 2000;32(7-8):747-763. doi:10.1016/S0895-7177(00)00169-2.
5. Fraser C, Donnelly CA, Cauchemez S, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. Science. 2009;324(5934):1557-1561. doi:10.1126/science.1176062.
6. Gibson GJ, Renshaw E. Estimating parameters in stochastic compartmental models using Markov chain methods. IMA J Math Appl Med Biol. 1998;15(1):19-40. doi:10.1093/imammb/15.1.19.
7. Hethcote HW. The mathematics of infectious diseases. SIAM Rev Soc Ind Appl Math. 2000;42(4):599-653. doi:10.1137/s0036144500371907.
8. Dietz K. Transmission and control of arbovirus diseases. In: Ludwig D, Cooke KL, eds. Epidemiology. Philadelphia: SIAM; 1975:104-121.
9. Just W, Callender HL. Differential equation models of disease transmission. March 2015. http://shorturl.at/dBZ36.
10. Hamer W. Epidemic disease in England. Lancet. 1906;(1):733-739.
11. Ross R. The Prevention of Malaria. 2nd ed. London: Murray; 1981.
12. Bailey NT. The Mathematical Theory of Infectious Diseases. 2nd ed. New York: Hafner; 1975.
13. Dietz K. Epidemics and rumours: A survey. J R Stat Soc Ser A. 1967;130(4):505-528. doi:10.2307/2982521.
14. Dietz K. The first epidemic model: a historical note on PD En'ko. Aust J Stat. 1988;30A(1):56-65. doi:10.1111/j.1467-842X.1988.tb00464.x.
15. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc R Soc Lond A. 1927;115(772):700-721. doi:10.1098/rspa.1927.0118.
16. Becker N. The uses of epidemic models. Biometrics. 1979;35(1):295-305. doi:10.2307/2529951.
17. Castillo-Chavez C. Mathematical and statistical approaches to AIDS epidemiology. Berlin: Springer-Verlag; 1989. doi:10.1007/978-3-642-93454-4.
18. Dietz K. Density-dependence in parasite transmission dynamics. Parasitol Today. 1988;4(4):91-97. doi:10.1016/0169-4758(88)90034-8.
19. Dietz K, Schenzle D. Mathematical models for infectious disease statistics In: Atkinson AC, Fienberg SE, eds. A Celebration of Statistics. New York: Springer-Verlag; 1985:167-204.

doi:10.1007/978-1-4613-8560-8_8.

20. Hethcote H.W. A thousand and one epidemic models. In: Levin SA, ed. Frontiers in mathematical biology. New York: Springer-Verlag; 1994:504-515. doi:10.1007/978-3-642-50124-1_29.

21. Hethcote HW, Levin SA. Periodicity in epidemiological models. In: Levin SA, Hallam TG, Gross LJ, eds. Applied mathematical ecology. New York: Springer-Verlag; 1989:193-211. doi:10.1007/978-3-642-61317-3_8.

22. Hethcote HW, Stech P. Periodicity and stability in epidemic models: A survey. In: Busenberg SN, Cooke KL, eds. Differential equations and applications in ecology, epidemics, and population problems. New York: Academic Press; 1981:65-82. doi:10.1016/B978-0-12-148360-9.50011-1.

23. Wickwire K. Mathematical models for the control of pests and infectious diseases: a survey. Theor Popul Biol. 1977;11(2):182-238. doi:10.1016/0040-5809(77)90025-9.

24. Dempster AP. A generalization of Bayesian inference. In: Yager RR, Liu L, eds. Classic works of the Dempster-Shafer theory of belief functions. Berlin: Springer; 2008. doi:10.1007/978-3-540-44792-4_1.

25. Greenland S. Generalized conjugate priors for Bayesian analysis of risk and survival regressions. Biometrics. 2003;59(1):92-99. doi:10.1111/1541-0420.00011.

26. Greenland S. Bayesian perspectives for epidemiological research: I. Foundations and basic methods. Int J Epidemiol. 2006;35(3):765-775. doi:10.1093/ije/dyi312.

27. Brooks S, Gelman A, Jones G, Meng XL, eds. Handbook of Markov Chain Monte Carlo. CRC Press; 2011. doi:10.1201/b10905.

28. Yang B. Stochastic dynamics of an SEIS epidemic model. Adv Differ Equ. 2016;2016(1):226. doi:10.1186/s13662-016-0914-3.

29. Nsoesie EO, Beckman RJ, Marathe MV. Sensitivity analysis of an individual-based model for simulation of influenza epidemics. PLoS One. 2012;7(10):e45414. doi:10.1371/journal.pone.0045414.

30. Halloran ME, Ferguson NM, Eubank S, et al. Modeling targeted layered containment of an influenza pandemic in the United States. Proc Natl Acad Sci U S A. 2008;105(12):4639-4644. doi:10.1073/pnas.0706849105.

31. Germann TC, Kadau K, Longini IM, Jr., Macken CA. Mitigation strategies for pandemic influenza in the United States. Proc Natl Acad Sci U S A. 2006;103(15):5935-5940. doi:10.1073/pnas.0601266103.

32. Brooks LC, Farrow DC, Hyun S, Tibshirani RJ, Rosenfeld R. Flexible Modeling of Epidemics with an Empirical Bayes Framework. PLoS Comput Biol. 2015;11(8):e1004382. doi:10.1371/journal.pcbi.1004382.

33. Kar TK, Batabyal A. Stability analysis and optimal control of an SIR epidemic model with vaccination. Biosystems. 2011;104(2-3):127-135. doi:10.1016/j.biosystems.2011.02.001.

34. Zaman G, Kang YH, Jung IH. Optimal treatment of an SIR epidemic model with time delay. Biosystems. 2009;98(1):43-50. doi:10.1016/j.biosystems.2009.05.006.