Available online at www.sciencedirect.com

Public Health

journal homepage: www.elsevier.com/puhe

Original Research

Studying the influence of mass media and environmental factors on influenza virus transmission in the US Midwest

S.R. Niakan Kalhori [a], M. Ghazisaeedi [a], R. Azizi [b], A. Naserpour [a,*]

[a] Department of Health Information Management, School of Allied Medical Sciences, Tehran University of Medical Sciences, Tehran, Iran

[b] Department of Health Information Management, School of Allied Medical Sciences, Lorestan University of Medical Sciences, Khorramabad, Iran

ARTICLE INFO

Article history:
Received 21 July 2018
Received in revised form
31 January 2019
Accepted 3 February 2019
Available online 20 March 2019

Keywords:
Influenza outbreak
Public health
Mass media
Climate parameters
Parameter estimation

ABSTRACT

Objectives: Disease burden and high financial cost of seasonal influenza emphasize the importance of studying the epidemics transmission dynamics. Our aim in this article is to extend the Susceptible Exposed Infectious Recovered (SEIR) model, a well-studied classical compartmental epidemic model, by incorporating socio-environmental factors. Particularly, the potential influence of mass media function and absolute humidity are examined on the model simultaneously.

Study design: The proposed model is fitted to Center for Disease Control and Prevention (CDC) influenza data of region five of the US for four outbreak seasons. Then, a full-performance comparison between the conventional and extended model is carried out.

Methods: Implementing the mass media and climate factors into the classical epidemic models, e.g., Susceptible Infectious Recovered (SIR) and SEIR, is a promising and ongoing research field in the public health area. In this article, we particularly address the potential effect of mass media and absolute humidity to modify the SEIR model.

Results: Computational simulations are carried out for both standard and extended models for four influenza seasons in CDC region five of the US. Moreover, the accuracy assessment is performed based on the following criteria: i) the root mean square error (RMSE); ii) the Akaike information criterion (AIC); iii) the outbreak peak time; and iv) the number of infected individuals at the peak time. Based on these criteria, the proposed model provided a better fit than a null model with smaller RMSE and AIC values for the last three study seasons. Specifically, RMSE values declined from 20 to 11.08 and from 26.87 to 19.15 for seasons 2010/11 and 2011/12, respectively; also, lower AIC values for these seasons indicate that the modified SEIR (referred to M-SEIR) model is a better-fitting model.

Conclusions: Parameter estimation techniques are important tools to determine the key parameters of the epidemic models. Based on our results, introducing the mass media and climate factors into the classic models will improve the model precision.

* *Corresponding author.*
E-mail address: anaserpour92@gmail.com (A. Naserpour).

https://doi.org/10.1016/j.puhe.2019.02.006

Introduction

Flu outbreaks occur around the world annually and cause a significant number of deaths.[1] According to US statistics, the annual prevalence of influenza is between 5% and 20% of the population, with 3000 to 49,000 death reports.[2,3] In spite of the widespread outbreak of influenza and its serious threat to the public health, various characteristics of the disease – for example, flu transmissibility and virus survival mechanisms – are still not fully understood. Researchers have long suspected the impact of socio-environmental and geographic factors[4] on the spread of influenza. This association might be attributed to the seasonal nature of the influenza outbreaks. Previous studies have revealed that the solar radiation reduces the survival time of the influenza virus.[5,6] Similar studies also show that low relative humidity and air temperature provide a favorite condition for influenza virus maintenance and transmission.[7,8] In a comprehensive study conducted in 365 US cities, authors provide evidence for an association between absolute humidity and influenza death rate; this correlation is expressed as a non-linear relationship between flu mortality rate and absolute humidity.[3] Our study also set out to examine the influence of mass media on the classical epidemic models. In recent years, more precise influenza surveillances have been developed by using social media data streams, e.g., Twitter and the google search query.[9–11] Providing a real-time influenza predictor, some scholars have also used Twitter data to extend the classic epidemic models.[12–14] (Fig. 1).

A data-driven approach is presented to achieve a mathematical paradigm developed by the implementation of social media and meteorological factors into the Susceptible Exposed Infectious Recovered (SEIR) model. In this investigation, a more accurate and efficient method is introduced to simulate the transmission of influenza disease. The modification is made through embedding a mass media function and absolute humidity into the classic SEIR models. Then, the Center for Disease Control and Prevention (CDC) data for the region five of the US are fitted with both the modified and null model. Our computation was carried out over the influenza season, spanning 25 weeks starting from week 50 of each year. It should be noted that for the sake of clarity, results will be illustrated from week 10 of the influenza season.

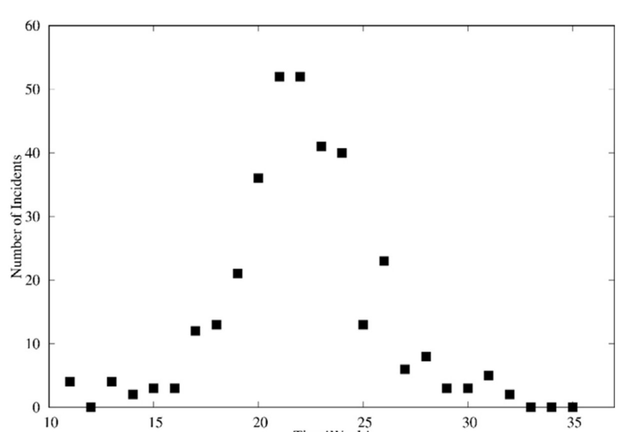

Fig. 1 – The number of H3N2 cases reported weekly by CDC for season 2006/07. CDC, Center for Disease Control and Prevention.

Methods

The CDC provides different forms of valuable sets of information. In this research, weekly laboratory-confirmed (Virologic Surveillance) data in US Environmental Protection Agency (EPA) region 5 (including Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin) are collected from the CDC homepage for four seasons from 2006/07 to 2011/12.[15] Pandemic seasons 2008/09 and 2009/10 were excluded from the data set. The proposed methodology investigates the behavior of subtype H3N2 because it was the dominant virus for the studied seasons.

Daily 2-m specific humidity data were obtained from the National Centers for Environmental Prediction-Department of Energy (NCEP-DOE) Reanalysis two data set. The data are publicly available at Earth System Research Laboratory (ESRL), which is a division of the National Oceanic and Atmospheric Administration (NOAA).[16,17] The daily data in a rectangular block of 36.5776N–48.9994N to 81.9097W-97.2382W, that covers the area of interest, have been downloaded for the time period between 2000 and 2016.[18] Then, to get a single value for every week, a temporal average filter on the original data was applied.

Mathematical modeling has provided various robust and effective techniques to study the transmission of infectious diseases.[19] To model the seasonal influenza behavior, the SEIR model is used as the basic model.[20–22] In the SEIR model, the population is divided into four subgroups: hosts who are susceptible to influenza (S), hosts who are exposed to influenza (E), hosts with influenza infection (I), and the rest of the population who are immune to the influenza disease (R). In this model, β represents the effective transmission rate, $^1/_\gamma$ is the average infectious period, and $^1/_\sigma$ is the average latent period. The SEIR system of equations is given in the following:

$$\begin{aligned} \frac{\partial S}{\partial t} &= -\beta SI \\ \frac{\partial I}{\partial t} &= \beta SI - \sigma E \\ \frac{\partial I}{\partial t} &= \sigma E - \gamma I \end{aligned} \tag{1}$$

where t is the time period. To examine the influence of social and environmental factors on the disease transmission, a modified SEIR (M-SEIR) model is developed by embedding media effect and specific humidity, i.e., an important environmental parameter and a common measure of absolute humidity. In this new model, transmission rate, β, is allowed to be adjusted by a time-dependent linear factor of average absolute humidity:[23]

$$\beta_t = \beta_1(1 + \beta_2 AH) \tag{2}$$

where β_1 and β_2 are constant coefficients and AH is the normalized absolute humidity. Moreover, the media coverage influence, on the other hand, is usually modeled by an exponential or a reciprocal function of I.[13,14,24] For the measurements of this study, a reciprocal function is chosen. Therefore, the final M-SEIR is presented in the following equations:

$$\frac{\partial S}{\partial t} = -\frac{\beta_t}{1 - pI} SI$$
$$\frac{\partial I}{\partial t} = \frac{\beta_t}{1 - pI} SI - \sigma E \qquad (3)$$
$$\frac{\partial I}{\partial t} = \sigma E - \gamma I$$

To examine the performance of the proposed model, the values of the model parameters are determined by means of a curve-fitting strategy. Dozens of research works have been conducted on parameter estimation of epidemiological models using computational techniques including Monte Carlo, maximum likelihood, and least square.[22,25] Least square fitting is a mathematical process to find the best fitting curve for a given set of data. M-SEIR parameters are estimated by minimizing the least square function. The optimization process of the least squares cost function is carried out in Matlab using the 'lsqcurvefit' function and the 'trust region reflective' algorithm.

Given all constant coefficients and initial conditions, a unique solution exists for Eqs. (1) and (3). Therefore, the unknown set of parameters for Eqs. (1) or (3) uniquely corresponds to an optimal state. Generally speaking, any least square optimization problem aims to minimize the standardized difference between predictions and observations. Subsequently, estimated parameters of SEIR and M-SEIR are $\theta = (S_0, E_0, I_0, \beta, \sigma, \gamma)$ and $\tilde{\theta} = (S_0, E_0, I_0, \beta, \sigma, \gamma, p, \beta_2)$, respectively. At each iteration of the algorithm, the estimated parameters obtained from the previous step are used to calculate S, E, I, and R by solving Eqs. (1) and (3). Afterward, SEIR-based and M-SEIR–based estimations of weekly number of infected cases are calculated as follows:[23]

$$\widehat{Y}_i = S(i) + E(i) - S(i-1) - E(i-1) \qquad (4)$$

where i is the week number in the studied influenza season. Then, the sum of the squared differences between estimations, $\widehat{Y}_i$, and observations is minimized by applying the trust region reflective algorithm in Matlab. For the sake of numerical stability of the optimization process, the following constrained conditions are implemented: $^1/_\sigma$ is varied from 1 to 3 days and $^1/_\gamma$ is varied from 1 to 5 days.[24] Also, it is supposed that β_2 is bounded in the interval [−1, 1] and p lies in the interval [−0.2, 0].

Results

As mentioned in the previous section, the present study implements absolute humidity and mass media into the SEIR model. The mass media effect on the influenza behavior dynamics is modeled by using a media function, f(I), in Eq. (3) which is inversely proportional to I (number of infected individuals). On the other hand, the transmission rate (β) is updated to a time-varying coefficient in terms of a linear factor of AH (Eq. (1)). The absolute humidity trend was obtained by averaging time-aligned data over 15 years (2000–2015); then, it is normalized between −1 and 1. We let AH adapt the transmission rate, β_t, by a scale factor varied from 0.8 to 1.2 (1± 0.2). Finally, the efficiency of the proposed model is determined by using root mean square error (RMSE), Akaike information criterion (AIC), and outbreak peak timing/magnitude (Fig. 2).

The weekly CDC data are fitted with both SEIR and M-SEIR models for four influenza seasons. It should be noted that there are 6 and 8 unknown parameters in the classical and modified models, respectively. The results of the curve-fitting procedure for both models are shown in Fig. 3. In this figure, the solid squares represent the CDC influenza data and red and green lines fit to the standard and extended models, respectively. The details of the model performance evaluations are provided in the following.

The RMSE, AIC, the peak time of the outbreak, and the outbreak magnitude at the peak time have been used as standards for performance comparisons. These criteria are summarized in Table 1. Generally speaking, the M-SEIR model captures more accurately the influenza dynamics than the SEIR model that is conventionally used. The proposed model considerably improves the outbreak peak time estimation and do slightly better in terms of the number of infections at the peak time. In addition, the extended model outperforms the basic model in terms of RMSE and AIC errors for the last three study seasons. Specifically, the RMSE has been decreased from 20 to 11.08 and from 26.87 to 19.15 for seasons 2010/11 and 2011/12, respectively. Furthermore, as shown in Table 1, AIC values declined for the last three seasons, and also, the peak time is estimated correctly for all study seasons.

As shown in Fig. 3 and Table 1, in some cases that the data are not smooth enough, the proposed approach to fitting the CDC data is more appropriate than the standard model. Finally, the M-SEIR–based estimations of Susceptible, Exposed, Infected, and Removed parameters for season 2010/11 are illustrated in Fig 4.

Discussion

The advance of mass media in the recent decades has shaped our social behavior and decisions. The role of modern media

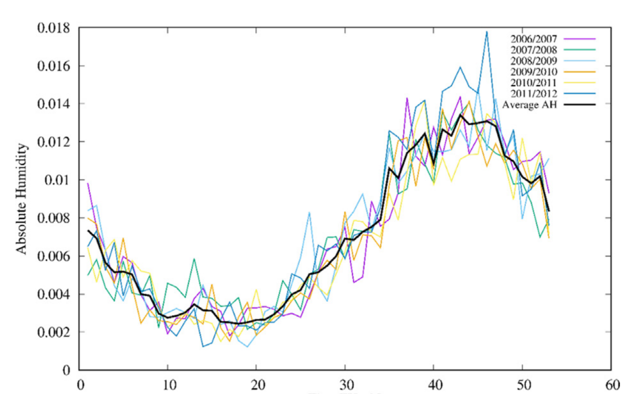

Fig. 2 – Absolute humidity for seasons 2006/2007 to 2011/2012 are displayed in terms of temporal functions. The black solid curve illustrates the average absolute humidity.

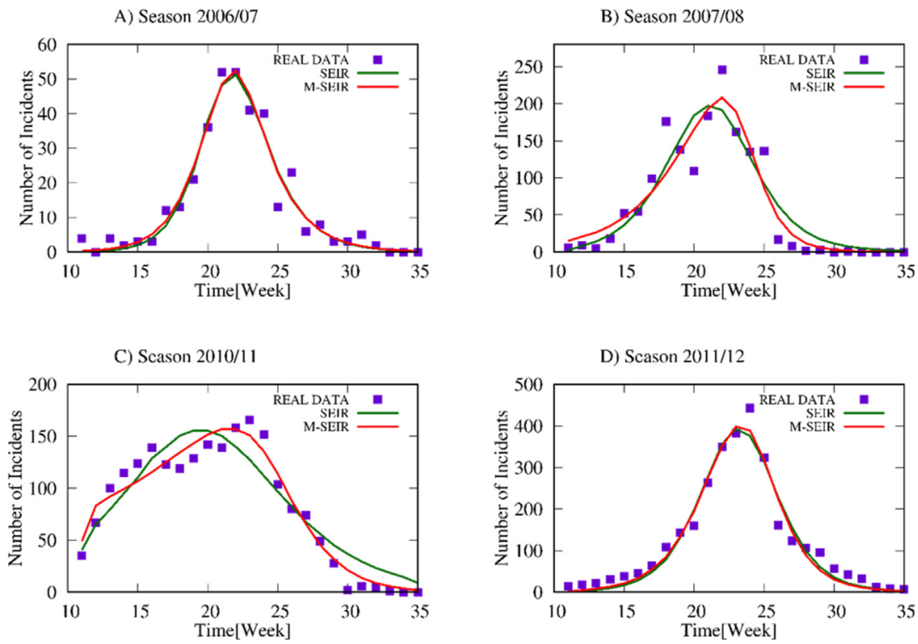

Fig. 3 – Weekly number of infected individuals and corresponding standard SEIR–based and modified SEIR–based estimations for studied outbreak seasons are depicted. Solid purple squares, green lines, and red lines represent the real data, SEIR-, and M-SEIR–based estimations, respectively. SEIR, Susceptible Exposed Infectious Recovered; M-SEIR, Modified Susceptible Exposed Infectious Recovered. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

in promoting the health in the society is inevitable. Media coverage of recent epidemics in news and social networks, for example, has helped to slow the effects of the flu virus. Because human behavior can influence the spreading dynamics of the epidemics, new computational models are therefore necessary to study this subject. In recent years, many researchers attempted to modify the conventional epidemic models through implementation of Twitter and Google Flu Trend as the complementary data streams.[26,27] To model this phenomenon, a time-varying scaling factor is introduced, referred to as the media function, into the SEIR model. Media function usually is a non-linear decreasing function of the number of infected cases, f(I). As given in Eq. (3), a first-order reciprocal media function was included in the model in which p is the corresponding media factor and I is the number of infected individuals at the present time (week). The results indicated that the extended model better fitted to the CDC data in the area of interest. According to our measurements, implementation of the appropriate media functions in the conventional epidemic models is inevitable. The findings of the present study is consistent with those of previous studies, suggesting that social media influence the transmission dynamics of influenza virus.[10,11] Mitchell and Ross[24] showed that implementation of mass media function into the Susceptible Infectious Recovered (SIR) model led to a considerable enhancement of influenza modeling. In another analogous study of the relationship between mass media and influenza dynamics, Paul et al. [27] reduced the forecasting error by using a basic linear autoregressive model incorporating secondary data from Twitter.

In cold and dry weather conditions, experimental results confirm that the influenza virus survives longer outside the human body due to the higher transmission rate.[3,28] However, the association of climate drivers with the virus behavior in different climates is not fully understood yet. Some studies address the role of absolute humidity in temperate regions, for example, Lowen et al.[7] for the first time showed that cold temperatures and low relative humidity are favorable to the spread of influenza virus. On the other hand, Chong et al. [29]

Table 1 – The performance comparison of the SEIR and M-SEIR models in terms of RMSE, AIC, the outbreak peak time, and the magnitude of the outbreak at the peak time.

Season/Criterion	2006/07	2007/08	2010/11	2011/12
RMSE				
SEIR	3.56	27.86	20.00	26.87
M-SEIR	3.55	22.23	11.08	9.15
AIC				
SEIR	77.12	184.00	166.78	182.00
M-SEIR	80.92	182.84	142.07	168.52
Peak time (week)				
SEIR	12	11	9	13
M-SEIR	11	12	13	14
Real	11	12	13	14
Peak mag. (case)				
SEIR	51	197	155	394
M-SEIR	52	209	156	418
Real	52	246	166	443

AIC, Akaike information criterion; RMSE, root mean square error; SEIR, Susceptible Exposed Infectious Recovered; M-SEIR, Modified Susceptible Exposed Infectious Recovered.

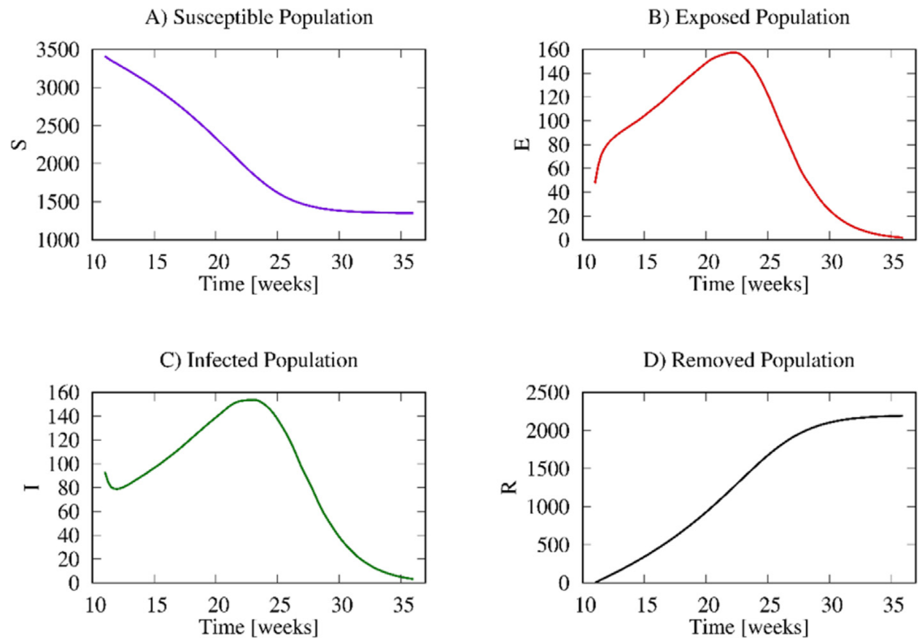

Fig. 4 – Plots A to D display M-SEIR–based Susceptible, Exposed, Infected, and Removed populations for season 2011/12, respectively. M-SEIR, Modified Susceptible Exposed Infectious Recovered.

studied the role of climate drivers in Hong Kong that has a subtropical climate. They showed that absolute humidity does not have any strong correlation with the number of infections. Besides mass media factor, the influence of absolute humidity to the influenza dynamics has also been investigated. As given in Eq. (2), we let the transmission rate to be updated by using an adjusting scale factor varied from 0.8 to 1.2. A substantial improvement was not observed in the absence of absolute humidity compared to mass media.

In this study, however, our results address the significant influence of social and environmental factors into the infectious disease modeling, and further investigations in more areas with different socio-environmental conditions are required. The use of spatially averaged climate data over a relatively vast area is a clear limitation of the present study. In addition, the homogeneity assumption of the studied population in the SEIR model is also obviously violated in the real-life data.

Conclusion

The influenza virus is a contagious respiratory illness that causes a tremendous amount of financial expenses and high morbidity and mortality rate annually. In this article, the effect of mass media and environmental parameters on the flu virus behavior has been closely examined. This research suggests that using mass media function associated to social network yields more remarkable findings than embedding absolute humidity into the epidemic model in EPA region 5. However, a comprehensive study of epidemic models for different regions is required. In the case of data availability, further similar studies can be conducted on smaller areas, possibly at county level, with different climate and social conditions.

Author statements

Acknowledgment and funding

This research was supported by the Tehran University of Medical Sciences. The authors would like to express their sincere gratitude for research and computational resources provide by the Department of Medical Informatics.

Ethical approval

Our investigation was performed using the CDC and ESRL publicly accessible data sets. Thus, no ethical approval was required for this study.

Competing interests

None declared.

REFERENCES

1. Cox N, Subbarao K. Global epidemiology of influenza: past and present. *Annu Rev Med* 2000;**51**(1):407–21.
2. Flu Statistics. *What are your odds of getting the flu?*. 2018. https://www.webmd.com/cold-and-flu/flu-statistics. [Accessed 12 February 2018].
3. Sullivan KM, Monto AS, Longini Jr IM. Estimates of the US health impact of influenza. *Am J Public Health* 1993;**83**(12):1712–6.
4. Šubelj M. Epidemiologic patterns of influenza outbreaks in institutional settings. *Publ Health* 2018;**155**:23–5.

5. Powell WF, Setlow RB. The effect of monochromatic ultraviolet radiation on the interfering property of influenza virus. *Virology* 1956;**2**(3):337–43.
6. Tamm I, Fluke DJ. The effect of monochromatic ultraviolet radiation on the infectivity and hemagglutinating ability of the influenza virus type A strain PR-8. *J Bacteriol* 1950;**59**(4):449.
7. Lowen AC, Mubareka S, Steel J, Palese P. Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog* 2007;**3**(10):e151.
8. Polozov IV, Bezrukov L, Gawrisch K, Zimmerberg J. Progressive ordering with decreasing temperature of the phospholipids of influenza virus. *Nat Chem Biol* 2008;**4**(4):248.
9. Gesualdo F, Stilo G, Gonfiantini MV, Pandolfi E, Velardi P, Tozzi AE. Influenza-like illness surveillance on Twitter through automated learning of naïve language. *PLoS One* 2013;**8**(12):e82489.
10. Separating fact from fear: tracking flu infections on twitter. In: Lamb A, Paul MJ, Dredze M, editors. *Proceedings of the 2013 conference of the North American chapter of the association for computational linguistics*. Human Language Technologies; 2013.
11. Salathé M, Khandelwal S. Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. *PLoS Comput Biol* 2011;**7**(10):e1002199.
12. Collinson S, Heffernan JM. Modelling the effects of media during an influenza epidemic. *BMC Public Health* 2014;**14**(1):376.
13. Cui J-a, Tao X, Zhu H. An SIS infection model incorporating media coverage. *Rocky Mt J Math* 2008:1323–34.
14. Tchuenche JM, Dube N, Bhunu CP, Smith RJ, Bauch CT. The impact of media coverage on the transmission dynamics of human influenza. *BMC Public Health* 2011;**11**(1):S5.
15. Centers for Disease Control and Prevention. *GIS FluView*. 2017. https://www.cdc.gov/flu/weekly/fluviewinteractive.htm. [Accessed 1 December 2017].
16. Kanamitsu M, Ebisuzaki W, Woollen J, Yang S-K, Hnilo J, Fiorino M, et al. Ncep–doe amip-ii reanalysis (r-2). *Bull Am Meteorol Soc* 2002;**83**(11):1631–43.
17. ESRL : PSD : NCEP-DOE AMIP-II Reanalysis (AKA Reanalysis 2).
18. Earth System Research Laboratory. *Physical sciences division (PSD) web transition information*. 2017. https://www.esrl.noaa.gov/psd/webswitch.html. [Accessed 15 December 2017].
19. Solgi M, Karami M, Poorolajal J. Timely detection of influenza outbreaks in Iran: evaluating the performance of the exponentially weighted moving average. *J Infect Pub Health* 2018;**11**(3):389–92.
20. Biswas MHA, Paiva LT, De Pinho M. A SEIR model for control of infectious diseases with constraints. *Math Biosci Eng* 2014;**11**(4):761–84.
21. Hethcote HW. The mathematics of infectious diseases. *SIAM Rev* 2000;**42**(4):599–653.
22. Leonenko VN, Ivanov SV. Fitting the SEIR model of seasonal influenza outbreak to the incidence data for Russian cities. *Russ J Numer Anal Math Model* 2016;**31**(5):267–79.
23. Zarebski AE, Dawson P, McCaw JM, Moss R. Model selection for seasonal influenza forecasting. *Infect Disease Model* 2017;**2**(1):56–70.
24. Mitchell L, Ross JV. A data-driven model for influenza transmission incorporating media effects. *Royal Soc Open Sci* 2016;**3**(10):160481.
25. Hotta LK. Bayesian melding estimation of a stochastic SEIR model. *Math Popul Stud* 2010;**17**(2):101–11.
26. Kang M, Zhong H, He J, Rutherford S, Yang F. Using google trends for influenza surveillance in South China. *PLoS One* 2013;**8**(1):e55205.
27. Paul MJ, Dredze M, Broniatowski D. Twitter improves influenza forecasting. *PLoS Curr* 2014;**6**.
28. Thompson M, Shay D, Zhou H, Bridges C, Cheng P, Burns E, et al. Estimates of deaths associated with seasonal influenza-United States, 1976-2007. *MMWR (Morb Mortal Wkly Rep)* 2010;**59**(33):1057–62.
29. Chong KC, Goggins W, Zee BCY, Wang MH. Identifying meteorological drivers for the seasonal variations of influenza infections in a subtropical city—Hong Kong. *Int J Environ Res Public Health* 2015;**12**(2):1560–76.