Journal of Theoretical Biology 390 (2016) 1–13

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

# Media coverage and hospital notifications: Correlation analysis and optimal media impact duration to manage a pandemic

Qinling Yan [a], Sanyi Tang [a,*], Sandra Gabriele [b], Jianhong Wu [c]

[a] College of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710062, PR China
[b] Department of Design, School of the Arts, Media, Performance & Design, York University, Toronto, Ontario, Canada M3J 1P3
[c] Centre for Disease Modelling, York University, Toronto, Ontario, Canada M3J 1P3

## HIGHLIGHTS

- We develop functional relationships between the media impact and intensity of mass media coverage.
- The multiple correlation analyses show the strong relations between mass media and H1N1 outbreak .
- Nonlinear least squares estimation to identify the best-fit parameter values from the observed data.
- Uncertainty and sensitivity analyses determine key parameters during early phase of disease outbreak.
- Main results show that media publicity should be focused on how to guide people's behavioral changes.

## ARTICLE INFO

*Article history:*
Received 8 August 2015
Received in revised form
3 November 2015
Accepted 4 November 2015
Available online 12 November 2015

*Keywords:*
Correlation analysis
A/H1N1
Media report
SEIR model
Basic reproduction number
Behavior change

## ABSTRACT

News reporting has the potential to modify a community's knowledge of emerging infectious diseases and affect peoples' attitudes and behavior. Here we developed a quantitative approach to evaluate the effects of media on such behavior. Statistically significant correlations between the number of new hospital notifications, during the 2009 A/H1N1 influenza epidemic in the Shaanxi province of China, and the number of daily news items added to eight major websites were found from Pearson correlation and cross-correlation analyses. We also proposed a novel model to examine the implication for transmission dynamics of these correlations. The model incorporated the media impact function into the intensity of infection, and enhanced the traditional epidemic SEIR model with the addition of media dynamics. We used a nonlinear least squares estimation to identify the best-fit parameter values in the model from the observed data. We also carried out the uncertainty and sensitivity analyses to determine key parameters during early phase of the disease outbreak for the final outcome of the outbreak with media impact. The findings confirm the importance of responses by individuals to the media reports, with behavior changes having important consequence for the emerging infectious disease control. Therefore, for mitigating emerging infectious diseases, media reports should be focused on how to guide people's behavioral changes, which are critical for limiting the spread of disease.



## 1. Introduction

In modern society, the interaction between mass media (television, radio, newspapers, billboards, and booklets) and emerging or re-emerging infectious diseases spread and control constitutes a complex and interconnected relationship. On one hand, the degree of mass media attention to a particular disease outbreak may be determined by the infection dynamics and informatics such as daily hospital notifications and disease relevant mortality; and on the other hand, mass media may be used to inform the public infection dynamics and enforce control and prevention measures. Massive media coverage of the infection dynamics can generate profound psychological impacts on the public and potentially alter individuals' behavior (De Silva et al., 2009). Whether and how the number of news items from different mass medias is correlated to the infection dynamics during different stages of a disease outbreak remains unclear, and inferring and quantifying this correlation, falls within the scope of this study.

Existing approaches to modeling the media impact on emerging infectious disease prevention and control have focused largely on exploring how media impact depends on the number of infected individuals at different disease stages (exposed, infected,

* Corresponding author. Tel.: +86 2985310232.
*E-mail addresses:* sytang@snnu.edu.cn, sanyitang219@hotmail.com (S. Tang).

http://dx.doi.org/10.1016/j.jtbi.2015.11.002
0022-5193/

hospitalized and dead) (Cui et al., 2008; Collinson and Heffernan, 2014; Li and Cui, 2009; Liu et al., 2007; Sun et al., 2011) and how this dependence affects the final outcome of an outbreak. The dependence, referred to as media function in some of these existing studies, is often incorporated into the incidence rates (or force of infection) (Cui et al., 2008; Li and Cui, 2009; Liu et al., 2007; Sun et al., 2011). The recent work of Collinson and Heffernan (2014) conducted a comparison of different choices of this dependence, and concluded that media functions used in the literature may not best represent the media effect during an epidemic. It was also noted, in the studies (Tchuenche and Bauch, 2012; Xiao et al., 2013, 2015), that this media function depends not only on the number of infected individuals but its change rate, namely, the media induced reduction of incidence rate should be represented by $\exp(-M(I, dI/dt))$ with $M(I, dI/dt) = \max\{0, p_1 I(t) + p_2 dI(t)/dt\}$ (where $p_1$ and $p_2$ are non-negative constants). The work of Xiao et al. (2015) shows that due to this media impact, the infection dynamics model becomes a switching system where media impact may switch on (new reports) and off (absence of reports) multiple times, thus contributing to the observed multiple waves of an endemic which have been reported for 2003 SARS outbreaks and in 2009 influenza pandemic.

Despite their success in describing how media impact contributes to multiple waves, delays in the outbreak's peak, and reductions in the outbreak, the aforementioned studies failed to develop functional relationships between the observed media impact and the intensity of mass media coverage in different stages of an outbreak, and the feedback relationship between media intensity (news items, duration, etc.) and the diversity in the type of media (website, TVs, radio, newspaper, etc.) and infection dynamics public information (hospital notifications, mortality, etc.). Developing these relationships and describing how to alter these relationships to achieve optimal disease control outcome (reducing the number of hospital notifications, for example) are objectives of this research. More specifically, we address the following issues: (1) whether the number of news items relevant to the disease is correlated to the disease infection? (2) how strong is this correlation? (3) how the correlation varies during an outbreak? and (4) what outcome can be should we change this correlation?

The rest of this paper is organized as follows. We obtained the number of daily (new) hospital notifications and the accumulated number of hospital notifications of the Shaanxi province (Peoples Republic of China) from the 8th Hospital of Xi'an (the dedicated hospital for the 2009 influenza pandemic in that province), and we obtained the number of daily news items relevant to the 2009 A/H1N1 influenza from eight major websites. We then used the Pearson correlation and cross-correlation analyses to determine the strength of the correlation between these numbers using a variety of statistical techniques (Huang et al., 2011; Haugh, 1976; Koch and Yang, 1986; Sampei and Midori, 2009; Taylor, 1990; William, 2006; Xie et al., 2014; Zhao et al., 2011). Then, we formulated a mathematical model that incorporates a new variable to represent dynamics of media coverage intensity (news items, coverage duration and a response factor of media to the diseases infection) in response to infection dynamics. We also employed the nonlinear least squares estimation to find the best-fit parameter values to the observe data, and calculated the basic reproductive ratio $R_0$ using the next generation matrix approach (Collinson and Heffernan, 2014; Wang and Zhao, 2008). Next, we numerically simulated the dynamic correlation between the intensity of mass media coverage and disease infection dynamics under different scenarios for media coverage duration and media switch on-off frequencies.

## 2. Media and infection dynamics: temporal correlation analysis

### 2.1. Data collection

The laboratory-confirmed cases of pandemic A/H1N1 influenza admitted to the 8th Hospital of Xi'an and the number of new hospital notifications and the accumulated number of hospital notifications for the Shaanxi province from September 3rd to November 16th 2009 are employed in this study. Note that the number of new hospital notifications and the accumulated number of hospital notifications were reported separately every two or three days, and no data were available during the weekends and holidays. To address this irregular reporting, we generated the number of new hospital notifications and the accumulated number of hospital notifications using the cubic spline interpolation method on the surveillance data, and showed in Fig. 1(A and B).

We also obtained the daily news items on the A/H1N1 pandemic from eight major popular websites in the Province: news.cn (xinhua net), sina.com, 163.com, qq.com, people.com, CCTV.com, nhfpc.gov.cn (National Health and Family Planning Commission of the People's Republic of China) and chinacdc.cn (Chinese Center for Disease Control and Prevention) using the key word “A/H1N1” included in the news item title, from September 3rd to November 16th 2009. This is shown in Fig. 1(C and D).

### 2.2. Data analysis

In what follows, $y$ denotes the number of new hospital notifications. We will also use, for the number of daily news items from each source, $x_1$ for news.cn, $x_2$ for sina.com, $x_3$ for 163.com, $x_4$ for qq.com, $x_5$ for people.com, $x_6$ for CCTV.com, $x_7$ for nhfpc.gov.cn and $x_8$ for chinacdc.cn, respectively. We adopt the Pearson correlation (Huang et al., 2011; Taylor, 1990) and the cross-correlation analyses (Huang et al., 2011; Sampei and Midori, 2009; William, 2006; Zhao et al., 2011) methods to explore the relationships between the number of daily news items and the number of new hospital notifications for A/H1N1 in the province during the specified period.

*Pearson correlation analysis*: Using SPSS software (version 19.0, SPSS Inc.), we conducted the Pearson correlation analysis to determine the association between the number of daily news items and the number of new hospital notifications from September 3rd to November 16th 2009. Some results are summarized in Table 1. We conclude that the number of daily news items from website of 163.com ($x_3$), qq.com ($x_4$) and CCTV.com ($x_6$) is statistically significantly correlated with the number of new hospital notifications over the study period. Among these three popular websites, the number of daily news items of qq.com is most closely correlated with the number of new hospital notifications ($r=0.45$, $p<0.01$), while the correlation for CCTV.com is relatively weak ($r=0.28$, $p<0.05$).

We also report in Table 1 the correlations between each pair of websites. This report shows that some of these websites have statistically significant correlations in terms of reporting the A/H1N1 infection dynamics, some are highly correlated or moderately correlated according to the labeling systems roughly categorized (low or weak correlations ($|\gamma| \le 0.35$), moderate correlations ($0.36 \le |\gamma| \le 0.67$) and strong/high correlations ($0.68 \le |\gamma| \le 1.0$), Taylor, 1990). In particular, we notice that $x_5$ is highly correlated to $x_1$, while $x_2$, $x_3$, $x_4$ and $x_8$ are moderately correlated to $x_1$. We also notice the moderate correlation pairs $(x_3, x_2), (x_5, x_2), (x_4, x_3), (x_5, x_3), (x_7, x_3)$, and $(x_6, x_4)$.

Although Pearson correlation analysis can reveal the statistically significant correlation between the number of daily news items and the number of new hospital notifications of A/H1N1, the

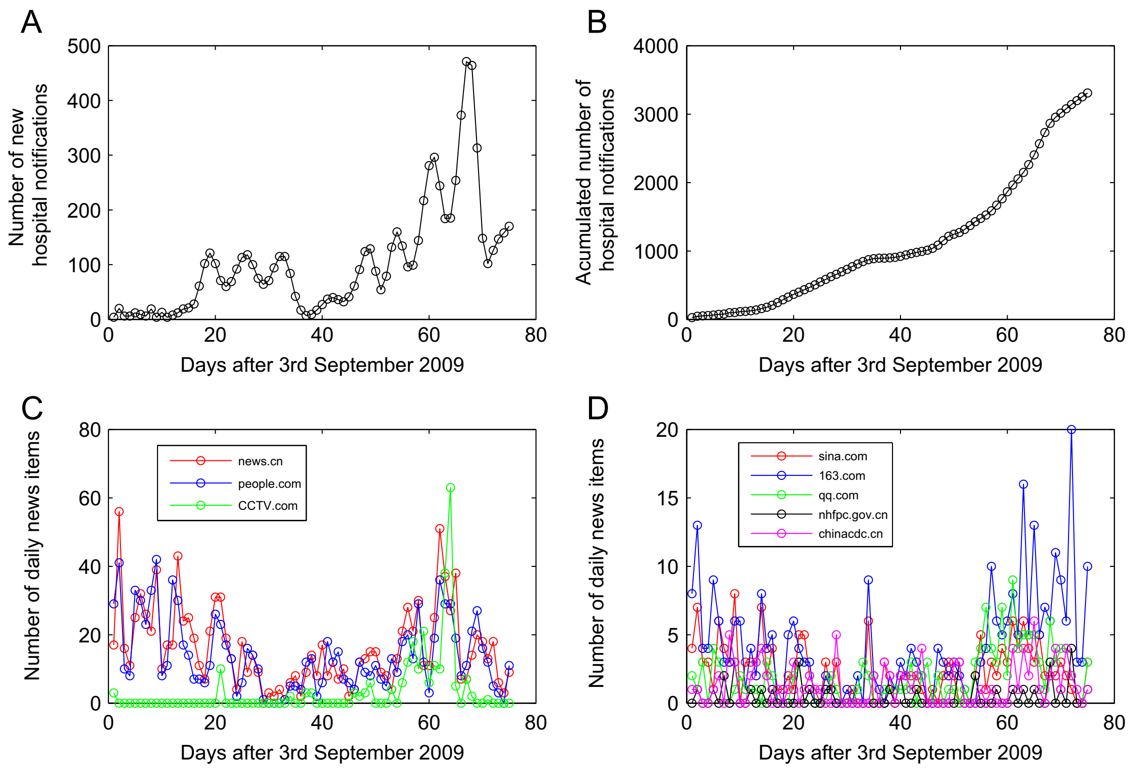


**Fig. 1.** The number of hospital notifications of the Shaanxi province and the number of daily news items from the eight popular sources during September 3rd to November 16th 2009: (A) and (B). The number of new hospital notifications and the accumulated number of hospital notifications for the Shaanxi Province, respectively. (C) and (D). The number of daily news items from news.cn, people.com, CCTV.com and sina.com, 163.com, qq.com, nhfpc.gov.cn and chinacdc.cn, respectively.

**Table 1**
Pearson correlation analysis between the number of daily news items and the number of new hospital notifications of A/H1N1 in Shaanxi Province from September 3rd to November 16th 2009.

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $y$ |
|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | 1 | 0.61** | 0.51** | 0.37** | 0.82** | 0.28* | 0.27* | 0.36** | −0.01 |
| $x_2$ | 0.61** | 1 | 0.45** | 0.31** | 0.59** | 0.27* | 0.22 | 0.22* | 0.06 |
| $x_3$ | 0.51** | 0.45** | 1 | 0.45** | 0.42** | 0.26* | 0.37** | 0.33** | 0.29* |
| $x_4$ | 0.37** | 0.31** | 0.45** | 1 | 0.33** | 0.48** | 0.14 | 0.17 | 0.45** |
| $x_5$ | 0.82** | 0.59** | 0.42** | 0.33** | 1 | 0.27* | 0.24* | 0.39** | −0.02 |
| $x_6$ | 0.28* | 0.27* | 0.25* | 0.49** | 0.27* | 1 | −0.04 | 0.08 | 0.28* |
| $x_7$ | 0.27* | 0.22 | 0.37** | 0.14 | 0.24* | −0.04 | 1 | 0.16 | 0.06 |
| $x_8$ | 0.36** | 0.28* | 0.33** | 0.17 | 0.39** | 0.08 | 0.16 | 1 | 0.04 |
| $y$ | −0.01 | 0.06 | 0.29* | 0.45** | −0.02 | 0.28* | 0.06 | 0.04 | 1 |

* Significance of correlation coefficient different from zero represents $p<0.05$.
** Significance of correlation coefficient different from zero represents $p<0.01$.

important issue of identifying time lags between the daily reported news and the daily hospital notifications needs further cross-correlation analysis.

*Cross-correlation analysis*: Cross-correlation is a spectral analysis technique that can be used to provide qualitative insights on the casual temporal interaction between two processes. This technique is thus particularly useful for us to detect causal interactions between the number of daily news items and the number of new hospital notifications of A/H1N1. In the following, we report our cross-correlation analysis, using the data from the Shaanxi province to reveal the range of the cross-correlation coefficients and to identify the specific value of time lags where these coefficients achieve their maximum values. This information is critical to fitting the data to our proposed dynamic models (see more details in Appendix A) in order to illustrate the media impact.

We calculated the cross-correlation function (CCF) between the number of daily news items of each website and the number of new hospital notifications. The results are summarized in Fig. 2. We observed that there are statistically significant cross-correlation between the number of daily news items at the sites $x_3, x_4$ and $x_6$ and the number of new hospital notifications $y$ at lags ranging from $-7$ to 5 days, $-4$ to 7 days, $-4$ to 7 days, respectively (see Fig. 2(C, D, F)). Therefore, there exists a feedback relationship (for $k\neq 0$) and a contemporaneous relationship (for $k=0$) between the number of daily news items at $x_3, x_4, x_6$ and the number of new hospital notifications, and the relation is linear in the range of time lags. We observed that the local maximal cross correlation coefficient exceeds 0.5 (moderately or highly correlated) occur at the time lags of $-3$, 4, 4 days (Fig. 2), respectively.

It follows from the Pearson correlation analysis and the cross-correlation analysis that the number of daily news items at the three most popular, websites (163.com, qq.com and CCTV.com) during the period, was strongly correlated with the number of new hospital notifications. Furthermore, a statistically significant cross-correlation between $x_1$ and $y$ at lags ranging from 4 to 6 days clearly indicates that the number of daily news items from the website news.cn caused the number of new hospital notifications, with a linear relationship when the delay is within this range. This delay should be considered as the incubation period for A/H1N1 infection plus the period of behavior changes in the population in response to the media reports.

As different individuals might only focus on the news from their respective preferred sites, we considered the average of the number of daily news items from the four statistically significantly correlated websites $x_1, x_3, x_4, x_6$, denoted by $(x)$, for additional cross-correlation analysis. We calculated the CCF of average number of daily news items $(x)$ and the number of new hospital notifications $(y)$. Fig. 3(A) shows that there are statistically significant cross-correlation between the average number of daily

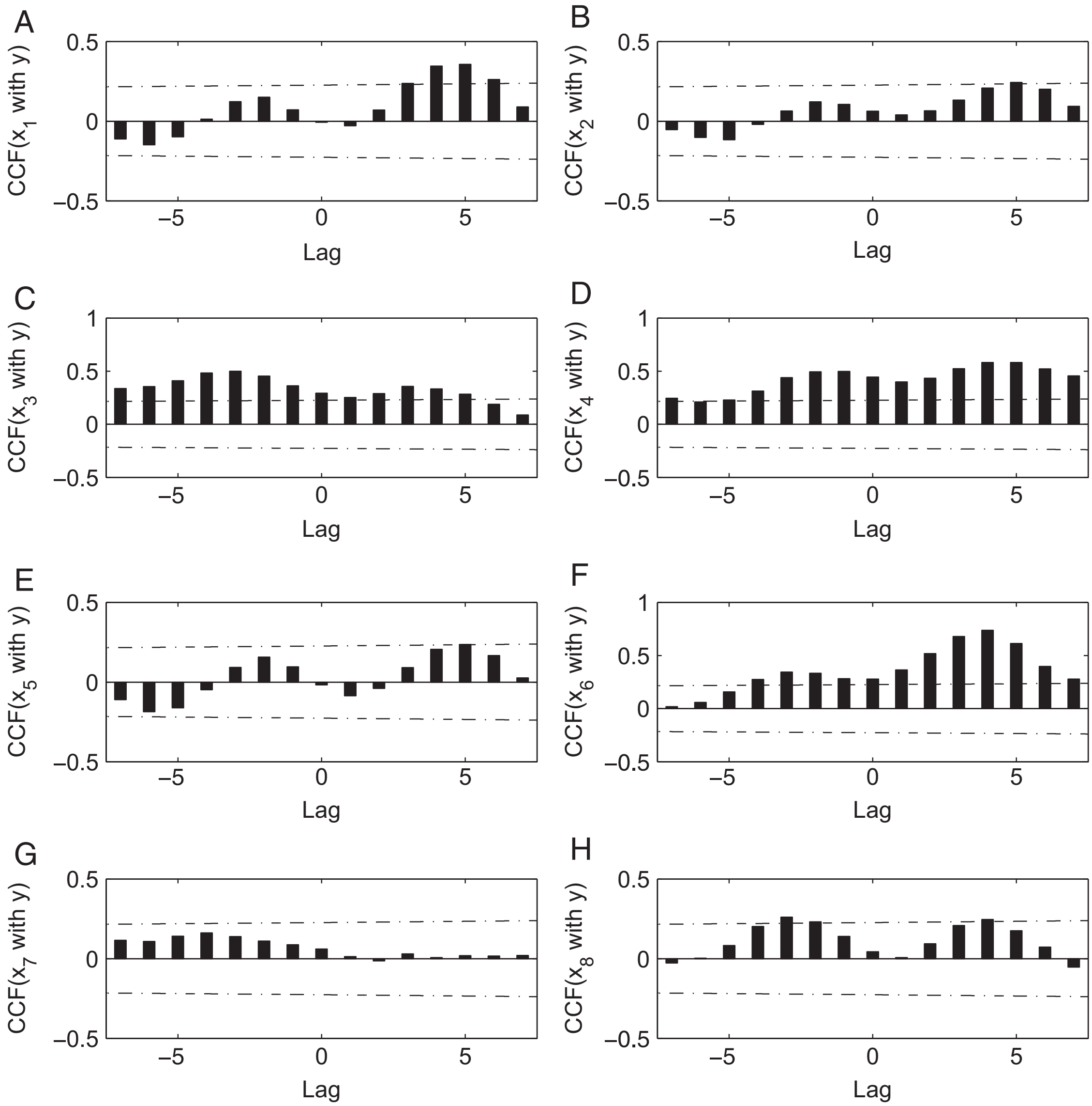


**Fig. 2.** Cross-correlation coefficients between the number of daily news items and the number of new hospital notifications of A/H1N1 in Shaanxi Province from September 3rd to November 16th 2009. The two dotted lines in the graphs represent the upper and lower confidence bounds of 95% confidence intervals.

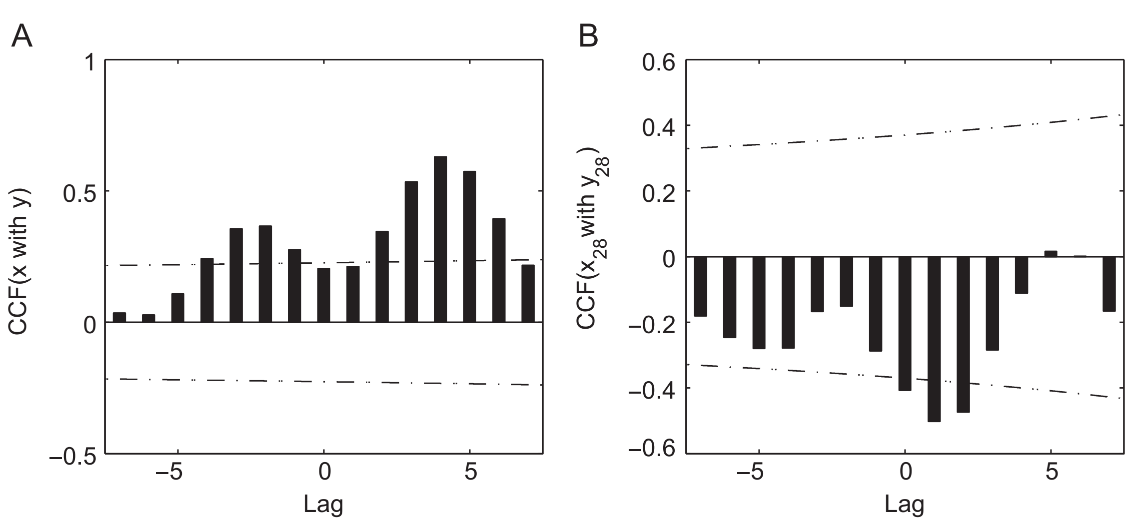


**Fig. 3.** Cross-correlation coefficients between the average number of daily news items and the number of new hospital notifications, from 3rd September to 16th November (A) and from 3rd to 30th September (B), respectively. The two dotted lines in the graphs represent the upper and lower confidence bounds of 95% confidence intervals.

news items $x$ and the number of new hospital notifications $y$ with the time lags ranging from $-3$ to 1 days and 2 to 6 days, and furthermore the local maximal cross-correlation coefficient is 0.63 (moderately correlated) and this occurred with the 4 days lag.

As discussed in Tang et al. (2010), and further illustrated in the data set shown in Fig. 1(A), the first wave declined quickly due to very strict interventions in early September 2009, but that declining trend was reversed following the October National Day holiday during which population mobility increased, the integrated non-pharmacological therapies including *Fengxiao* (school closure) and isolation were suspended. As such, we chose to focus on the period from 3rd to 30th September (the exponential growth phase of the first wave) to understand the casual interaction between the average number of daily news items and the number of new hospital notifications.

We denote the average number of daily news items and the number of new hospital notifications during the first 28 days by $x_{28}$ and $y_{28}$, and the cross-correlation analysis result is shown in Fig. 3(B). The analysis shows a statistically significant cross correlation between $x_{28}$ and $y_{28}$ at lags ranging from 1 to 2 days, which suggested that the number of new hospital notifications delay, and are possibly caused by, the average number of daily news items. And also a contemporaneous relationship (for $k=0$) and a linear relationship exist in both at the range of lags days investigated.

All of these confirm that not only the average number of daily news items had impact on the disease control (in terms of the number of new hospital notifications), but also the disease dynamics statistics (in terms of new hospital notifications) affected the news coverage, in the long run (see Figs. 2 and 3(A)). On the other hand, we noticed that in the short term, only the average

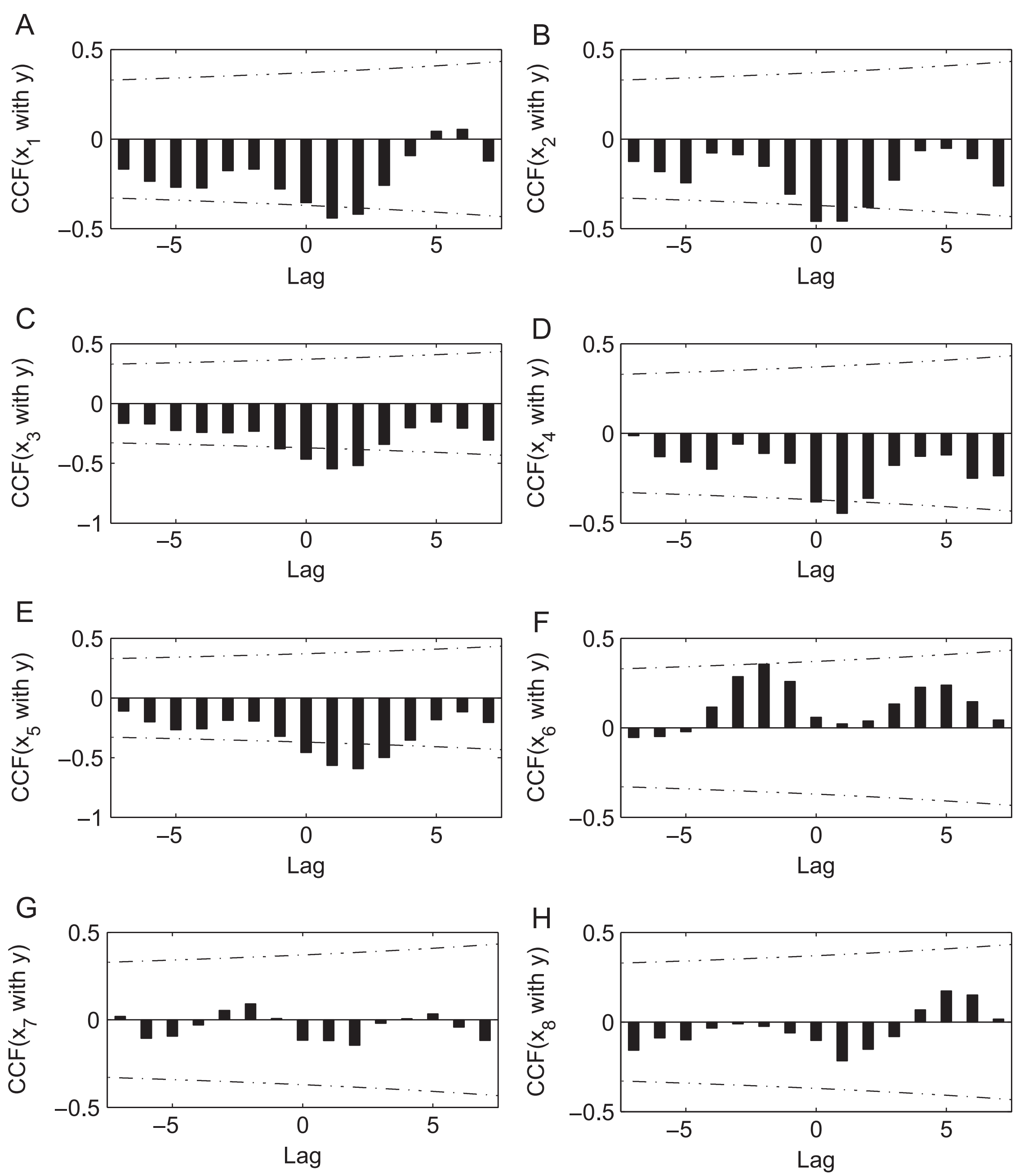


**Fig. 4.** Cross-correlation coefficients between the number of daily news items and the number of new hospital notifications of A/H1N1 in Shaanxi Province from September 3rd to 30th 2009. The two dotted lines in the graphs represent the upper and lower confidence bounds of the 95% confidence intervals.

number of daily news items caused the number of new hospital notifications (see Figs. 4 and 3(B)). These findings are confirmed by the patterns (two distinct waves) of the A/H1N1 outbreak in mainland China. During the first wave, the outbreak resulted in substantially increasing media coverage, which in turn helped increase the population awareness and alter their behaviors for a better implementation of mitigations. However, shortly after the peak of the first wave, coinciding with the National Day holiday, the media coverage reduced significantly. It was only until after this National Day holiday, and during the second wave, the number of new hospital notifications increased and so did the media coverage as shown in Figs. 4 and 3(B).

In conclusion, our Pearson correlation and cross-correlation analyses confirmed the correlation and interaction between the number of daily news items and the number of new hospital notifications and estimated the time lags between these interactions. However, the standard cross-correlation theory assumes that the observations were independent and stationary, whereas, in reality, the number of new hospital notifications and the number of daily news items were temporally correlated and non-stationary. It is important to incorporate these statistical analysis into a dynamic model to address the issue how the average number of daily news items and the number of new hospital notifications affect each other dynamically. This is what we aim to achieve in the next section.

## 3. Model formulation

### 3.1. SEIRM framework

We now incorporate media impact into the classical SEIR (susceptible-exposed-infective-recovery) type epidemiological model. We stratify the usual susceptible ($S(t)$), exposed ($E(t)$), infected ($I(t)$), and recovered ($R(t)$) compartments in the classical SEIR model, to include the media impacts as a new variable, denoted by $M(t)$, which is the average number of news items related to the outbreak. Let $N=S(t)+E(t)+I(t)+R(t)$ be the total population, and assume that the susceptible individuals are infected by infectious individuals with a rate of $\beta$, and become exposed; exposed individuals become infectious with a rate $\sigma$; recovery individuals are removed with a rate $\gamma$ from the possibility of infection through immunity.

Note that during the initial stage (from September 3rd to 30th) of the A/H1N1 outbreak in Shaanxi province, infected individuals reported through hospital notifications were mainly university/college/school students and were hospitalized in the 8th hospital of Xi'an (Tang et al., 2012), and there was no any death cases reported. Therefore, we assume that the total number of population $N$ is a constant and consider the infectious class $I(t)$ as the hospitalized compartment. But with the October National Day holiday arrival (from 1st October), population mobility increased and *Fengxiao* – that virtually isolated all universities/colleges from the general population – was suspended. After that the A/H1N1 spread from universities/colleges to general population, and consequently the number of susceptible population increased significantly. So, in this study, we shall focus on the data set from 3rd to 30th September (as shown in Fig. 1, the exponential growth phase of the first wave) to fit the proposed model and to examine those issues such as how the average number of daily news items and the accumulated number of hospital notifications affects each other in a dynamic feedback fashion. In what follows, we denote the time interval from 3rd to 30th September as $T_1=[0,28]$.

The mass media can significantly affect disease transmission during the initial stage of an emerging infectious disease (Xiao et al., 2015), and play a key role in defining health issues and providing major source of information which results in changing behavior in the public. We assume that susceptible individuals' response to a disease threat depends on risk perception that is gained largely through information (such as the number of infections, hospitalizations and deaths) through mass media. Therefore, we describe the effects of media reports on population infection dynamics by modifying the transmission rate $\beta$ following (Xiao et al., 2015) as $\beta f(M,p)$, where $f(M,p)$ is a function of the average number of news items $M$, and $p$ is a positive constant which represents the weight of media impact which changes with the accumulated number of hospital notifications. In general, $f(M,p)$ is a decreasing function of $M$, which indicates that, as the number of news items increases, susceptible individuals will practice social distancing or control measures, and consequently decreases the contact transmission rate. In this study, for the sake of simplicity, we use the simplest form of the function $f(M,p)=\exp(-pM)$.

It is important to emphasize that the average number of daily news items and the accumulated number of hospital notifications are inevitably intertwined. In particular, a number of factors including the number of infections, hospitalization and death cases may result in increasing mass media reports, and vice versa, as shown in Fig. 1.

In general, when an emerging infectious disease such as the 2009 A/H1N1 pandemic starts, the number of infected cases is small but the number of daily news items relevant to the disease is high (Xie et al., 2014). After the initial stage, the disease cases increase quickly, but the media coverage begins to decline, as shown in Fig. 1((A), (C) and (D)). Meanwhile, for the data set, we also note that the number of daily news items during weekdays was significantly greater comparing with weekends.

Therefore, we assume that the changing rate of the average number of daily news items depends on the number of new hospital notifications. We will use $\eta(t)\sigma E(t)$ for this dependence, where $T_2$ is the period during which the number of daily news items increased significantly, this period is termed as "media sensitive period" in what follows. More specifically, we have the function $\eta(t)$ defined as

$$\eta(t)=\begin{cases}\eta_0\ \exp(q(T_2-t)), & t\in T_{11}\subset T_1,\\ \eta_1\ \exp(q(T_2-t)), & t\in T_{12}\subset T_1,\end{cases} \tag{3.1}$$

where $\eta_0$ and $\eta_1$ are positive constants which characterize media coverage in response to the number of new hospital notifications during the weekends and workdays, respectively. It is reasonable to assume $\eta_0<\eta_1$. In the above formulation, $q$ is the scaling parameter measuring the sensitivity of the media coverage to the hospital notifications during the study period. As mentioned, during the period from 3rd to 30th September 2009, we had the weekends corresponding to $T_{11}=[3,4]\bigcup[10,11]\bigcup[17,18]\bigcup[24,25]$ and the rest $T_{12}=T_1\backslash T_{11}$ for the weekdays. This leads to the following SEIRM model:

$$\begin{cases}\dfrac{dS(t)}{dt}=-f(M,p)\beta S(t)I(t)/N,\\ \dfrac{dE(t)}{dt}=f(M,p)\beta S(t)I(t)/N-\sigma E(t),\\ \dfrac{dI(t)}{dt}=\sigma E(t)-\gamma I(t),\\ \dfrac{dR(t)}{dt}=\gamma I(t),\\ \dfrac{dM(t)}{dt}=\eta(t)\sigma E(t)-\delta M(t),\end{cases} \tag{3.2}$$

where $\delta$ represents the spontaneous disappearance rate of media reports (media wading rate). The more detailed definitions of variables and parameters for model (3.2) are provided in Table 2.

For a general periodic function $\eta(t)$, we can define the basic reproduction number for the periodic system (3.1), which is actually the spectral radius of an infection operator. This basic reproduction number can be explicitly calculated as $R_0=\beta/\gamma$ (see Appendix B). It is natural that this basic reproduction number is independent of parameters relevant to media impact, confirming the fact that media impacts hardly affect the basic reproduction number (Xiao et al., 2013, 2015), rather it affects the outbreak size and the peak timing.

### 3.2. Parameter estimation and model fitting

Due to irregular data reporting in the province of Shaanxi (e.g. reporting delays at weekends) and reporting policy changes, we had to generate the accumulated number of hospital notifications using the cubic spline interpolation method, implemented as a Matlab program, as shown in Fig. 1(B). Note that the sensitive period $T_2$ and the two parameters $\eta_0$ and $\eta_1$ which describe the response of mass media to the accumulated number of hospital notifications are highly correlated. This shows that we cannot estimate those parameters along with the estimation of the parameter $q$ simultaneously.

As estimated in the previous section, the period of media attention to an outbreak at the very early stage of the outbreak is about 2 or 3 days, thus we first fix $T_2$=2.5 days, $\eta_1=2$ with $\eta_0=\eta_1/3$ and then carry out the uncertainty and sensitivity analyses for those parameter values. By fitting data on the accumulated number of hospital notifications, we obtain other unknown parameter values. To do so, we utilized the nonlinear least-square (NLES) method in Matlab to fit the aforementioned real data sets which correspond to the model solution time series, i.e. $H(t)$ and $\eta(t)\sigma E(t)$, where $H(t)$ follows the $dH(t)/dt=\sigma E(t)$, as shown in Fig. 5. The estimated parameter values are listed in Table 2.

**Table 2**
Initial conditions and parameter values for model (3.2).

| | Description | Value (source) | Mean | Std |
|---|---|---|---|---|
| Population | | | | |
| $S$ | Susceptible | 100 000 | | |
| $E$ | Exposed | 10 | | |
| $I$ | Infected | 4 | | |
| $R$ | Recovered | 0 | | |
| $M$ | Media | 8 | | |
| $H$ | Accumulated number of hospital notifications | 25 | | |
| Parameter | | | | |
| $p$ | Weight of media effects sensitive to accumulated number of hospital notifications | 0.0018 (LS) | 0.0017 | 0.0003 |
| $\beta$ | Contact transmission rate $((person-day)^{-1})$ | 0.63 (LS) | 0.6297 | 0.0031 |
| $\sigma$ | Transition rate E to I $((day)^{-1})$ | 0.6864 (LS) | 0.7056 | 0.0536 |
| $\gamma$ | Recovered rate $((day)^{-1})$ | 0.33 (LS) | 0.3510 | 0.0466 |
| $\delta$ | Media spontaneous disappearance rate $((day)^{-1})$ | 0.01 (LS) | 0.1783 | 0.3737 |
| $q$ | Weight of hospital notifications effects sensitive to the average number of daily news items | 0.001 (LS) | 0.001 | 0.0001 |
| $T_2$ | Sensitive period $(day)$ | 2.5 (fixed) | | |
| $\eta_1$ | Response of mass media on the number of new hospital notifications during weekdays | 2 (fixed) | | |
| $\eta_0$ | Response of mass media on the number of new hospital notifications during weekend | $\eta_1/3$ (fixed) | | |

### 3.3. Basic reproduction number $R_0$

In the absence of available data on the distribution function of all input parameters ($p, \beta, \sigma, \gamma, \delta$ and $q$), it is impossible to estimate the basic reproduction number $R_0$ and its confidence interval from model (3.2), so a stochastic simulation technique was used. Firstly, we generated 500 samples of the accumulated number of hospital notifications and the average number of daily news items from a Poisson process as the counting process is a Poisson process for both of them. Secondly, we used the NLES approach with the 500 samples to fit model (3.2), and consequently we got 500 groups of values for those estimated parameters, which allowed us to calculate the confidence interval of $R_0$. Further, normal distributions were fitted to the frequency histograms of all of the parameters, and the mean values and standard deviations obtained are shown in Table 2.

Based on the above parameter estimations, we then calculated the basic reproductive number $R_0$, which is defined as the number of newly infected individuals produced by a single infected individual in a totally susceptible population during the infectious period. According to the parameter values given in Table 2, the mean value of the $R_0$ was estimated as 1.794 (95% confidence interval is 1.3858–1.9091).

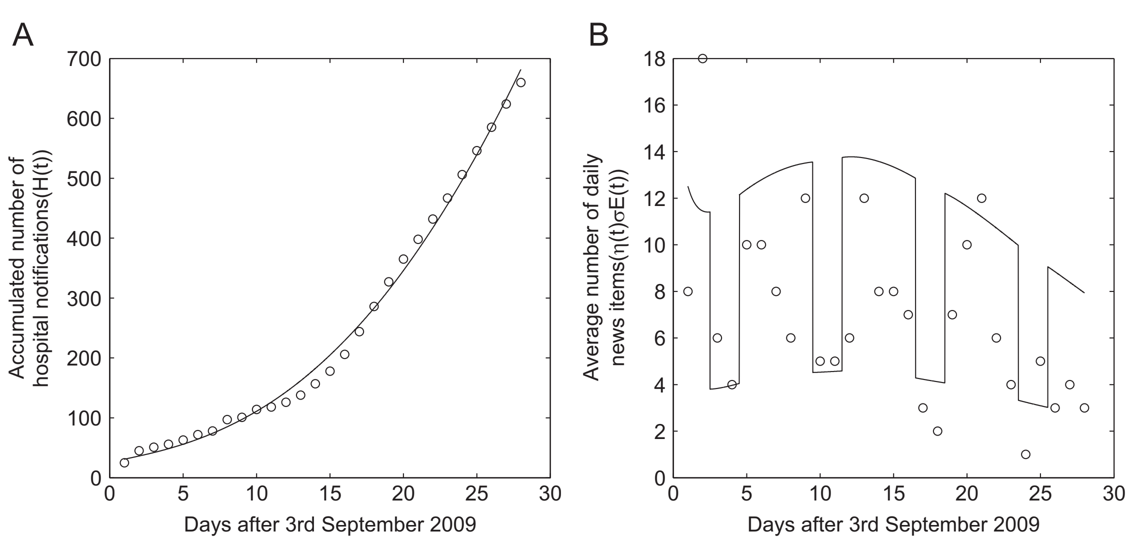


**Fig. 5.** Model fitting for the accumulated number of hospital notifications and the average number of daily news items from 3rd to 30th September.

## 4. Uncertainty and sensitivity analyses

### 4.1. The effects of media

Previous work revealed that media coverage significantly delayed the epidemics peak and decreased the severity of the outbreak (Cui et al., 2008; Liu et al., 2007; Xiao et al., 2013, 2015). Moreover, media impacts are not always effective in lowering disease transmission during an entire outbreak, they can play a key role only during the early stages of an outbreak (Xiao et al., 2015), and the data shown in Fig. 5 revealed that the average number of daily news items significantly decreased after the high concern period, which further confirmed the important role of media in mitigating infections during the early stage.

In this work, we used both the accumulated number of hospital notifications and the average number of daily news items and proposed a novel model including a sensitive period ($T_2$), the response of mass media on the accumulated number of hospital notifications during weekdays ($\eta_1$) and with the weight of hospital notifications effects being sensitive to the average number of daily news items ($q$). This allowed us to evaluate the effects of the parameters related to media coverage on the disease transmission during the entire outbreak.

For different media parameters, we used a wide range of parameter values (i.e. $T_2 \in [0, 10], \eta_1 \in [0.5, 3.5]$ and $\delta \in [0, 0.5]$) to show the effects of media coverage on the disease outbreak. Meanwhile, the weight of media effects sensitive to the accumulated number of hospital notifications $p$ ranged from 0 to 0.01, and the weight of hospital notifications effects sensitive to the average number of daily news items $q$ ranged from 0 to 0.003.

Note that coverage duration $T_2 = 0$ means that the mass media does not highlight the emerging infectious disease, but individuals would behave as if all was normal during the outbreak. The effects of coverage duration $T_2$ on infections with A/H1N1 are shown in Fig. 6(A and B), from which we can see that the simulated average number of daily news items related to A/H1N1 significantly increases as coverage duration increases, and consequently the outbreak of infectious disease is largely reduced. The results indicated that in order to reduce the accumulated number of hospital notifications and therefore weaken the disease outbreak, the mass media not only needs to strengthen their news reports concerning the disease threat such as number of infections, hospitalizations and deaths, but also persistently broadcast news concerning the disease (i.e., increasing $T_2$).

Most importantly, considering the incubation period for A/H1N1 infection is between three to seven days (Tang et al., 2010,

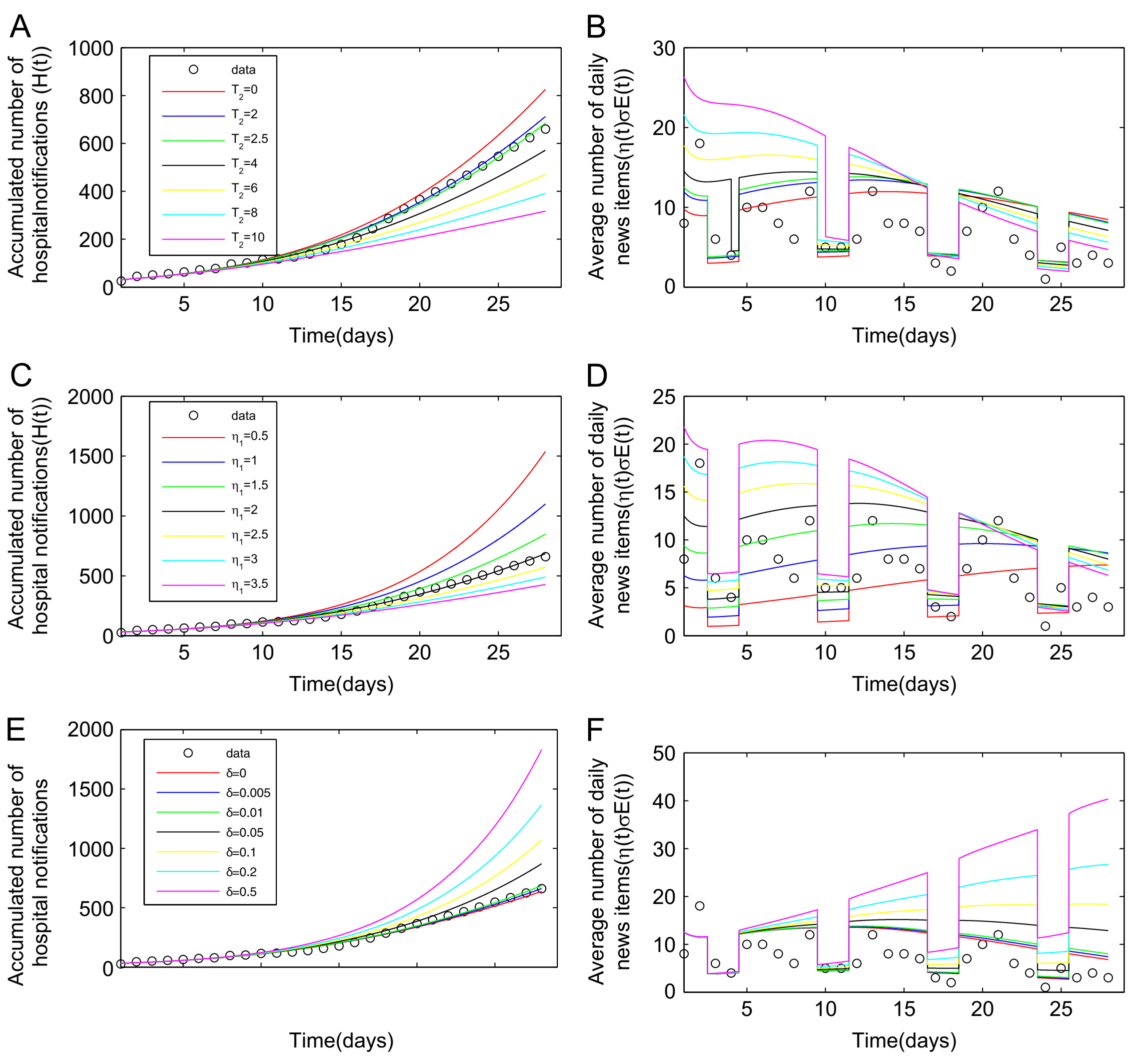


**Fig. 6.** The best-fit solution with the effects of media. (A) and (B) The effects of coverage duration $T_2$ on the accumulated number of hospital notifications and the average number of daily news items, respectively (the baseline $T_2$=2.5). (C) and (D) The effects of the response of mass media on the accumulated number of hospital notifications during weekdays $\eta_1$ and on the average number of daily news items, respectively (the baseline $\eta_1 = 2$). (E) and (F) The effects of the mass media spontaneous disappearance rate $\delta$ on the accumulated number of hospital notifications and on the average number of daily news items, respectively (the baseline $\delta = 0.01$).

2012) and the period of people taking action in responding to the media reports, it is reasonable to reach the conclusion that the effects of media reports were significant, and so media reports need to be timely and to continue for longer periods than they did.

As mentioned before, we chose a wide range of parameter values for $\eta_1$ (and of course for $\eta_0$) to show the significance of the media response to the accumulated number of hospital notifications. Obviously, the simulated number of daily news items increases quickly as $\eta_1$ increases (as shown in Fig. 6(D)), while the simulated accumulated number of hospital notifications decreases with increasing $\eta_1$, as shown in Fig. 6(C). In particular, if the $\eta_1$ decreases below 0.5, then the simulated accumulated number of hospital notifications is increased by 125% by the day 30th September 2009; while the $\eta_1$ reaches up to 2.5, then the simulated accumulated number of hospital notifications is decreased by 16%. But the larger the amplification effect of the media the fewer are its effects on decreasing the accumulated number of hospital notifications and excessive amplification induces panic, so it is important and meaningful for the amplification effect of media to be appropriate in relation to disease control.

Fig. 6 (E) and (F) show the variation in the simulated accumulated number of hospital notifications ($H(t)$) and the average number of reported news items ($\eta(t)\sigma E(t)$) with the rate of disappearance of media reports (representing the duration of mass media concerning the disease infection). It follows from Fig. 6 (E) and (F) that the quicker media reports disappear (i.e. the more the value of $\delta$), the larger the simulated accumulated number of hospital notifications is, and the larger the simulated average number of reported news is. In particular, if the spontaneous disappearance rate of media reports decreases below 0.01 (i.e. mass media can last a quite long time), then the simulated accumulated number of hospital notifications is relatively low. If the spontaneous disappearance rate of media reports reaches up to 0.5, then $\delta$ has few effects of media on the accumulated number of hospital notifications because effects of the average number of daily news items fell away so quickly.

The effects of the parameters $p$ and $q$ on A/H1N1 infection are shown in Fig. 7(A) and (B). It implies that the larger the weight of media effects is sensitive to the accumulated number of hospital notifications ($p$), the less the simulated accumulated number of hospital notifications is and the less the simulated number of reported news items is. Noting that Fig. 7(C) and (D) show that the larger the value of the parameter $q$, the larger the simulated number of hospital notifications is, and the less the simulated number of reported news items is. In particular, the estimated accumulated number of hospital notifications by the 30th September 2009 can be controlled at around 684, and decreased by 64% if the weight of media effects sensitive to the accumulated number of hospital notifications $p$ increased 3 times; while decreased by 30% if the effects of weight of hospital notifications effects sensitive to the number of daily news items $q$ decreased to one-third of the baseline.

All of these results confirm the importance of responses of individuals to the media reports, that is, behavior changing is more important in emerging infectious disease control than the high media attention to unexpected events and reports. Therefore, for mitigating emerging infectious disease, the media publicity should be focused on how to guide people's behavioral changes.

### 4.2. Uncertainty and sensitivity analyses

We explored the parameter space by performing an uncertainty analysis using a Latin hypercube sampling (LHS) method, an extension of Latin sampling. Sensitivity analysis was done by evaluating the partial rank correlation coefficients (PRCCs) (see detail in Blower and Dowlatabadi, 1994; Marino et al., 2008; Mckay et al., 1979) for various input parameters (such as $p, \beta, \sigma, \gamma, \delta$ and $q$) against the output variables (here, the simulated accumulated number of hospital notifications ($H(t)$) and the average number of daily news items ($\eta(t)\sigma E(t)$) with the LHS method over time, and then parameters which have the most significant influence are identified. The mean and standard deviation of the parameters are given in Table 2, and normal distribution functions

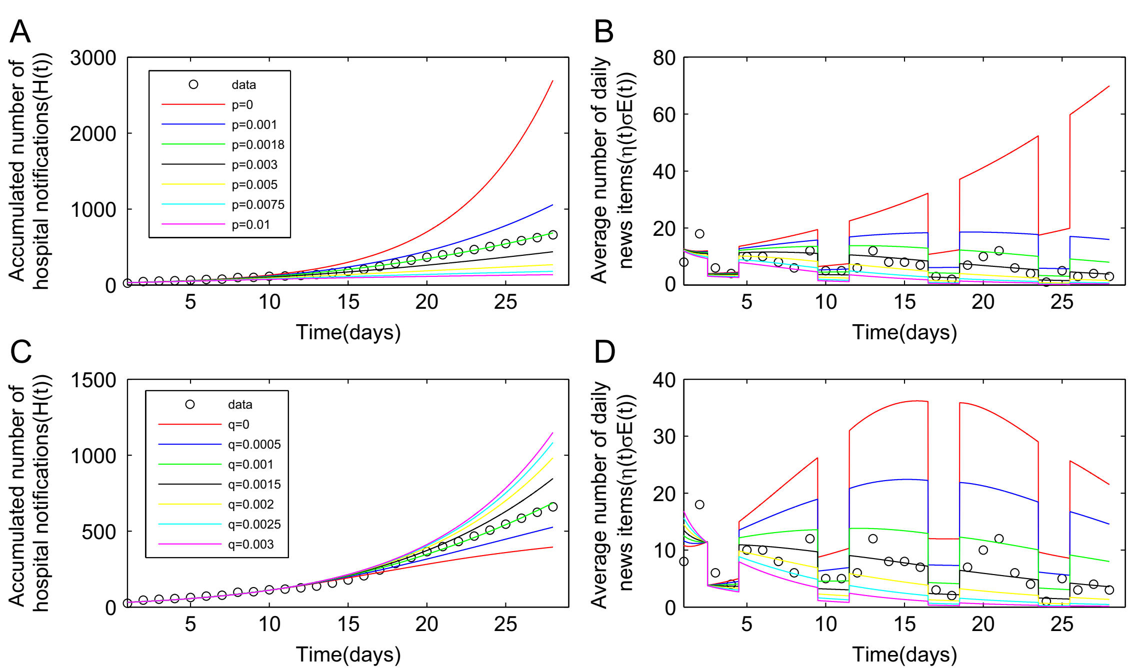


**Fig. 7.** The best-fit solution with the effects of media. (A) and (B) The effects of the weight of media effects sensitive to the accumulated number of hospital notifications $p$ on the accumulated number of hospital notifications and the average number of daily news items, respectively (the baseline $p=0.0018$). (C) and (D) The effects of the weight of hospital notifications effects sensitive to the number of daily news items $q$ on the accumulated number of hospital notifications and on the average number of daily news items, respectively (the baseline $q=0.001$).

were used for all parameters. Significance tests were done to assess if a PRCC was statistically significantly different from zero (see Appendix C). PRCC indices can be calculated for multiple time points and plotted versus time. This allows us to assess whether significance of one parameter occurs over an entire time interval during progression of the model dynamics. Fig. 8 shows the results of PRCCs plotted over the time interval (from 3rd to 30th on September 2009) for the above six parameters, respectively. Absolute values of PRCC greater than 0.064 (determined in Appendix C) indicated a statistically significant correlation between input parameters and output variables, with positive (negative) values of the PRCC meaning positive (negative) correlations between the input parameters and the output variable.

Fig. 8 (A) and (B) show the simulated accumulated number of hospital notifications and the average number of daily news items for the period (from 3rd to 30th, September 2009) corresponding to the LHS matrix and scheme defined by varying all input parameters. From Fig. 8(C) we can see the effects of changes in the above parameters on the simulated accumulated number of hospital notifications over time. It shows that the contact transmission rate $\beta$ and the transition rate $\sigma$ (weight of media effects sensitive to accumulated number of hospital notifications $p$) are positively (negatively) correlated with strong PRCCs until the time of the first peak. As the media spontaneous disappearance rate $\delta$ and the weight of hospital notifications effects sensitive to the average number of daily news items $q$ change, the PRCC values for them are not statistically significantly different from zero at the early time, then they become significantly correlated with high absolute PRCCs until the peak, and finally, are positively correlated to $H(t)$. The most significant parameter is the recovered rate $\gamma$, which is negatively correlated with strong PRCCs throughout the considered period.

The effects of the parameters ($\beta$, $\delta$, $\sigma$, $p$ and $\gamma$) on the simulated number of daily news items are similar to those on the simulated number of hospital notifications (see Fig. 8(D)). But as $q$ changes, the PRCC values are positively correlated and different from zero at the early time, then become not statistically significant, and finally, become negatively correlated with strong PRCCs.

In conclusion, we remark that the effectiveness of the media reports was the greatest when the weight of media effects sensitive to the accumulated number of hospital notifications increased, the rate of spontaneous disappearance of media decreased, the individuals' behavior altered. When these happened, the contact transmission rate and the transition rate, which are responsible for the lowering accumulated number of hospital notifications, decrease. We also note that the larger the weight of hospital notifications effects sensitive to the average number of daily news items is, the larger the effects of the disease on the media reports are. Also, the higher the recovered rate is, the smaller the accumulated number of hospital notifications is. Further, these results are consistent with the previous results about the effects of the media on the parameters $p, \delta, q$, which are important for how media influences can guide people's behavioral changes and, hence, critical for the control of the disease.

## 5. Discussion and conclusions

Mass media plays an increasing role in our daily life and can be used to inform the public during pandemics and epidemics. A comprehensive understanding of the effects of media during an epidemic or pandemic threat can aid in promoting public health communication strategies and disease mitigation measures. For public health promotion, it is important to recognize that media reports have effects on some important epidemic characteristics

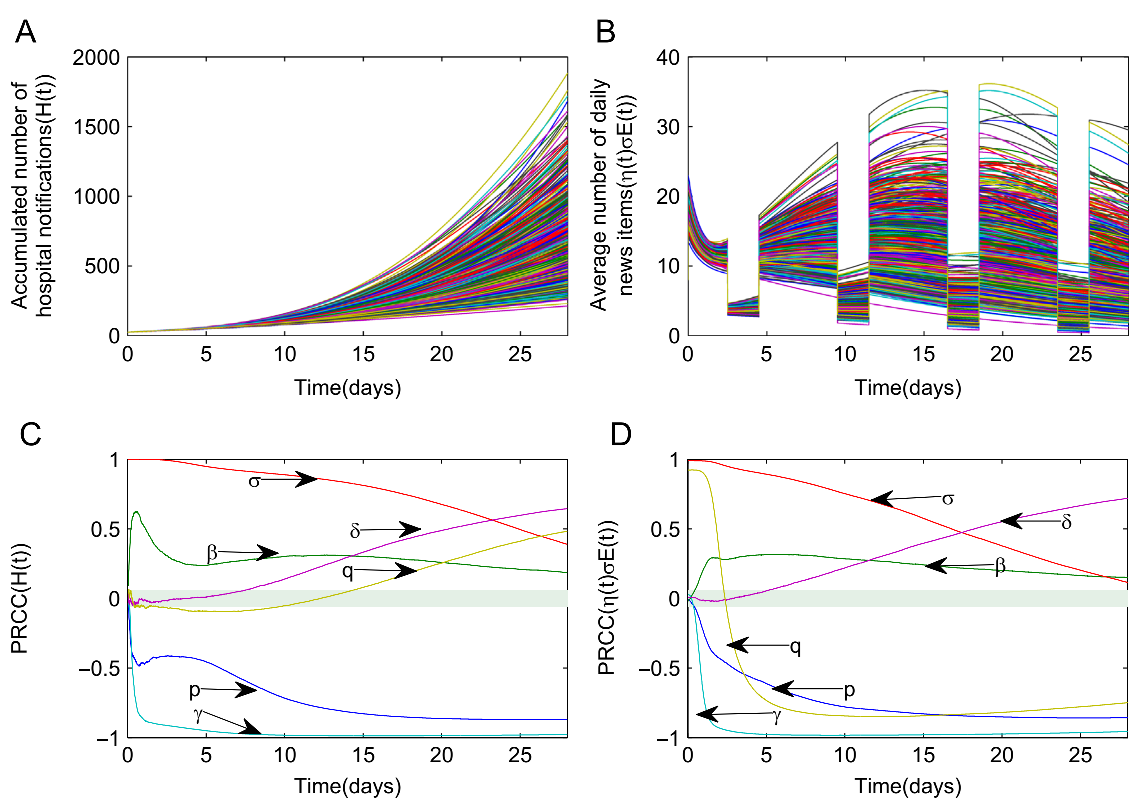


**Fig. 8.** Sensitivity analysis. (A) and (B) are plots of output (1000 runs) of the model (3.2) for the accumulated number of hospital notifications and the average number of daily news items. (C) and (D) are PRCCs of the model (3.2) for the accumulated number of hospital notifications and the average number of daily news items for parameters $p$, $\beta$, $\sigma$, $\gamma$, $\delta$, $q$. The gray region represents the confidence intervals which enclose only those PRCC values which are not significantly correlated.

such as the magnitude of the peak, its timing and the final level of infections, which need to be taken into account when planning interventions. Our analyses and model quantified how the number of reported news items and the number of hospital notifications affected each other and provided new insights into the progression of the 2009 A/H1N1 pandemic. In our research, cross-correlation analysis, media functions ($f(M,p)=\exp(-pM)$) and the dynamics of the number of reported news items were used to assess interaction effects between the number of daily news items and the accumulated number of hospital notifications. However, the time lags of the media reports have not been taken into consideration in our model, and we will address the effects of this important factor on the emerging infectious disease in near future.

We obtained a statistically significant correlation between the number of daily news items and the number of new hospital notifications by Pearson correlation analysis. Among the eight popular websites that we studied, some, but not all, had numbers of their news items that were statistically significantly highly or moderately correlated with the number of new hospital notifications. Further, by taking the CCF between them, we showed that there exists a feedback relationship ($k \neq 0$) and a contemporaneous relationship ($k=0$). We also demonstrated linear relationships for a range of lags, and the local maximum cross correlation coefficient exceeded 0.5 (moderately correlated or highly correlated) at specific lag $-3$, 4, 4 days between the number of daily news items from the websites 163.com, qq.com, CCTV.com and the number of new hospital notifications, respectively. In addition, the number of daily news items from news.cn caused the number of new hospital notifications and a linear relationship at lags ranging from 4 to 6 days.

We further noted that different individuals may only focus on the news from the websites which they usually access. Therefore, we averaged the number of daily news items for the above four websites and obtained a feedback relationship, a linear relationship at lags ranging from $-3$ to 1 days and 2–6 days and the local maximum cross correlation coefficient was 0.63 (moderately correlated) at only a specific lag of 4 days between the average number of daily news items and the number of new hospital notifications.

From the literature (Tang et al., 2010) and the data set shown in Fig. 1(A), the first 28 days of average number of daily news items and the number of new hospital notifications were correlated and the results suggested that there were lags in the number of new hospital notifications, which were possibly caused by the average number of daily news items. Also, a contemporaneous relationship (for $k=0$) and a linear relationship between them at lags ranging from 0 to 2 days were found.

Although Pearson correlation analysis and cross-correlation analysis helped us to confirm that the average number of daily news items and the number of new hospital notifications affect each other in the long run, it is difficult to quantify how they influence each other dynamically. Therefore, we further proposed a novel mathematical model, to study the effects of mass media on epidemics by employing a media function in the transmission terms and the dynamics of the numbers of news reports embedded into the classical infectious disease SEIR model.

By fitting the accumulated number of hospital notifications during the 2009 pandemic of A/H1N1 influenza in Shaanxi Province and the average number of daily news items to our proposed model, we were able to obtain reasonable estimates of the unknown parameter values (in Table 2) and the mean basic reproduction number was estimated as 1.794 (95% CI 1.3858–1.9091). The latter and its confidence interval obtained here based on the model (3.2) are about the same as those estimated by Tang et al. (2010), with small differences probably associated with differences in methodology. Besides, many factors such as differences in population size, seasonal factors, realization of control measures and mobility of the population among regions have led to a wide range of reproduction number estimates: Mexico (1.58, 95% CI 1.34–2.04, Fraser et al., 2009), the United States of American (95% CI 1.5–3.2, White et al., 2009), Thailand (2.07, 95% CI 1.92–2.22, De Silva et al., 2009), Japan (2.3, 95% CI 2.0–2.6, Nishiura et al., 2009).

Coverage duration $T_2$, the responses of mass media to the accumulated number of hospital notifications during weekdays $\eta_1$, the media spontaneous disappearance rate $\delta$, the weight of media effects sensitive to the accumulated number of hospital notifications $p$ and the weight of hospital notifications effects sensitive to the average number of daily news items $q$ all markedly affect the outbreak of A/H1N1 influenza. In particular, the estimated accumulated number of hospital notifications by the 30th September 2009 can be controlled at around 684, and decreased by 54% if coverage duration $T_2$ increased to 10 days (i.e. mass media persistent reports for a long time); or increased by 125% if response of mass media on the accumulated number of hospital notifications during weekdays $\eta_1$ decreased to 0.5 (i.e the mass media shrink the number of infections); or increased by 293% if the weight of media effects sensitive to the accumulated number of hospital notifications $p$ decreased to 0 (i.e. the mass media have no effects on the number of infections).

The analysis based on the PRCCs, identifying the key input variables ($p$, $\beta$, $\sigma$, $\gamma$, $\delta$, $q$) that contribute to the infection outcome (accumulated number of hospital notifications and average number of daily news items), strongly support the implementation of an integrated strategy for different mitigation measures, including reporting news in a timely manner and increasing the duration of highlighted mass media reports, to curb the outbreak during different phases of the epidemic. The PRCCs showed that the most important parameters correlated to the disease outbreak and control were $p, \beta, \sigma, \gamma, \delta, q$ during the exponential growth phase of the first wave. Most important were the strength of the effects of media reports (i.e. increase in $p$ and decrease in $\delta$), which led to the behavior changes ($\beta$, $\sigma$ decreased) responsible for lowering the number of hospital notifications and critical for the control of the disease.

In conclusion, this study presents a novel methodology through using cross-correlation analysis and embedding a media function and the dynamics of the number of news reports into the classical SEIR model, showed that combining statistical analysis with a mathematical model is beneficial for analyzing media impacts. It demonstrated that the media reports affect the accumulated number of hospital notifications by reducing the transmission rate and the rate of the spontaneous disappearance media reports, and increasing coverage duration of the media. All these results confirmed the importance of the responses of individuals to the media reports, with behavior changes being more important in emerging infectious disease control than the substantial media attention given to unexpected events and reports. Therefore, for mitigating emerging infectious diseases, the media publicity should be focused on how to guide people's behavioral changes which are critical for the control of the disease.

## Acknowledgements

This work is supported by the National Natural Science Foundation of China (NSFCs 11171199, 11471201), and by the Fundamental Research Funds for the Central Universities (GK201003001, GK201401004).

## Appendix A. The details of CCF

*Estimation of the cross-correlation function* (*CCF*): Cross correlation is the most commonly used method to detect causal interactions between two simultaneously recorded, stochastic processes $\{x(t)\}$ and $\{y(t)\}$, $t=0, \pm 1, \pm 2, \ldots$ . We assume without loss of generality that they are jointly stationary. The cross correlation is defined as the joint second-order moment (El-Gohary and McNames, 2007; William, 2006; Zhao et al., 2011)

$$\gamma_{xy}(k)=E(x_t-\mu_x)(y_{t+k}-\mu_y), \quad k=0, \pm 1, \pm 2, \ldots \tag{A.1}$$

where $k$ is the lag, $\mu_x$, $\mu_y$ and $\sigma_x$, $\sigma_y$ are the means and standard deviations of $x_t$ and $y_t$, respectively, and $E[\cdot]$ denotes the expected value. This is often normalized to a range from $-1$ to 1 and expressed as the (CCF)

$$\rho_{xy}(k)=\frac{\gamma_{xy}(k)}{\sigma_x\sigma_y}, \quad k=0, \pm 1, \pm 2, \ldots \tag{A.2}$$

A very common estimator of the CCF is by the sample CCF,

$$\hat{\rho}_{xy}(k)=\frac{\hat{\gamma}_{xy}(k)}{S_xS_y}, \quad k=0, \pm 1, \pm 2, \ldots \tag{A.3}$$

where $S_x$, $S_y$ are the sample variances of the two sets of observed data, and $S_x=\sqrt{\hat{\gamma}_{xx}(0)}, S_y=\sqrt{\hat{\gamma}_{yy}(0)}$.

$$\hat{\gamma}_{xy}(k)=\begin{cases}\dfrac{1}{n}\displaystyle\sum_{t=1}^{n-k}(x_t-\bar{x})(y_{t+k}-\bar{y}), & k\geq 0,\\ \dfrac{1}{n}\displaystyle\sum_{t=1-k}^{n}(x_t-\bar{x})(y_{t+k}-\bar{y}), & k\leq 0,\end{cases} \tag{A.4}$$

where $\bar{x}$, $\bar{y}$ are the samples of the two sets of observed data. If the two processes are independent or uncorrelated, their CCF ($\rho_{xy}(k)$) is zero for all lags. If the CCF $\rho_{xy}(k)$ is statistically significant and $\hat{\rho}_{xy}(k)$ is the local maximum at only a specific lag $k=d$, this implies that there is a pure delay of $d$ samples and a linear association between the two processes.

*Statistical significance*: The CCF is generally asymmetric and measures both the strength and direction of an association. In order to test if a CCF $\rho_{xy}(k)$ value is significantly different from zero, Bartlett proved that the CCF $\hat{\rho}_{xy}(k)$ is approximately normally distributed with zero mean and a variance of $1/\sqrt{n-k}$. So we obtain the following statistic:

$$U=\frac{\hat{\rho}_{xy}(k)}{\sqrt{\dfrac{1}{n-k}}}\sim N(0,1), \tag{A.5}$$

where $n$ is the number of observed samples. We used this statistic to calculate a 95% critical region which would encompass the estimated cross correlation of uncorrelated processes over 95% of the lags. If the estimated CCF $\hat{\rho}_{xy}(k)$ markedly differs from the critical region at a lag $d$, we interpret the value as being significantly different from zero (i.e. $\rho_{xy}(k)\neq 0$), otherwise $\rho_{xy}(k)=0$

*Detection criterion of causal interactions*: To see the complete picture of the relationship between two time series $x_t$ and $y_t$ , it is important to examine $\rho_{xy}(k)$ for both positive lags $k>0$ and negative lags $k<0$ (William, 2006).

(i) Series $x_t$ is said to cause series $y_t$, if $\rho_{xy}(k)\neq 0$ for some $k>0$ and $\rho_{xy}(k)=0$ for all $k<0$.

(ii) Series $y_t$ is said to cause series $x_t$, if $\rho_{xy}(k)\neq 0$ for some $k<0$ and $\rho_{xy}(k)=0$ for all $k>0$.

(iii) Two series $x_t$ and $y_t$ are said to have a contemporaneous relationship if $\rho_{xy}(k)\neq 0$. Particularly, they are said to be related only contemporaneously if $\rho_{xy}(k)\neq 0$, and $\rho_{xy}(k)=0$ for all $k\neq 0$.

(iv) A feedback relationship between $x_t$ and $y_t$ exists, if $\rho_{xy}(k)\neq 0$ for some $k>0$ as well as for some $k<0$.

## Appendix B. The calculation of $R_0$

We simplify the model (3.2) as follows:

$$\begin{cases}\dfrac{dE(t)}{dt}=f(M,p)\beta S(t)I(t)/N-\sigma E(t),\\ \dfrac{dI(t)}{dt}=\sigma E(t)-\gamma I(t),\\ \dfrac{dM(t)}{dt}=\eta(t)\sigma E(t)-\delta M(t),\\ \dfrac{dS(t)}{dt}=-f(M,p)\beta S(t)I(t)/N.\end{cases} \tag{B.1}$$

Let the right-hand side of (B.1) be zero, then the model (B.1) has one disease-free equilibrium (DFE) at $E_0=(0,0,0,S_0)$. Using the notation in Wang and Zhao (2008) and Becar et al. (2014), we have two vectors $\mathcal{F}$, $\mathcal{V}$ to represent the new infection term and the remaining transfer terms, respectively:

$$\mathcal{F}=\begin{pmatrix}\exp(-pM)\beta SI/N\\ 0\\ 0\\ 0\end{pmatrix}, \quad \mathcal{V}=\begin{pmatrix}\sigma E\\ -\sigma E+\gamma I\\ -\eta(t)\sigma E+\delta M\\ \exp(-pM)\beta SI/N\end{pmatrix} \tag{B.2}$$

The infected compartments are $E$ to $I$, hence, a straightforward calculation gives

$$\mathbf{F}=\begin{pmatrix}0&\beta&0\\0&0&0\\0&0&0\end{pmatrix}, \quad \mathbf{V}=\begin{pmatrix}\sigma&0&0\\-\sigma&\gamma&0\\-\eta(t)\sigma&0&\delta\end{pmatrix} \tag{B.3}$$

where $\mathbf{F}$ is non-negative and $\mathbf{V}$ is a non-singular M-matrix. Based on the definition of the spectral radius, therefore, we can define the spectral radius of the infection operator $\mathbf{L}$ as the basic reproduction number (Wang and Zhao, 2008)

$$\mathbf{R}_0:=\rho(\mathbf{L}) \tag{B.4}$$

for the periodic epidemic model (B.1). Generally, we compute $R_0$ for the following linear $\omega$-periodic equation (Becar et al., 2014):

$$\frac{d\omega(t)}{dt}=\left[-\mathbf{V}(t)+\frac{\mathbf{F}(\mathbf{t})}{\lambda}\right]\omega, \quad \omega\in\mathcal{R}^3,\ t\in\mathcal{R} \tag{B.5}$$

with parameters $\lambda\in(0,+\infty)$. Let $W(t,s,\lambda)$ be the evolution operator of the system on $\mathcal{R}^m$. Then the positive root of $\rho(W(\omega,0,\lambda))=1$ is the value of $R_0$, so

$$R_0=\frac{\beta}{\gamma}. \tag{B.6}$$

We use the 500 groups of values of $\beta$ and $\gamma$ (obtained in Section 3.3), and the mean reproduction number is estimated as 1.794 (95% CI 1.3858–1.9091) based on the model (3.2).

## Appendix C. Significance tests of PRCC

Significance tests have be done to assess if a PRCC is significantly different from zero. Each PRCC ($r$) generates a value $T$ according to the following statistic (Marino et al., 2008):

$$T=\hat{r}\sqrt{\frac{N-2-p}{1-\hat{r}^2}}\sim t(N-2-p), \tag{C1}$$

where $T$ follows a student's $t$ distribution with $(N-2-p)$ degrees of freedom. $N$ is the sample size and $p$ is the number of input parameters.

## References


Becar, R., Gonzalez, P.A., Vasquez, Y., 2014. Fermionic greybody factors of two and five-dimensional dilatonic black holes. Gen. Relativ. Quant. Cosmol. 2, 1–11.

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

Cui, J.G., Sun, Y.H., Zhu, H.P., 2008. The impact of media on the control of infectious diseases. J. Dyn. Differ. Equ. 20, 31–53.

Collinson, S., Heffernan, J.M., 2014. Modeling the effects of media during an influenza epidemic. BMC Public Health 14, 1–10.

De Silva, U.C., Warachit, J., Waicharoen, S., Chittaganpitch, M., 2009. A preliminary analysis of the epidemiology of influenza A(H1N1) virus infection in Thailand from early outbreak data, June–July 2009. Euresurveillance 14, 1–3.

El-Gohary, M., McNames, J., 2007. Establishing causality with whitened cross-correlation analysis. IEEE Trans. Biomed. Eng. 54, 2214–2222.

Fraser, C., Donnelly, C.A., Cauchemez, S., et al., 2009. Pandemic potential of a strain of influenza A (H1N1): early findings. Science 324, 1557–1561.

Huang, F., Zhou, S.S., Zhang, S.S., et al., 2011. Temporal correlation analysis between malaria and meteorological factors in Motuo County. Tibet Malar. J. 10, 1–8.

Haugh, L.D., 1976. Checking the independence of two covariance-stationary time series: a univariate residual cross-correlation approach. J. Am. Stat. Assoc. 71, 378–385.

Koch, P.D., Yang, S.-S., 1986. A method for testing the independence of two time series that accounts for a potential pattern in the cross-correlation function. J. Am. Stat. Assoc. 81, 533–544.

Li, Y.F., Cui, J.G., 2009. The effect of constant and pulse vaccination on SIS epidemic models incorporation media coverage. Commun. Nonlinear Sci. Numer. Simul. 14, 2353–2365.

Liu, R.S., Wu, J.H., Zhu, H.P., 2007. Media/psychological impact on multiple outbreaks of emerging infectious diseases. Comput. Math. Methods Med. 8, 153–164.

Marino, S., Hogue, I.B., Ray, C.J., Kirschner, D.E., 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J. Theor. Biol. 254, 178–196.

Mckay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21, 239–245.

Nishiura, H., Wilson, N., Baker, M.G., 2009. Estimating the reproduction number of the novel influenza A virus (H1N1) in a Southern Hemisphere setting: preliminary estimate in New Zealand. N. Z. Med. J. 122, 73–77.

Sampei, Y.K., Midori, A.-U., 2009. Mass-media coverage, its influence on public awareness of climate-change issues, and implications for Japans national campaign to reduce greenhouse gas emissions. Global Environ. Change 19, 203–212.

Sun, C.J., Yang, W., Arino, J., Khan, K., 2011. Effect of media-induced social distancing on disease transmission in a two patch setting. Math. Biosci. 230, 87–95.

Taylor, R., 1990. Interpretation of the correlation coefficient: a basic review. J. Diag. Med. Sonogr. 1, 35–39.

Tang, S.Y., Xiao, Y.N., Yang, Y.P., et al., 2010. Community-based measures for mitigating the 2009 H1N1 pandemic in China. PLoS ONE 5, 1–11.

Tang, S.Y., Xiao, Y.N., Yuan, L., et al., 2012. Campus quarantine (Fengxiao) for curbing emergent infectious diseases: lessons from mitigating A/H1N1 in Xi'an, China. J. Theor. Biol. 295, 47–58.

Tchuenche, J.M., Bauch, C.T., 2012. Dynamics of an infectious disease where media coverage influences transmission. ISRN Biomath. 2012, 1–10.

Wang, W.D., Zhao, X.Q., 2008. Threshold dynamics for compartmental epidemic models in periodic environments. J. Dyn. Differ. Equ. 20, 699–717.

White, L.F., et al., 2009. Estimated of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the USA. Influenza Other Respir. Viruses 3, 267–276.

William, W.S., 2006. Time Series Analysis: Univariate and Multivariate Methods. Transfer Function Models, 2nd ed. Pearson Education Inc., New York, pp. 345–360.

Xiao, Y.N., Zhao, T.T., Tang, S.Y., 2013. Dynamic of an infectious diseases with media/psychology induced non-smooth incidence. Math. Biosci. Eng. 10, 445–461.

Xiao, Y.N., Tang, S.Y., Wu, J.H., 2015. Media impact switching surface during an infectious disease outbreak. Sci. Rep. 5, 7838. http://dx.doi.org/10.1038/srep07838.

Xie, T.S., Yang, Z.X., Yang, S.H., et al., 2014. Correlation between reported human infection with avian influence A H7N9 virus and cyber user awareness: what can we learn from digital epidemiology?. Int. J. Infect. Dis. 22, 1–3.

Zhao, Y., Billings, S.A., Wei, H.L., Sarrigiannis, P.G., 2011. Detecting and tracking time-varying causality with applications to EEG data. Res. Rep. 1027, 1–15.