

# Predicting the outbreak of hand, foot, and mouth disease in Nanjing, China: a time-series model based on weather variability

Sijun Liu<sup>1,2,3</sup> · Jiaping Chen<sup>2</sup> · Jianming Wang<sup>1,2</sup> · Zhuchao Wu<sup>1</sup> · Weihua Wu<sup>4</sup> · Zhiwei Xu<sup>3,5</sup> · Wenbiao Hu<sup>3,5</sup> · Fei Xu<sup>2,6</sup> · Shilu Tong<sup>3,5</sup> · Hongbing Shen<sup>2</sup>

Received: 14 September 2016 / Revised: 17 October 2017 / Accepted: 20 October 2017 / Published online: 30 October 2017  
© ISB 2017

**Abstract** Hand, foot, and mouth disease (HFMD) is a significant public health issue in China and an accurate prediction of epidemic can improve the effectiveness of HFMD control. This study aims to develop a weather-based forecasting model for HFMD using the information on climatic variables and HFMD surveillance in Nanjing, China. Daily data on HFMD cases and meteorological variables between 2010 and 2015 were acquired from the Nanjing Center for Disease Control and Prevention, and China Meteorological Data Sharing Service System, respectively. A multivariate seasonal autoregressive integrated moving average

(SARIMA) model was developed and validated by dividing HFMD infection data into two datasets: the data from 2010 to 2013 were used to construct a model and those from 2014 to 2015 were used to validate it. Moreover, we used weekly prediction for the data between 1 January 2014 and 31 December 2015 and leave-1-week-out prediction was used to validate the performance of model prediction. SARIMA (2,0,0)<sub>S2</sub> associated with the average temperature at lag of 1 week appeared to be the best model ( $R^2 = 0.936$ , BIC = 8.465), which also showed non-significant autocorrelations in the residuals of the model. In the validation of the constructed model, the predicted values matched the observed values reasonably well between 2014 and 2015. There was a high agreement rate between the predicted values and the observed values (sensitivity 80%, specificity 96.63%). This study suggests that the SARIMA model with average temperature could be used as an important tool for early detection and prediction of HFMD outbreaks in Nanjing, China.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00484-017-1465-3>) contains supplementary material, which is available to authorized users.

✉ Wenbiao Hu  
w2.hu@qut.edu.au

✉ Fei Xu  
frankxufei@163.com

<sup>1</sup> Department of Social Medicine and Health Education, School of Public Health, Nanjing Medical University, Nanjing 211166, China

<sup>2</sup> Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing 211166, China

<sup>3</sup> School of Public Health and Social Work, Queensland University of Technology, Kelvin Grove, Brisbane, Queensland 4059, Australia

<sup>4</sup> Department of Infectious Diseases, Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing Medical University, Nanjing 210008, China

<sup>5</sup> Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland 4059, Australia

<sup>6</sup> Department of Non-communicable Disease Prevention, Nanjing Municipal Center for Disease Control and Prevention, Nanjing 210003, China

**Keywords** SARIMA · Forecasting · Hand, foot and mouth disease (HFMD) · Temperature · Infectious disease

## Introduction

Hand, foot, and mouth disease (HFMD) is an infectious disease caused by various enteroviruses, with coxsackievirus A16 (CV-A16) and enterovirus 71 (EV71) being the most commonly reported (Puenpa et al. 2011; Wu et al. 2010). HFMD is transmitted through close personal contact, exposure to feces, contaminated objects, and surfaces of an infected person (Wei et al. 2015). It usually affects infants and children under 5 years old (Qian et al. 2016), and most patients experience self-limiting illness typically including fever, mouth ulcers, and vesicles on the hands, feet, and mouth (Jiang

et al. 2012). However, some patients will rapidly develop neurological and cardiopulmonary symptoms that can be fatal, particularly when the cases are associated with EV71 (Ma et al. 2010).

Over the past decades, HFMD epidemics are widespread in Asian countries (Ho et al. 1999; Chan et al. 2000; Chan et al. 2003; Tu et al. 2007; Chua and Kasri 2011), with deaths predominantly among children. Numerous outbreaks of HFMD have recently been reported in mainland China (Zhang et al. 2010; De et al. 2011; Liu et al. 2011) and HFMD has become one of the leading causes of children's mortality in China (Chen et al. 2015a). In 2008, a large-scale outbreak of HFMD occurred in Fuyang city of Anhui province (Zhang et al. 2010), resulting 353 severe cases and 22 deaths, and then rapidly spread throughout the mainland China. Since 2008, the Chinese Ministry of Health has listed HFMD as a class C communicable disease and cases must be reported to the National Disease Surveillance Reporting and Management System within 24 h of diagnosis. According to this system, there were more than 7.2 million cases of HFMD during 2008–2012 in China, and the disease causes 500–900 deaths each year (Xing et al. 2014). However, there is no vaccine or specific curative treatment for HFMD at present. Thus, it is important to develop an early warning system of HFMD outbreaks which could help to target timely interventions to control and prevent the disease.

Many statistical models have been used for identifying and predicting disease outbreaks, including time-series forecasting models such as seasonal autoregressive integrated moving average (SARIMA) model. This model has been widely used to predict the incidences of various infectious diseases, such as dengue fever (Shi et al. 2016), malaria (Midekisa et al. 2012), and hepatitis E (Ren et al. 2013). Such predictions are helpful for mobilizing public health resources to prevent and control these diseases. A previous study indicated that targeting preventive measures for HFMD can be developed by using SARIMA model since it fits and predicts the changing trends of HFMD and achieves good results (Liu et al. 2016). The other study reported the climate-based SARIMA model is useful for the implementation of routine surveillance of HFMD and evaluation of new intervention strategies (Feng et al. 2014). Many studies have demonstrated that climatic parameters are among the most significant environmental factors associated with the incidence of HFMD (Deng et al. 2013; Xing et al. 2014; Chen et al. 2015b). The relationship between temperature and HFMD has been well documented. Studies have suggested that increased temperature has been associated with the occurrence of HFMD in Japan (Onozuka and Hashizume 2011), China (Huang et al. 2013), and Singapore (Hii et al. 2011), and non-linear dose-response curves with temperature have also been noted (Xiao et al. 2017). Besides temperature, there are positive associations between HFMD infection and relative humidity and precipitation (Wang et al.

2016). In addition, evidence has also shown that high wind speed was a risk factor for HFMD (Ma et al. 2010). Moreover, a study found air pressure was negatively linked with HFMD occurrence (Li et al. 2014). The impact of temperature and air pressure on the incidence of HFMD might contribute to changes of survival and reproduction of pathogens and influence people's immune function and activity. However, emergence of HFMD infection due to temperature variability in China has been reported (Cheng et al. 2016; Xu et al. 2016) but few studies have integrated HFMD surveillance and temperature with SARIMA model to provide early signals for the prediction of this disease.

Nanjing is the capital of Jiangsu province and the larger mega city in eastern China. The HFMD incidence in the whole population in Nanjing (1.6 cases per 1000 persons) was higher than the national average in China (1.4 cases per 1000 persons) in 2010 and 2.0 times higher than that reported in Japan in 2010 (0.8 cases per 1000 persons) (Guo et al. 2016). Liu et al. (2015a) analyzed spatiotemporal dynamics of HFMD in Jiangsu Province, China, and their findings indicated that Nanjing was the main epidemic area in Jiangsu province which should strengthen surveillance and public health interventions to control and prevent this disease. However, no study has focused on the prediction of HFMD epidemics in Nanjing city of eastern China. Hence, in this study, we developed a weather-based forecasting model for HFMD using SARIMA model coupled with the information on climatic variables and disease surveillance in Nanjing, China.

## Material and methods

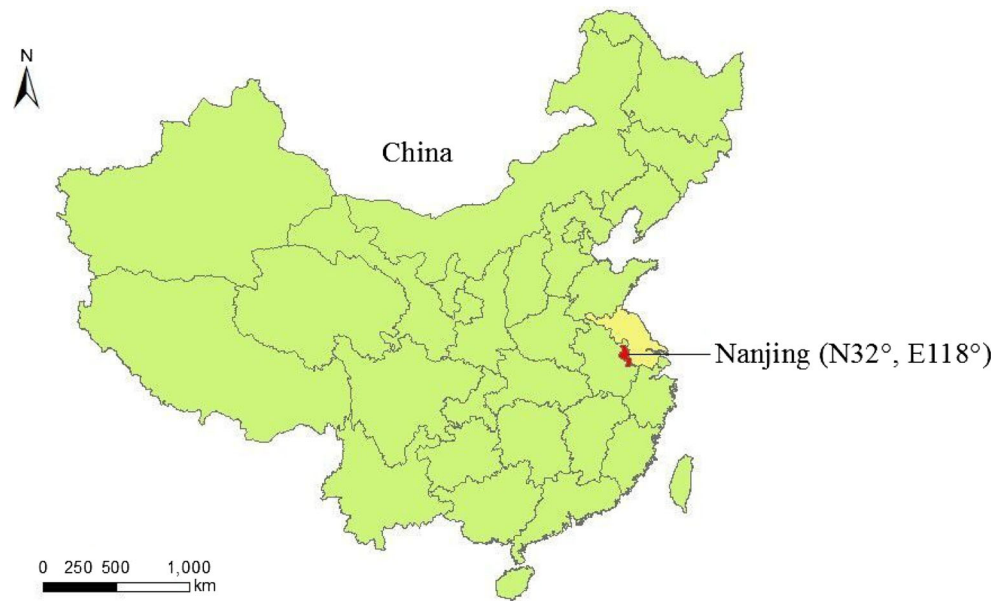
### Study site

Figure 1 shows the geographic location of Nanjing in eastern China. Nanjing is an important central city in the economic core area of the Yangtze River Delta and also the transportation and communication center of eastern China. It has a total population of 8.2 million and an area of 6587.02 km<sup>2</sup>. Nanjing is situated at 32° north latitude and 118° east longitude. The city of Nanjing has a subtropical climate characterized by a hot, humid summer and a cold, dry winter. Nanjing has four distinctive seasons: spring (1 April–7 June), summer (8 June–17 September), autumn (18 September–11 November), and winter (12 November–31 March). The annual average temperature ranges from 13.0 to 20.8 °C. The average annual rainfall is 1091.1 mm, and the total sunshine is 1863.8 h.

### Data collection

Daily reported cases of HFMD from January 2010 to December 2015 were obtained from the Nanjing Centre for Disease Control and Prevention. According to China's

**Fig. 1** Geographical location of Nanjing, Jiangsu Province of China



notifiable infectious disease regulations, all HFMD cases are required to be reported online to the infectious disease surveillance system within 24 h of diagnosis, including demographic information (name, sex, date of birth, and address); case classification (probable or confirmed); date of symptom onset; date of diagnosis; date of death (if applicable); and virus type (EV71, CV-A16, or other enterovirus) for confirmed cases (Xing et al. 2014). The clinical criteria for diagnosis of HFMD were provided in a guidebook published by the Chinese Ministry of Health in 2010. Patients with the following clinical symptoms are defined as having HFMD: papular or vesicular rash on hands, feet, mouth, or buttocks, with or without fever. A laboratory-confirmed case was defined as a clinical case with laboratory evidence of enterovirus infection completed by RT-PCR, real-time PCR, or virus isolation. The HFMD cases include clinical-confirmed cases and laboratory-confirmed cases. We also carefully conducted data checking and cleaning, including deleting suspected cases and cases from other areas.

Meteorological data of daily average temperature (°C), relative humidity (%), wind velocity (m/s), duration of sunshine (h), total rainfall (mm), and air pressure (hPa) from January 2010 to December 2015 were downloaded from China Meteorological Data Sharing Service System (<http://data.cma.gov.cn>). We aggregated the daily data and obtained average weekly data for temperature, relative humidity, wind velocity, duration of sunshine, and air pressure. Total weekly rainfall was aggregated using daily data.

### Data analysis

Data analysis included two stages: descriptive and time-series analyses. Daily counts of reported cases and daily

meteorological variable data were used to produce weekly average values because there were zeros if daily counts of HFMD cases were used. Descriptive analysis was used to illustrate the distribution and correlation of meteorological variables and HFMD cases. We estimated the correlations between meteorological variables and HFMD at lag 1 week because this infection typically has an incubation period of 3–7 days (Wong et al. 2010).

### Time-series analysis

We developed the multivariate SARIMA model to predict the incidence of HFMD by using the weekly numbers of HFMD cases as the dependent variable and climatic variables as the independent variables. To facilitate the selection of climatic variables to be used as external repressors, we computed Pearson's correlation coefficient between numbers of HFMD cases and climatic variables. Average temperature and air pressure that were highly correlated with HFMD cases were included in the model ( $r = 0.413$  and  $-0.460$  respectively,  $P < 0.01$ ). We then computed cross-autocorrelation analysis to assess associations between HFMD cases and average temperature and air pressure over a range of time lags. Lagged climatic variables found to be significantly associated with HFMD cases were tested as external predictors in multivariate SARIMA model. To control for the impact of seasonality, we decomposed the HFMD incidence into three series. That is  $Y_t = T_t + S_t + E_t$ , where  $Y_t$  denotes the HFMD incidence,  $T_t$  denotes the trend component,  $S_t$  denotes the seasonal component, and  $E_t$  denotes the residual component. To control for the impact of seasonality in SARIMA model, we input the  $S_t$  into the SARIMA model as a seasonal factor. In general, three important parts were included in the SARIMA model:

autoregressive (AR), data of present and past time points to be included in the model; differencing, transforming time series from non-stationary to stationary; moving average (MA), errors of present and past time points to be included in the model. The SARIMA model is typically represented by  $(p, d, q) \times (P, D, Q)_s$  (Box and Jenkins 1970);  $p, d, q$  are the orders of the AR, the differencing, and MA process, respectively, whereas  $P, D, Q$  are the seasonal orders of AR, differencing, and MA process, respectively;  $s$  is the seasonal period ( $s = 52$  weeks in this study).

We used the Box-Jenkins approach to evaluate the SARIMA model, which consists of a four-step process. First, it is necessary to identify the stationary state of data in series. Natural logarithm transformation was performed to stabilize the variance of the time-series data. Second, the orders of the model were identified initially by the analysis of autocorrelation function (ACF) and partial autocorrelation function (PACF). Parameters of the model were estimated by using the maximum likelihood method. Third, the goodness-of-fit of the models was examined through Ljung-Box test for white noise in residuals. The Bayesian Information Criteria (BIC) and the stationary  $R$  square ( $R^2$ ) were also conducted to compare the goodness-of-fit of SARIMA models. A model with the highest  $R^2$  and the lowest BIC was considered as the most appropriate model to fit the data. Finally, the model was validated through the validation process. The data file was divided into two datasets: the data from 1 January 2010 to 31 December 2013 (208 weeks in total) were used to construct a SARIMA model and those between 1 January 2014 and 31 December 2015 (104 weeks in total) were used to validate the model. The predictive validity of the models was evaluated by using the mean absolute percentage error (MAPE). The smaller the MAPE, the better the model is for forecasting. There were two approaches to estimate the predicted values which were called the 104-steps-ahead prediction and 1-step-ahead prediction (Luz et al. 2008). The 104-steps-ahead prediction used the fitted SARIMA model to predict the consecutive 104 observations that constitute the year 2014 and 2015. The 1-step-ahead prediction used the SARIMA model to predict 1 week ahead, which means the next week predicted value (the 1st week 2014). Then, as the observed value for the 1st week 2014 was obtained, we updated the data to 1st week 2014, re-estimated the parameters of the SARIMA model, and computed the next 1-step-ahead predicted value, the 2nd week 2014. This process was continued until the 52nd week of 2015. Both approaches were used to compare in this study.

### Sensitivity analysis

Sensitivity analysis was also conducted in the validation process. We defined an outbreak where the HFMD incidence rate exceeded the third quartile (Q3) of the incidence rates plus one standard deviation (SD). That is  $\geq Q3 + 1$  SD were set up on

outbreak of HFMD in Nanjing. To further test the model utility and predictive validity, mean + 1 SD was also used to compare the predictive results. We used the sensitivity, specificity, and overall agreement rates to compare the model predictive power (Table S1). Finally, Q3 + 1 SD definition appears better predictive capability and was thus selected as the outbreak threshold of HFMD in this study.

Descriptive analysis and SARIMA modeling were performed using IBM SPSS version 23 (SPSS Inc.; Chicago, IL, USA) with a significance level of  $P$  value  $< 0.05$  (two sides).

## Results

### Descriptive analysis

There were totally 102,751 HFMD cases in Nanjing between 1 January 2010 and 31 December 2015. Table 1 shows the summary statistics of weekly data of HFMD and the meteorological variables during the study period in Nanjing. The weekly average number of HFMD was 328.3 (range, 7–1093). The weekly average of temperature, relative humidity, sunshine, wind velocity, and air pressure were 16.4 °C (range,  $-0.9$ –33.3 °C), 70.1% (range, 42–91%), 5.3 h (range, 0–12.2 h), 2.8 m/s (range, 1.5–5.8 m/s), and 1012.4 hPa (range, 996.4–1029.0 hPa), respectively. The weekly total rainfall was 22.5 mm (range, 0–326.3 mm).

Table 2 reveals the Spearman correlations between weekly meteorological variables at lag of 1 week and HFMD cases. We found that the HFMD incidence was significantly associated with temperature, rainfall, and air pressure ( $r = 0.413, 0.155, \text{ and } -0.460$  respectively,  $P < 0.01$ ), but not significantly with relative humidity, sunshine, and wind velocity ( $P > 0.05$ ).

Figure 2 illustrates the weekly incidences of HFMD and weekly average temperature during the study period. Figure 3 shows seasonal distribution of weekly number of HFMD cases, indicating a seasonal pattern. The peaks reached a

**Table 1** Descriptive statistics of weekly data on meteorological variables and HFMD cases in Nanjing, 2010–2015

Variables	<i>n</i>	Mean	SD	Min.	Median	Max.
Weekly HFMD cases	313	328.3	252.2	7	235	1093
Meteorological variables						
Temperature (°C)	313	16.4	9.1	-0.9	17.8	33.3
Relative humidity (%)	313	70.1	10.0	42.0	70.7	91.0
Sunshine (hour)	313	5.3	2.3	0	5.3	12.2
Wind velocity (m/s)	313	2.8	0.6	1.5	2.8	5.8
Rainfall (mm)	313	22.5	37.3	0	9.1	326.3
Pressure (hPa)	313	1012.4	8.6	996.4	1013.3	1029.0

**Table 2** Spearman correlations between weekly meteorological variables at lag of 1 week and HFMD cases in Nanjing, 2010–2015

	Pressure	Relative humidity	Rainfall	Wind velocity	Sunshine	Temperature
Relative humidity	− 0.275**					
Rainfall	− 0.364**	0.668**				
Wind velocity	− 0.134*	0.048	− 0.215**			
Sunshine	− 0.195**	− 0.649**	0.476**	− 0.082		
Temperature	− 0.926**	0.257**	0.277**	0.032	0.293**	
HFMD cases	− 0.460**	0.087	0.155**	0.101	0.058	0.413**

\* $P < 0.05$ ; \*\* $P < 0.01$

maximum during April to July and again spiked in October to December.

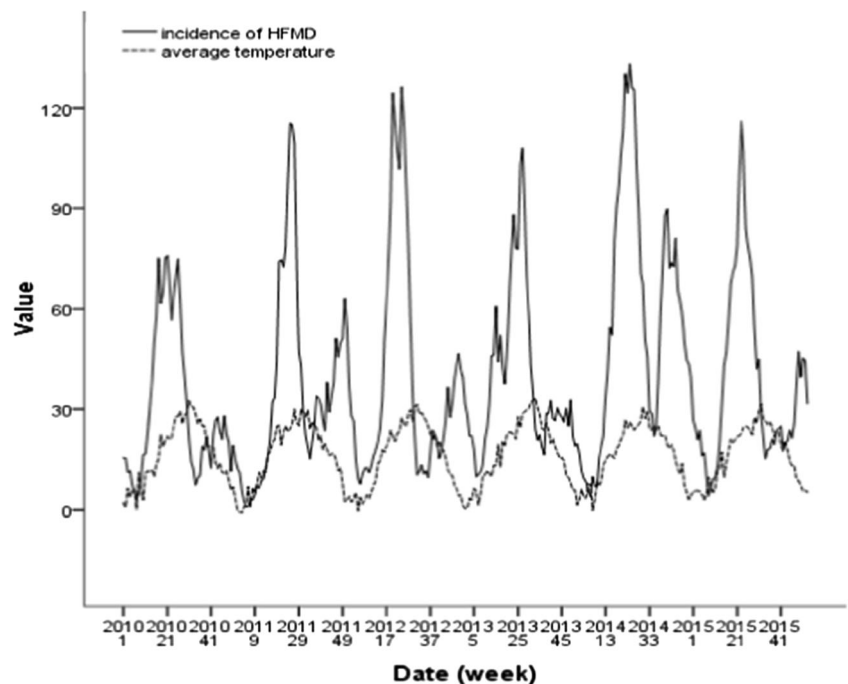
**Time-series analysis**

In the development of a multivariate SARIMA model, natural logarithm transformation was performed to stabilize the variance of the series. Average temperature and air pressure that significantly correlated with HFMD cases were included as external independent variables. Cross-autocorrelation analysis was used to assess the associations between HFMD cases and average temperature and air pressure over a range of time lags. To estimate the parameters of the SARIMA model, we drew the graphs of ACF and PACF based on the modeling results. Furthermore, we fitted the data with different orders and excluded the models in which the residual was not likely to be white noise. Among these models, the SARIMA (2,0,0)<sub>52</sub> associated with the average temperature at lag of 1 week is the most appropriate, which had the lowest BIC and the

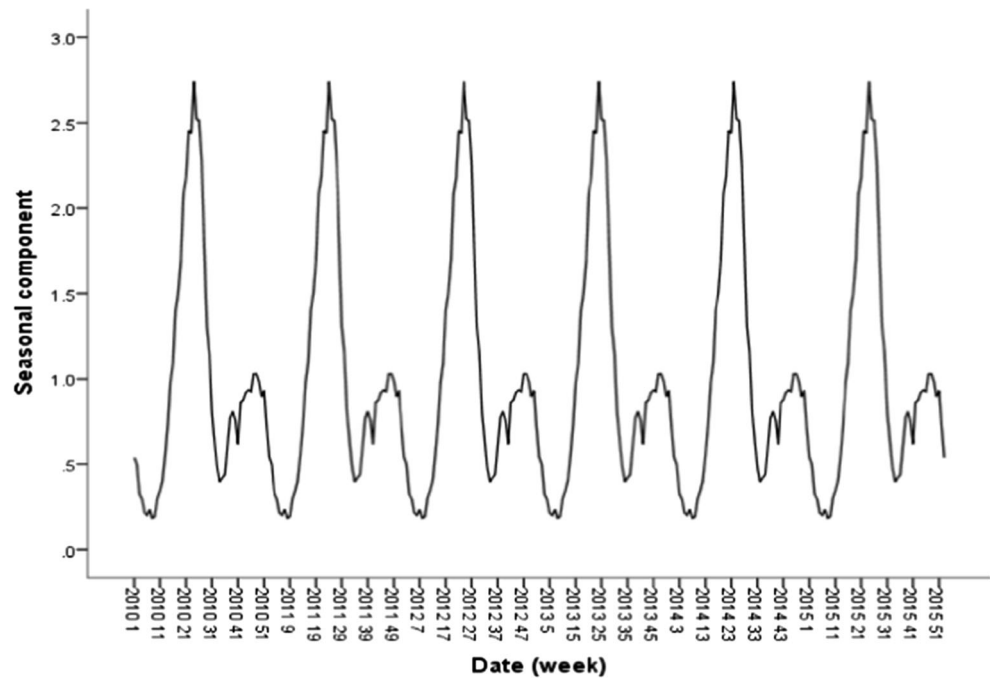
highest  $R^2$  values (8.465 and 0.936, respectively). Estimation of the SARIMA model parameters and their testing results are presented in Table 3. The results of the SARIMA model (2,0,0)<sub>52</sub> show that autogression1 ( $\beta = 1.149$ ,  $P$  value  $< 0.0001$ ), autogression2 ( $\beta = - 0.284$ ,  $P$  value  $< 0.0001$ ), and the average temperature at lag of 1 week ( $\beta = 0.012$ ,  $P$  value  $< 0.05$ ) were statistically significantly associated with the weekly counts of HFMD cases. The ACF and PACF of the residuals of SARIMA (2,0,0)<sub>52</sub> are presented in Fig. 4, which reveals that autocorrelations and partial autocorrelations of residuals were near zero, indicating that the residuals did not deviate significantly from a zero-mean white noise process. Results of the Ljung-Box test confirmed that the autocorrelation coefficients were not statistically significantly different from zero ( $Q = 14.67$ ,  $P = 0.549$ ). Goodness-of-fit analysis indicated that the SARIMA (2,0,0)<sub>52</sub> model fitted the data reasonably well.

Finally, the model was used to predict the weekly incidences of HFMD in Nanjing for the 104 weeks between

**Fig. 2** Weekly incidences of HFMD and weekly average temperature in Nanjing, China, from January 2010 to December 2015. Solid lines indicate incidence of HFMD. Dash lines indicate the average temperature



**Fig. 3** Seasonal distribution of weekly number of HFMD cases after seasonal decomposition in Nanjing, China, from January 2010 to December 2015. The x-axis represents time (2010–2015) and the y-axis the seasonal component



January 2014 and December 2015 and was validated using the actual observations. Figure 5 shows a broadly good fit between observed and predicted data (MAPE = 18.024). The SARIMA model's 1-step-ahead predictions were closer to the observed HFMD values than the 104-steps-ahead predictions. Sensitivity, specificity, and overall agreement rate of 1-step-ahead predictions were 80, 96.63, and 94.23%, respectively, and those of 104-steps-ahead predictions were 60, 96.63, and 91.35%, respectively (Table 4).

## Discussion

HFMD is one of the most common infectious diseases globally. In recent years, the outbreaks of HFMD in the Asia-Pacific region attracted increasing attention. HFMD has been an important public health issue in mainland China (Xing et al. 2014), and an accurate prediction of epidemic can improve the effectiveness of disease control and prevention. Thus, it is important to develop an early warning system of HFMD

**Table 3** Parameter estimates and their testing results of the SARIMA (2,0,0)<sub>52</sub> model

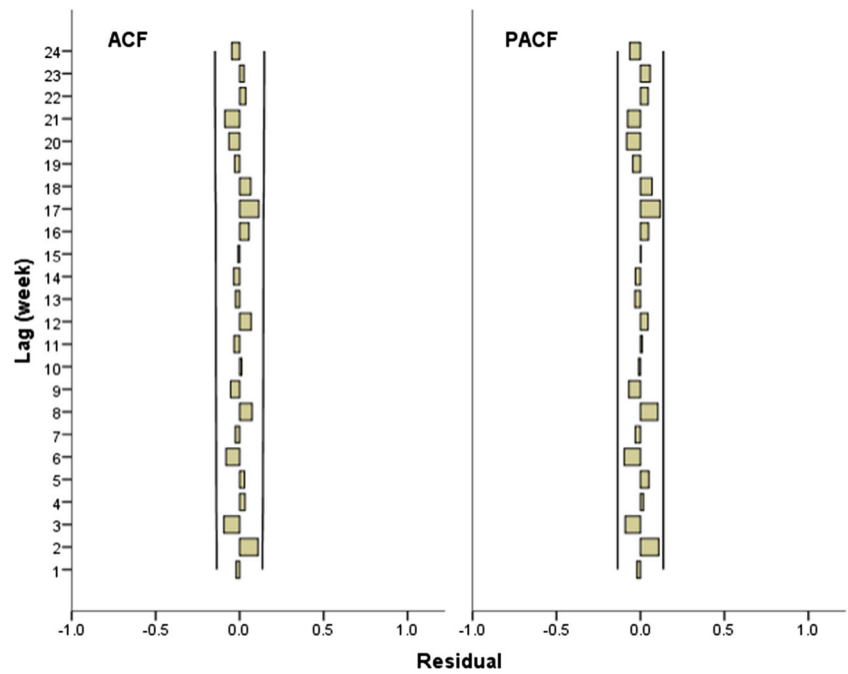
Parameters	Coefficients	Standard error	<i>t</i>	<i>P</i> value
AR1	1.149	0.067	17.036	0.000
AR2	-0.284	0.067	-4.217	0.000
T{avg.}-Lag1 week	0.012	0.005	2.415	0.017

SARIMA seasonal autoregressive integrated moving average model, AR autoregressive, T{avg.}-Lag1 week average temperature at lag of 1 week

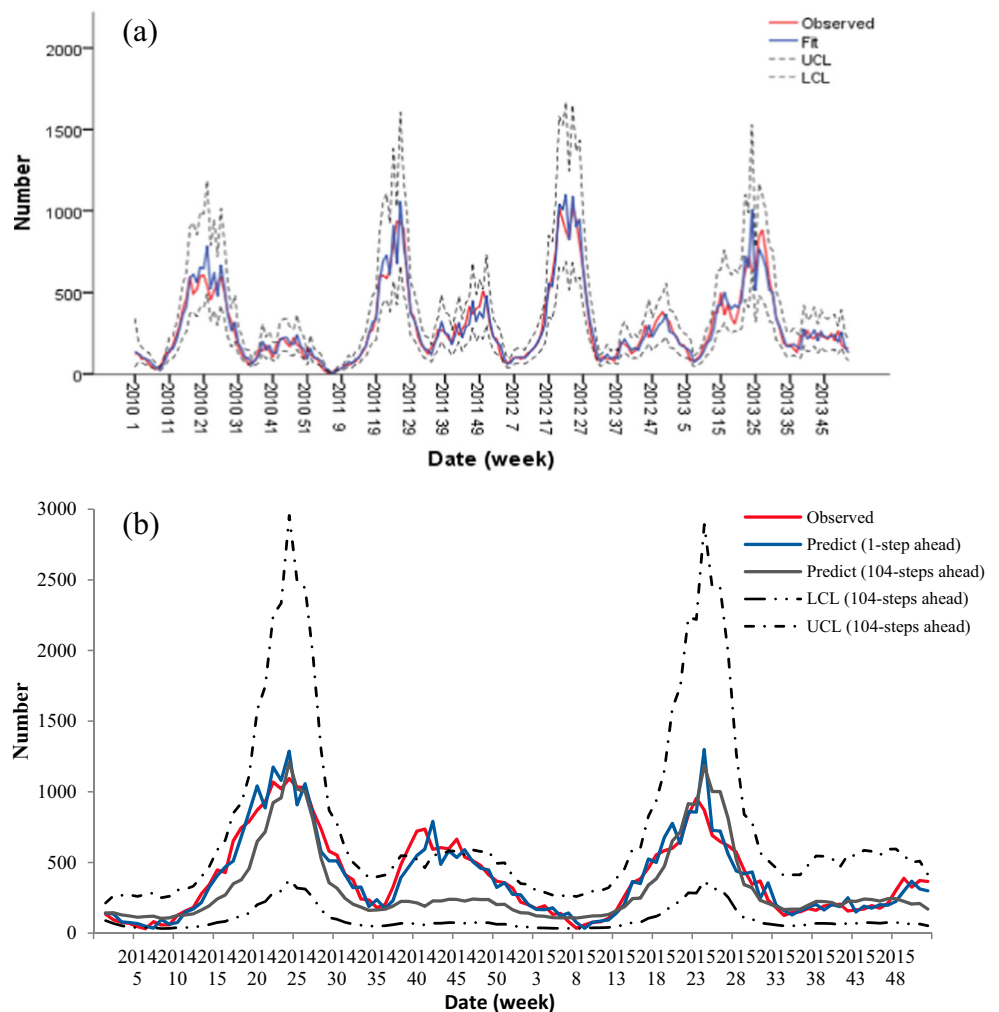
outbreaks which could help to target timely public health interventions to minimize the impact of this disease.

HFMD seemed to have two peaks per year. The first peak occurred in April to July, followed by a second small peak in October to December. According to Rajtar et al. (2008), enteroviruses are resilient to the environmental conditions of the gastrointestinal tract, and that their stability in external environmental conditions is dependent on temperature, humidity, and UV radiation. Therefore, enteroviral infections tend to develop during summer and early autumn in temperate areas; meanwhile, enteroviral infections might maintain a constant level throughout the year in tropical and subtropical areas (Onozuka and Hashizume 2011). Furthermore, Bertrand et al. (2012) also reported that temperature and UV radiation are two main factors leading to enterovirus inactivation. Hence, extremely high temperature may shorten the survival time of enterovirus in the environment and then reduce the risk of HFMD transmission. Another possible explanation is that warm weather may increase the chances for individuals' contact with infectious individuals or contaminated environment. However, in summer time, people tend to stay in air-conditioned settings which will decrease the chances for transmission. The seasonality of HFMD in Nanjing is similar to the patterns seen in other areas of China (Xie et al. 2014). However, there are different seasonal patterns in other regions. For instance, the disease peaks during the summer in Japan (Onozuka and Hashizume 2011) and during the autumn in Finland (Blomqvist et al. 2010). This discrepancy might be attributed to some risk factors of the disease such as climatic, geographic, social, and economic factors (Liu et al. 2015b).

**Fig. 4** Autocorrelation function (ACF) and partial autocorrelation function (PACF) of the residual series of the SARIMA (2,0,0)<sub>52</sub> model. The x-axis gives the value of the correlation coefficient comprised between -1 and 1 and, the y-axis, the number of lags in weeks. Solid lines indicate 95% confidence interval



**Fig. 5 a** Observed and fitted value of the multivariate SARIMA model between 2010 and 2013. The x-axis represents time (2010–2013) and the y-axis the weekly number of HFMD cases. The red lines indicate observed values, the blue lines indicate fitted values, and the dash lines indicate the upper control limit and lower control limit of fitted values. **b** Prediction of the number of HFMD cases from 2014 to 2015 based on the multivariate SARIMA model. The x-axis represents time (2014–2015) and the y-axis the weekly number of HFMD cases. The red lines indicate observed values, the blue lines indicate predicted values of 1-step-ahead prediction, the gray lines indicate predicted values of 104-steps-ahead prediction, and the dash lines indicate the upper control limit and lower control limit of 104-steps-ahead prediction



**Table 4** Sensitivity analysis of the observed values and predicted values

Observed values	1-step-ahead prediction values		104-steps-ahead prediction values	
	Outbreak	Unoutbreak	Outbreak	Unoutbreak
Outbreak	12	3	9	6
Unoutbreak	3	86	3	86

Weather factors have been proven to have significant influence on occurrence and transmission of HFMD. The current study, which was conducted in Nanjing, demonstrated that high temperature and low air pressure presented a higher risk of HFMD infection. These findings are consistent with a previous study in Guangzhou, China, which suggested that each 1 °C rise in temperature corresponded to an increase of 9.38% in the weekly number of HFMD cases and a 1 hPa rise in air pressure corresponded to a decrease in the number of cases by 6.80% (Li et al. 2014). Other studies in mainland China (Chen et al. 2015b; Feng et al. 2014) and Hong Kong (Ma et al. 2010) also reported the positive association between average temperature and HFMD. However, few studies have been published revealing the relationship between air pressure and HFMD. A similar finding in the north of China indicated that air pressure behaved in the opposite way to the HFMD incidence (Wang et al. 2011). The exact mechanism about the association between air pressure and HFMD has rarely been investigated. Evidence has shown that low air pressure may have the adverse impact on human immune system (Maes and De Meyer 2000). For example, the study by Styra et al. (2009) indicated that the correlation between atmospheric pressure decrease and atmospheric pressure increase of cardiovascular disease in 1–2 days and on the same day was 25–44%.

In this study, we developed a multivariate SARIMA model using average temperature and HFMD surveillance data from 2010 to 2013 for predicting the outbreak of HFMD between 2014 and 2015 in Nanjing, China. Our results suggest that the SARIMA (2,0,0)<sub>52</sub> model well reflected the trend of the incidence of HFMD in Nanjing, China. The model also showed that the average temperature was a key determinant of HFMD transmission. As far as we know, this is the first study to apply SARIMA model to forecast weekly incidences of HFMD in Nanjing, China, whereas only Yu et al. (2015) examined the relationship between HFMD and meteorological factors by distributed lag non-linear model (DLNM) in this region. Similar to our findings, Yu et al. (2015) reported average temperature and air pressure were correlated with the incidence of HFMD. However, comparing with the SARIMA model which was widely used to predict the incidences of infectious disease, DLNM is more appropriate to estimate the total relative risks of delayed exposure to climatic factors on HFMD. The results of study by Yu et al. (2015) further indicated that lower temperature and air pressure may have lag effects on HFMD incidences. In

current study, we have focused on forecasting the outbreak of HFMD by using the SARIMA model and our findings suggest that the predicted values matched the observed values reasonably well.

To date, few studies used SARIMA model to predict HFMD incidence in other regions, and the results varied across different studies. For example, Liu et al. (2016) developed a univariate SARIMA model based on the monthly surveillance data collected between 2010 and 2014 in Sichuan, China, and found that the SARIMA (1,0,1) (0,1,0)<sub>12</sub> model was the best fitting model to estimate the HFMD incidence trend. By using the data of 2932 hospitalized patients from 2008 to 2012, Feng et al. (2014) showed that the SARIMA (0,1,0) (1,0,0)<sub>52</sub> associated with average temperature at lag of 2 weeks was suitable to predict the weekly number of HFMD hospitalizations in Zhengzhou, China. According to our results, the SARIMA (2,0,0)<sub>52</sub> model with average temperature at lag of 1 week was reliable with high validity (stationary  $R^2 = 0.936$ ), which suggests that the SARIMA model could be used to predict the future changing trends of HFMD in Nanjing, China. We found that HFMD incidence is better predicted when average temperature at lag of 1 week is included as an external independent variable in the model. For the lag effects, it is compatible with the incubation period of enteroviruses and the possible delay for parents' response to the clinical symptoms of children (Huang et al. 2013). Furthermore, the relationship between average temperature and HFMD has also been reported in other studies (Huang et al. 2013; Lin et al. 2013; Chen et al. 2014). Currently, although the exact mechanism remains unknown, there are several explanations for this association. Temperature could not only influence children's behavior but also has a potential effect on the survival and transmission of the enteroviruses, and then affects the transmission of the disease (Yin et al. 2016).

In this study, we have constructed the SARIMA model which fitted HFMD data reasonably well in Nanjing, China. Both the approach of 104-steps-ahead prediction and the approach of 1-step-ahead prediction were compared in the validation process. The forecast results indicate that the performance of the 1-step-ahead prediction was slightly better than that of the 104-steps-ahead prediction, although both sets of results were generally acceptable (sensitivity, 80 and 60%, respectively; specificity, 96.63% for both; overall agreement rate, 94.23 and 91.35, respectively). The approach of 104-



steps-ahead prediction failed to detect the peak of HFMD during September and December 2014, while the approach of 1-step-ahead prediction did well. According to Luz et al. (2008), 1-step-ahead prediction can really reflect the on-going nature of disease surveillance. The model allows the incorporation of the newly collected data, therefore increasing its predictive power. This real-time forecasts could be used by public health authorities to identify the outbreaks of HFMD and consequently to implement timely interventions. With the help of the SARIMA model, it is reasonable for the public health authorities to allocate health resources to control the epidemic efficiently. If prediction results continue to rise, more human and material resources should be prepared for health interventions in advance. HFMD morbidity and mortality would be minimized through accurate forecasts.

This study has two strengths. Firstly, a sophisticated time-series model was used in the attempt to develop an epidemic forecasting system for the control and prevention of HFMD in Nanjing, China. Secondly, the model developed in this study appears to have a high degree of accuracy. Nevertheless, several limitations of our study should also be acknowledged. First, our surveillance data of HFMD do not capture all cases because some asymptomatic and mildly affected patients may not seek health care; thus, actual cases could be many times higher and underreporting biases may not be completely excluded using existing official statistics. Second, this study is based on the data from one city, so the results may not be generalizable to other places with different climates. Multi-city studies in different areas with diverse weather patterns are needed to be undertaken in the future. Third, this investigation is an ecological study, which does not enable us to assess individual exposure level and exclude the potential unmeasured confounding variables. Finally, our study develops a forecasting model for HFMD of Nanjing city without considering the intra-city difference. More attention in further research should be paid to spatiotemporal characteristics of HFMD so that the communities at high risk can be identified. Thus, the local health authorities could formulate the more regional specific interventions and mobilize limited resources to prevent and control HFMD outbreaks. The model incorporating social economic factors and internet-based data will improve our prediction capability to more precisely forecast HFMD outbreaks in the future.

## Conclusions

Our findings demonstrate that the SARIMA model coupling with the data on climatic factors and disease surveillance could be an important tool for early detection and prediction of the HFMD outbreak. An effective warning system for HFMD could reduce childhood morbidity and should be incorporated in disease control and prevention strategies for

HFMD in Nanjing and other areas with similar climatic conditions.

**Acknowledgements** This study was supported by the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions and Jiangsu Government Scholarship for Overseas Studies.

**Compliance with ethical standards** This study was approved by the Ethical Committee of Nanjing Medical University.

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Bertrand I, Schijven JF, Sanchez G, Wyn-Jones P, Ottoson J, Morin T et al (2012) The impact of temperature on the inactivation of enteric viruses in food and water: a review. *J Appl Microbiol* 112(6): 1059–1074
- Blomqvist S, Klemola P, Kajjalainen S, Paananen A, Simonen ML, Vuorinen T et al (2010) Co-circulation of coxsackieviruses A6 and A10 in hand, foot and mouth disease outbreak in Finland. *J Clin Virol* 48(1):49–54
- Box G, Jenkins G (1970) *Time-series analysis: forecasting and control*. Holden-Day, San Francisco
- Chan KP, Goh KT, Chong CY, Teo ES, Lau G, Ling AE (2003) Epidemic hand, foot and mouth disease caused by human enterovirus 71, Singapore. *Emerg Infect Dis* 9(1):78–85
- Chan LG, Parashar UD, Lye MS, Ong FG, Zaki SR, Alexander JP et al (2000) Deaths of children during an outbreak of hand, foot, and mouth disease in Sarawak, Malaysia: clinical and pathological characteristics of the disease. For the outbreak study group. *Clin Infect Dis* 31(3):678–683
- Chen B, Sumi A, Toyoda S, Hu Q, Zhou D, Mise K et al (2015a) Time series analysis of reported cases of hand, foot, and mouth disease from 2010 to 2013 in Wuhan, China. *BMC Infect Dis* 15:495
- Chen C, Lin H, Li X, Lang L, Xiao X, Ding P et al (2014) Short-term effects of meteorological factors on children hand, foot and mouth disease in Guangzhou, China. *Int J Biometeorol* 58(7):1605–1614
- Chen Z, Sun H, Yan Y, Wang Y, Zhu C, Zhou W et al (2015b) Epidemiological profiles of hand, foot, and mouth disease, including meteorological factors, in Suzhou, China. *Arch Virol* 160(1):315–321
- Cheng J, Zhu R, Xu Z, Wu J, Wang X, Li K et al (2016) Impact of temperature variation between adjacent days on childhood hand, foot and mouth disease during April and July in urban and rural Hefei, China. *Int J Biometeorol* 60(6):883–890
- Chua KB, Kasri AR (2011) Hand foot and mouth disease due to enterovirus 71 in Malaysia. *Virol Sin* 26(4):221–228
- De W, Changwen K, Wei L, Monagin C, Jin Y, Cong M et al (2011) A large outbreak of hand, foot, and mouth disease caused by EV71 and CAV16 in Guangdong, China, 2009. *Arch Virol* 156(6):945–953
- Deng T, Huang Y, Yu S, Gu J, Huang C, Xiao G et al (2013) Spatiotemporal clusters and risk factors of hand, foot, and mouth disease at the district level in Guangdong province, China. *PLoS One* 8(2): e56943
- Feng H, Duan G, Zhang R, Zhang W (2014) Time series analysis of hand-foot-mouth disease hospitalization in Zhengzhou: establishment of forecasting models using climate variables as predictors. *PLoS One* 9(1):e87916
- Guo C, Yang J, Guo Y, Ou QQ, Shen SQ, Ou CQ et al (2016) Short-term effects of meteorological factors on pediatric hand, foot, and mouth

- disease in Guangdong, China: a multi-city time-series analysis. *BMC Infect Dis* 16(1):524
- Hii YL, Rocklöv J, Ng N (2011) Short term effects of weather on hand, foot and mouth disease. *PLoS One* 6(2):e16796
- Ho M, Chen ER, Hsu KH, Twu SJ, Chen KT, Tsai SF et al (1999) An epidemic of enterovirus 71 infection in Taiwan. Taiwan Enterovirus Epidemic Working Group. *N Engl J Med* 341(3):929–935
- Huang Y, Deng T, Yu S, Gu J, Huang C, Xiao G et al (2013) Effect of meteorological variables on the incidence of hand, foot, and mouth disease in children: a time-series analysis in Guangzhou, China. *BMC Infect Dis* 13:134
- Jiang M, Wei D, WL O, Li KX, Luo DZ, Li YQ et al (2012) Autopsy findings in children with hand, foot, and mouth disease. *N Engl J Med* 367(1):91–92
- Li T, Yang Z, Di B, Wang M (2014) Hand foot and mouth disease and weather factors in Guangzhou, southern China. *Epidemiol Infect* 142(8):1741–1750
- Lin H, Zou H, Wang Q, Liu C, Lang L, Hou X et al (2013) Short-term effect of El Niño-Southern Oscillation on pediatric hand, foot and mouth disease in Shenzhen, China. *PLoS One* 8(7):e65585
- Liu L, Luan RS, Yin F, Zhu XP, Lu Q (2016) Predicting the incidence of hand, foot and mouth disease in Sichuan province, China using the ARIMA model. *Epidemiol Infect* 144(1):144–151
- Liu MY, Liu W, Luo J, Liu Y, Zhu Y, Berman H et al (2011) Characterization of an outbreak of hand, foot, and mouth disease in Nanchang, China in 2010. *PLoS One* 6(9):e25287
- Liu W, Ji H, Shan J, Bao J, Sun Y, Li J et al (2015a) Spatiotemporal dynamics of hand-foot-mouth disease and its relationship with meteorological factors in Jiangsu province, China. *PLoS One* 10(6):e0131311
- Liu Y, Wang X, Pang C, Yuan Z, Li H, Xue F (2015b) Spatio-temporal analysis of the relationship between climate and hand, foot, and mouth disease in Shandong province, China, 2008–2012. *BMC Infect Dis* 15:146
- Luz PM, Mendes BV, Codeco CT, Struchiner CJ, Galvani AP (2008) Time series analysis of dengue incidence in Rio de Janeiro, Brazil. *Am J Trop Med Hyg* 79(6):933–939
- Ma E, Lam T, Wong C, Chuang SK (2010) Is hand, foot and mouth disease associated with meteorological parameters? *Epidemiol Infect* 138(12):1779–1788
- Maes M, De Meyer F (2000) Relationships of climatic data to immune and hematologic variables in normal human. *Neuro Endocrinol Lett* 21(2):127–136
- Midekisa A, Senay G, Henebry GM, Semuniguse P, Wimberly MC (2012) Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia. *Malar J* 11:165
- Onozuka D, Hashizume M (2011) The influence of temperature and humidity on the incidence of hand, foot, and mouth disease in Japan. *Sci Total Environ* 410–411:119–125
- Puenpa J, Theamboonlers A, Korkong S, Linsuwanon P, Thongmee C, Chatproedprai S et al (2011) Molecular characterization and complete genome analysis of human enterovirus 71 and coxsackievirus A16 from children with hand, foot and mouth disease in Thailand during 2008–2011. *Arch Virol* 156(11):2007–2013
- Rajtar B, Majek M, Polanski L, Polz-Dacewicz M (2008) Enteroviruses in water environment—a potential threat to public health. *Ann Agric Environ Med* 15(2):199–203
- Qian H, Huo D, Wang X, Jia L, Li X, Li J et al (2016) Detecting spatial-temporal cluster of hand foot and mouth disease in Beijing, China, 2009–2014. *BMC Infect Dis* 16(1):206
- Ren H, Li J, Yuan ZA, JY H, Yu Y, YH L (2013) The development of a combined mathematical model to forecast the incidence of hepatitis E in Shanghai, China. *BMC Infect Dis* 13:421
- Shi Y, Liu X, Kok SY, Rajarethinam J, Liang S, Yap G et al (2016) Three-month real-time dengue forecast models: an early warning system for outbreak alerts and policy decision support in Singapore. *Environ Health Perspect* 124(9):1369–1375
- Styra D, Usovaite A, Damauskaitė J, Juozulynas A (2009) Leaps in cardiovascular diseases after a decrease of hard cosmic ray flux and atmospheric pressure in Vilnius city in 2004–2007. *Int J Biometeorol* 53(6):471–477
- Tu PV, Thao NT, Perera D, Huu TK, Tien NT, Thuong TC et al (2007) Epidemiologic and virologic investigation of hand, foot, and mouth disease, southern Vietnam, 2005. *Emerg Infect Dis* 13(11):1733–1741
- Wang JF, Guo YS, Christakos G, Yang WZ, Liao YL, Li ZJ et al (2011) Hand foot and mouth disease: spatiotemporal transmission and climate. *Int J Health Geogr* 10:25
- Wang P, Goggins WB, Chan EY (2016) Hand, foot and mouth disease in Hong Kong: a time-series analysis on its relationship with weather. *PLoS One* 11(8):e0161006
- Wei J, Hansen A, Liu Q, Sun Y, Weinstein P, Bi P (2015) The effect of meteorological variables on the transmission of hand, foot and mouth disease in four major cities of Shanxi province, China: a time series data analysis (2009–2013). *PLoS Negl Trop Dis* 9(3):e0003572
- Wong SS, Yip CC, Lau SK, Yuen KY (2010) Human enterovirus 71 and hand, foot and mouth disease. *Epidemiol Infect* 138(8):1071–1089
- Wu Y, Yeo A, Phoon MC, Tan EL, Poh CL, Quak SH et al (2010) The largest outbreak of hand; foot and mouth disease in Singapore in 2008: the role of enterovirus 71 and coxsackievirus A strains. *Int J Infect Dis* 14(12):e1076–e1081
- Xiao X, Gasparrini A, Huang J, Liao Q, Liu F, Yin F et al (2017) The exposure-response relationship between temperature and childhood hand, foot and mouth disease: a multicity study from mainland China. *Environ Int* 100:102–109
- Xie YH, Chongsuvivatwong V, Tang Z, McNeil EB, Tan Y (2014) Spatio-temporal clustering of hand, foot, and mouth disease at the county level in Guangxi, China. *PLoS One* 9(2):e88065
- Xing W, Liao Q, Viboud C, Zhang J, Sun J, JT W et al (2014) Hand, foot, and mouth disease in China, 2008–12: an epidemiological study. *Lancet Infect Dis* 14(4):308–318
- Xu J, Zhao D, Su H, Xie M, Cheng J, Wang X et al (2016) Impact of temperature variability on childhood hand, foot and mouth disease in Huainan, China. *Public Health* 134:86–94
- Yin F, Zhang T, Liu L, Lv Q, Li X (2016) The association between ambient temperature and childhood hand, foot, and mouth disease in Chengdu, China: a distributed lag non-linear analysis. *Sci Rep* 6:27305
- Yu Y, Li J, Zhou L, Chen XD (2015) Association of meteorological factors with incidence of hand, foot and mouth disease in Nanjing city from 2010 to 2013 [in Chinese]. *J Environ Occup Med* 32:1132–1135
- Zhang Y, Zhu Z, Yang W, Ren J, Tan X, Wang Y et al (2010) An emerging recombinant human enterovirus 71 responsible for the 2008 outbreak of hand foot and mouth disease in Fuyang city of China. *Virol J* 7:94