International Journal of Infectious Diseases 74 (2018) 61–70

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

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

Epidemiology and ARIMA model of positive-rate of influenza viruses among children in Wuhan, China: A nine-year retrospective study

Zhirui He[a,b], Hongbing Tao[a,*]

[a] Department of Health Management, School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No. 13 Hongkong Road, Wuhan 430030, China

[b] Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital), Tongji Medical College, Huazhong University of Science & Technology, No. 100 Hongkong Road, Wuhan 430016, China

ARTICLE INFO

Article history:
Received 11 March 2018
Received in revised form 1 July 2018
Accepted 2 July 2018
Corresponding Editor: Eskild Petersen, Aarhus, Denmark

Keywords:
Influenza virus
Children
Time series analysis
ARIMA model

ABSTRACT

Objective: Influenza is a common childhood disease and protecting children by predicting the positive rate of influenza virus is important as vaccines are not routinely administered in China. Our study aims to describe the epidemiology of influenza viruses among children in Wuhan, China during the past nine influenza seasons (2007–2015) and to predict the positive rate of different types of influenza virus in the future.
Methods: During the last nine influenza seasons (2007–2015), a total of 10,232 nasopharyngeal swabs collected from pediatric outpatients (age <15 years) with influenza-like illness (ILI) infections in two sentinel children's hospitals were examined for influenza A and B by real-time one step RT-PCR. An autoregressive integrated moving average (ARIMA) model was used to fit the time series and to predict the future (first half of 2016) positive rates of different types of influenza virus.
Results: A total of 1,341 specimens were positive for influenza A and 490 for influenza B. The majority of infected patients were 1–11 years old (87.7%). The ARIMA model could effectively predict the positive rate of influenza virus in a short time. ARIMA(0,0,11), SARIMA$(1,0,0)(0,1,1)_{12}$, ARIMA(0,0,1) and SARIMA$(0,0,1)(1,0,1)_{12}$ were suitable for B(Victoria), B(Yamagata), A(H1N1)pdm09, and A(H3N2), respectively.
Conclusion: Additional policies must be formulated to prevent and control influenza. The wide use of influenza vaccines, especially for influenza B, especially for influenza B(Yamagata) and B(Victoria), can potentially reduce the effects of influenza on children of China.

Introduction

Influenza virus affects individuals of all ages and is the leading cause of respiratory illness among children (Munoz, 2002; O'Brien et al., 2004). Influenza virus causes 3–5 million severe cases and 250,000–500,000 deaths per year, affecting 20%–30% of children worldwide (Lin et al., 2015). Children exhibit the highest rate of hospitalization attributable to influenza, and the annual incidence rate in children is estimated to be 20% (Munoz, 2002; Neuzil et al., 2002; Aymard et al., 2003; Bonney et al., 2012; Gioula et al., 2013). An estimated 9243–105,690 influenza-associated respiratory deaths occur annually among children younger than 5 years (Iuliano et al., 2017). During epidemics, the incidence of influenza virus infection may exceed 40% in preschool children and 30% in school-age children (Longini et al., 1982; Glezen et al., 1991; Monto and Sullivan, 1993).

The infection rates of influenza are highest in young children because of a lack of both prior exposure and immunity to the virus. Children introduce and spread influenza virus to household members and the community (Hurwitz et al., 2000; Mansiaux et al., 2015). Influenza vaccination is the most effective way to prevent the disease. Substantial evidence indicates that immunization of young children can reduce the incidence of influenza (Reichert, 2002; Jordan et al., 2006). In China, influenza vaccination is not routinely recommended, and immunization of children is uncommon (Ji et al., 2010). To provide effective strategies for influenza prevention, control, and vaccination among young children in China, the epidemiological and virological characteristics of different influenza virus subtypes infecting children must be understood. An early warning model of influenza plays a key role in influenza prevention and control. Time series analysis has been widely used in epidemiology to predict epidemics of infectious

* Corresponding author.
E-mail addresses: hzr@hust.edu.cn (Z. He), hhbtao@hust.edu.cn (H. Tao).

https://doi.org/10.1016/j.ijid.2018.07.003

diseases, including influenza (Crighton et al., 2004; Saltyte and Hofoss, 2008). Incidence data are used for estimation in most studies but are not accurate enough for children in China because of the existence of influenza-like illness (ILI). Few scholars have studied the application of time series analysis using the positive rate of influenza virus. In our opinion, establishing a fitting model and using a positive rate of influenza virus in China to predict the future one are more accurate approaches than using incidence data.

Influenza surveillance is an important tool for monitoring epidemiological trends and obtaining the positive rate of different types of influenza virus. Influenza surveillance in Wuhan was launched simultaneously with the national surveillance in 2000. Wuhan is located in the central area of China (113°41′–115°05′E, 29°58′–31°22′N) and is the capital of Hubei Province. Its land area is 8,569.15 km^2, with over 10,760,000 permanent residents, including more than 1,100,000 children under 15 years of age (2016). In the present study, we have summarized influenza surveillance data and analyzed the epidemiological and virological characteristics of influenza virus among children (<15 years of age) in Wuhan from 2007 to 2015. Subsequently, we performed a time series analysis of the positive rate of different subtypes of influenza virus to construct four time series models for short-time prediction. We hope this study serves as a reference for influenza prevention and control thereby improving public health.

Materials and methods

Case identification

Patients with ILI were identified based on a standard case definition (PRC, 2011), including body temperature ≥38 °C with either cough or sore throat, in the absence of an alternative diagnosis. Sore throat was defined as an inflamed or red pharynx upon examination by a physician.

Sample source

Influenza surveillance was undertaken in two large tertiary care children's hospitals (Wuhan Children's Hospital; Maternal and Child Health Hospital of Hubei Province), which are the only two children's hospitals serving Wuhan district and have been the national sentinel hospitals of influenza surveillance in children since 2000. In each sentinel hospital, 20 throat swabs per surveillance week were collected by trained nurses from ILI patients who visited outpatient and/or emergency departments. The swabs were stored at 4 °C in sterile viral transport medium for influenza virus testing in 48 h.

Laboratory diagnosis and virus identification

From January 2007 to March 2009, all samples from sentinel hospitals were tested for influenza viruses by virus isolation in Madin–Darby canine kidney (MDCK) cells. MDCK cell culture was replaced by Real-time reverse transcription polymerase chain reaction (RT-PCR) assay to identify the types/subtypes of influenza virus in accordance with a standard protocol since April 2009 (PRC, 2010). The testing and viral identification were performed by the laboratory of the Center for Disease Control and Prevention of Wuhan and Hubei. The laboratory results were submitted to the national influenza surveillance network after the completion of tests.

Data collection

We obtained monthly influenza surveillance data from the national influenza surveillance network, obtaining the number of laboratory-confirmed influenza cases by virus type, that is, influenza A and B. For an accurate study, we selected the cases of patients who lived in Wuhan, with ILI, and under 15 years old from January 2007 to December 2015.

Time series analysis

Autoregressive integrated moving average (ARIMA)(p, d, q) model is an extension of autoregressive (AR), moving average (MA), and ARMA models (Rojas and Pomares, 2016). Both of these models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting). The letters p, d, and q correspond to the meaning of the order of autoregression, degree of difference, and order of moving average, respectively. If the data showed no seasonality, the ARIMA model was considered to be perfect for forecasting. However, if evident seasonality in the data existed, we utilized the seasonal ARIMA (SARIMA $[p, d, q][P,D,Q]_s$) model. The letters p, q, and d correspond to the meaning of the order of autoregression, degree of difference, and order of moving average, respectively. P, D, Q, and s stand for seasonal autoregression, seasonal integration, seasonal moving average, and seasonal period length. Three key steps were involved in the ARMA modeling, namely, identification, estimation, and diagnosis. First, we checked stationarity and seasonality by drawing a time series plot of the monthly positive rate of influenza virus. Thus, we constructed the model and depicted the autocorrelation function (ACF) and partial autocorrelation function (PACF) of model residuals to confirm autoregressive and moving average parameters. Second, we obtained the best ARIMA model using SPSS professional modeler. The final automatic model selection provided the optimal model with the lowest value in Bayesian information criterion (BIC) and highest R^2 (the coefficient of determination, a statistic that will tell the goodness of fit of a model). The ACF and PACF of residuals, as well as the test of white noise, were determined to evaluate the goodness of fit. Finally, we applied this model to forecast the monthly positive rate of influenza virus from January to June 2016. To validate the predictive ability of the model, we used the data from 2007 to 2015 as a training set and the model was fitted only to these data. The predictive ability of the model was then assessed by generating predictions for the first half of 2016 as the test set and comparing the predictions made based on data until June 2016 to the observed values in 2016. A line chart for actual and predicted values was constructed to show the goodness of fit.

Statistical methods

Cochran–Armitage trend test was used to analyze the trend change of influenza virus positivity with the increase in age. The monthly data of positive influenza by subtype and the percentage of specimens that tested positively were plotted to describe the seasonality and circulation of influenza types/subtypes among ILI cases. Time series analysis was performed to construct an ARIMA model for forecasting the positive rate in the first half of 2016. Two-sided P-values <0.05 were considered to be statistically significant. Descriptive statistics, Cochran-Armitage trend test, and time series analysis were conducted using Excel 2016 (Microsoft Corp., USA), SAS version 9.1.3 (SAS Institute, Cary, NC, USA), and SPSS Statistics 19 (IBM Corp., USA), respectively.

Ethical considerations

The influenza surveillance was a nation-wide governmental public health activity. Ethical approval was given by the Medical Ethics Committee of Wuhan Children's Hospital with the reference number: 2000012; The Research Project of "Time Series Analysis of Positive-Rate Influenza Viruses among Children