Original Article
Asian Pacific Journal of Tropical Medicine
journal homepage: www.apjtm.org

doi: 10.4103/1995-7645.281528
Impact Factor: 1.77

Analysis of temporal trends of human brucellosis between 2013 and 2018 in Yazd Province, Iran to predict future trends in incidence: A time-series study using ARIMA model

Vahid Rahmanian[1,2], Saied Bokaie[3✉], Karamatollah Rahmanian[4], Saeed Hosseini[5], Aliakbar Taj Firouzeh[6]

[1]*Zoonoses Research Center, Jahrom University of Medical Sciences, Jahrom, Iran*
[2]*Health Policy Research Center, Institute of Health, Shiraz University of Medical Sciences, Shiraz, Iran*
[3]*Epidemiology and Zoonosis Division Department of Food Hygiene, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran*
[4]*Zoonoses Research Center, Jahrom University of Medical Sciences, Jahrom, Iran*
[5]*Health Monitoring Research Center, School of Medicine, Shahid Sadoughi University of Medical Sciences, Yazd, Iran*
[6]*Shahid Sadoughi University of Medical Sciences, Yazd, Iran*

ABSTRACT

Objective: To determine the temporal patterns of cumulative incidence of brucellosis using autoregressive integrated moving average models.

Methods: This cross-sectional study employed yearly and monthly data of 1 117 laboratory-confirmed human brucellosis cases from January 2013 to December 2018 using the Yazd brucellosis national surveillance system. The monthly incidences constructed a time-series model. The trend of cumulative incidence was perceived by tracing a line plot, which displayed a seasonal trend with periodicity. Thus, the ARIMA models were selected. Thereafter, Akaike information criteria (AIC) and Bayesian information criterion (BIC) values among different models indicated a preferable model from models which were expanded by diverse lags [(3, 0, 3), (2, 0, 3), (3, 0, 2), (4, 0, 3) and (3, 0, 4)]. Then, the achieved ARIMA model was applied to the forecasting cumulative incidence of monthly brucellosis incidences. All analyses were performed using Stata, version 11.2.

Results: For the ARIMA (3, 0, 4) model, MAPE value was 56.20% with standard error 0.009–0.016, and white noise diagnostic check (Q=19.79, P=0.975) for the residuals of the selected model showed that the data were completely modelled. The monthly incidences that were fitted by the ARIMA (3, 0, 4) model, with AIC (25.7) and BIC (43.35) with a similar pattern of actual cases from 2013 to 2018 and forecasting incidences from January 2019 to December 2019 were, respectively, 0.50, 0.44, 0.45, 0.49, 0.55, 0.58, 0.56, 0.51, 0.46, 0.44, 0.45 and 0.49 per 100 000 people.

Conclusions: In summary, the study showed that the ARIMA (3, 0, 4) model can be applied to forecast human brucellosis patterns in Yazd province, supplementing present surveillance systems, and may be better for health policy-makers and planners.

KEYWORDS: Malta fever; Forecasting; Public health surveillance; Iran

1. Introduction

Brucellosis (Malta fever) is endemic in the Middle East (Iran), Africa, Latin America, Central Asia, and the Mediterranean Basin. Most provinces of Iran are endemic with brucellosis, especially the regions where human lives are in close contact with livestock[1,2]. Brucellosis is a neglected zoonotic disease caused by *Brucella* bacteria. Transmission in humans can occur through the use of unpasteurized milk and dairy products, laboratory inhalation, accidental skin penetration, and conjunctiva contact (rarely), blood transfusion, both transplacentally and person-to-person[3,4]. Direct person-to-person transmission infrequently happens, albeit it has been described that transmission may occur via breast-feeding and sexual intercourse[5-7]. Human brucellosis is one of the common

✉To whom correspondence may be addressed. E-mail: sbokaie@ut.ac.ir

For reprints contact: reprints@medknow.com

How to cite this article: Rahmanian V, Bokaie S, Rahmanian K, Hosseini S, Firouzeh AT. Analysis of temporal trends of human brucellosis between 2013 and 2018 in Yazd Province, Iran to predict future trends in incidence: A time-series study using ARIMA model. Asian Pac J Trop Med 2020; 13(6): 272-277.

Article history: Received 27 July 2019 Revision 20 November 2019 Accepted 15 December 2019 Available online 27 March 2020

multisystem diseases with more than 500 000 cases occurring worldwide every year[8,9]. The prevalence of diseases are more than 10 cases per 100 000 population in some countries[10,11]. According to the report of the Center for Disease Control and Prevention (Iran); the distribution of human brucellosis based on cumulative incidence are as follows: very high incidence regions, (31-41 per 100 000), high incidence regions, (21-30 per 100 000), moderate incidence regions (11-20 per 100 000) and low incidence regions (0-10 per 100 000)[12].

Many cases of the disease are reported in Iran due to long borders with endemic countries and lack of supervision on imports of livestock, a large number of tribal population, traditional husbandry methods, failure to supervise the production and distribution of dairy products and the lack of systematic implementation of vaccination, testing, and slaughter of livestock[13]. In reaction to the outbreak of human brucellosis, and according to the zoonotic and economic importance of brucellosis, a fast alert system was immediately required in the epidemic region[14]. Consequently, brucellosis surveillance systems that registered data brucellosis cases were established. Each case of human brucellosis should be reported within 7 days through the National Notifiable Disease Surveillance System by healthcare providers in Iran. Time-series analysis of existing surveillance data in infectious disease is essential in triggering novel hypotheses, the prediction of the observed events and subsequently, creating a quality control system. This form of time data can be investigated by an autoregressive integrated moving average (ARIMA) model[15]. This model can assess trend and seasonality designs with cumulative incidence of brucellosis, and is suitable for forecasting. This study was designed to determine the temporal patterns of brucellosis incidence from 2013 to 2018 in Yazd in the central region of Iran, using ARIMA models. We created a time-series model for brucellosis and forecasted the brucellosis incidence for 2019. This model will be useful for short-term forecasting the epidemic trend of human brucellosis and providing a brucellosis reference guide for interventions.

2. Materials and methods

2.1. Study site

Yazd province is located in central Iran, and, based on the National Population and Housing Census, it had a population of 1138 533 in 2016. The city of Yazd is the administrative capital of Yazd province. It also consists of ten counties. Yazd has a climate that is similar to a dry desert climate. Small rains with high water evaporation, comparatively low humidity, changes in heat and temperature, are among the factors that make this province one of the driest provinces of Islamic Republic of Iran. Yazd province shares its borders with Isfahan, Semnan, Razavi Khorasan, South Khorasan, Kerman and Fars provinces. The production of 233 000 tons of raw milk during 2016 ranked the province fifteenth in the country in terms of production milk. A total of 436 980, 426 000 and 154 400 sheep, goats and cows respectively, were reported in 2016 in this province by the Yazd veterinary organization.

2.2. Data source

This cross-sectional study employed yearly and monthly data of 1 117 laboratory-confirmed human brucellosis cases from January 2013 to December 2018 using the Yazd brucellosis national surveillance system. The cumulative incidence for human brucellosis was calculated on a monthly basis (number of human brucellosis/ population size) and stated per 100 000 people. Then the monthly incidences during study period constructed the time-series model. Before the selection of the model, the trend was observed by plotting the line diagram, and this scheme shows the seasonal trend with a periodicity. Therefore, the ARIMA model was selected.

2.3. Statistical analysis

2.3.1. ARIMA model

When there were no missing data in the stationary seasonal time series and when there were trend changes, periodic changes, and random disorders in the time series of infectious diseases in other regions, the ARIMA model was selected[16]. The ARIMA model was defined via three parameters (p, d, q): p was the number of autoregressive (AR) terms, d was the number of times the model was differenced, and q was the number of moving average (MA) terms[17]. Analysis of the autocorrelation functions (ACF) and partial autocorrelation functions (PACF) plots allowed estimation of the AR and MA parameters for the time series of incidence and therefore identification of plausible models[18,19]. After the identification step, lower Akaike Information Criteria (AIC) value and Bayesian information criterion (BIC) among diverse models demonstrated a preferable model from models which were expanded via diverse lags[20]. Next, the achieved ARIMA model was applied to the predicting cumulative incidence of annually and monthly separate brucellosis incidences eventually, the forecasting accuracy was evaluated *via* the mean absolute percentage error (MAPE), which was calculated using the formula for MAPE:

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{n} \frac{\text{Actual cases-predicted cases}}{\text{Actual cases}} \times 100$$

Where: N = the number of prediction.

2.3.2. Statistical methods

Data analysis was performed in Stata, version 11.2 (Stata Corp, College Station, TX, USA) using commands including: time series, ARIMA and ARIMA models, forecasting, correlogram and partial correlogram graphs and portmanteau white noise test. To investigate the stationary in variance and mean, box cox regression and Dickey-Fuller test were used, respectively. The significance level of 0.05 was considered.

2.4. Ethical considerations

This research was approved by the Health Policy Research Center, Institute of Health, Shiraz University of Medical Sciences, (Ref: 2018/01/62/19278). In using raw data from the national brucellosis surveillance system, all principles were considered to protect the confidentiality of personal information of individuals.

3. Results

3.1. Descriptive data

Most of the patients (95.1%) had been detected for the first time (new cases). A total of 630 (56.4%) males and 487 (43.6%) females with the median age of 35 years (IQR: 31) ranginged from 5 to 94 years were enrolled. Also, 730 (65.4%) of the patients lived in urban regions, 661 (59.2%) of patients had a history of close contact with animals, and 908 (81.3%) had a history of consumption of unpasteurized dairy products. The trend of monthly numbers of brucellosis from the study period is displayed in Figure 1. Absolute case counts per month ranged from 1 to 58. Human brucellosis was most common in the months of March, April, May and June and overall, cumulative incidence peaked in March 2014. It was observed that the disease incidence had a periodic and seasonal trend with the peak in March and the lowest point in December. Time-series plot showed an analogous seasonal trend each year, proposing that the ARIMA model was suitable to fitting the data.

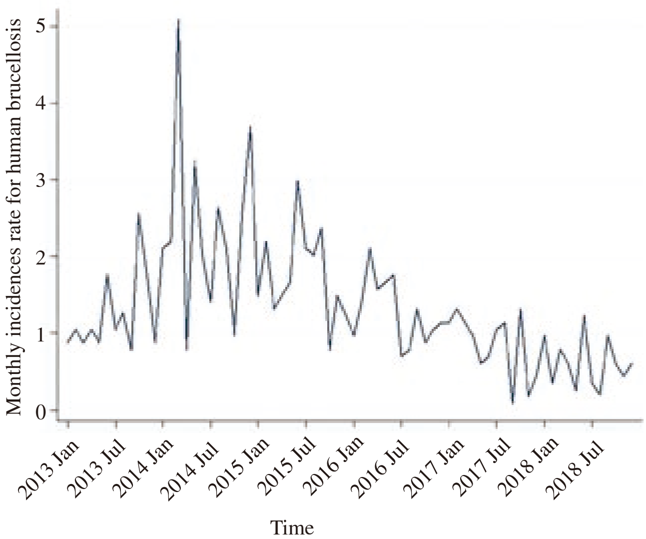

Figure 1. Trend and distribution of human brucellosis incidence (per 100 000 population).

3.2. ARIMA model

This model was created by the data of brucellosis cumulative incidence from 2013 January to 2018 December. It showed that the series was non-stationary in variance, and square root transformation was used. The mean was stationary in this series (d=0). Transformed data were done on all further statistical methods. To define key parameters of the ARIMA model (p, d, q), ACF and PACF plots were drawn and it was proposed that a composition between MA (3) and AR (3) terms could be added in the model (Figure 2 and 3). In fact, the gray space in these graphs shows a 95% confidence interval and the lines that are out of this range sequentially, will be significant. The possible models for ARIMA in this study were ARIMA (3, 0, 3); ARIMA (2, 0, 3); ARIMA (3, 0, 2); ARIMA (4, 0, 3) and ARIMA (3, 0, 4). After this step, the ARIMA (3, 0, 4) model was the most appropriate of all examined models, with lower AIC (25.7) and BIC (43.35) value (Table 1). In fact, in Table 1, for each potential model, a time-series analysis was performed, and the AIC and BIC values of that fitted model were extracted and then compared. The model's fitted cumulative incidence from January 2013 to December 2018 are shown in Figure 4. In general, the predicted brucellosis cases followed an analogous pattern of actual brucellosis cases, indicating that this model could well predict human brucellosis. The MAPE value was 56.20% with standard error 0.009–0.0160 and the portmanteau test for white noise (Q=19.79, P=0.975) for the residuals of the selected model showed that the data were completely modelled. The total cumulative incidence of human brucellosis in 2019 in central Iran was predicted 5.93 per 100 000 people, and monthly incidence was 0.50, 0.44, 0.45, 0.49, 0.55, 0.58, 0.56, 0.51, 0.46, 0.44, 0.45 and 0.49 per 100 000 people, respectively.

Table 1. ARIMA models for human brucellosis incidence in central Iran.

Component	ARIMA (3,0,3)	ARIMA (2,0,3)	ARIMA (3,0,2)	ARIMA (4,0,3)	ARIMA (3,0,4)
Constant	5.26	5.00	5.14	5.10	4.98
L1.AR	0.21	0.05	0.09	-0.60	2.29
L2.AR	0.02	0.81	0.46	0.64	-2.22
L3.AR	0.61	-	0.29	0.55	0.90
L4.AR	-	-	-	0.17	-
L1.MA	-0.22	-0.02	-0.12	0.62	-2.51
L2. MA	0.20	-0.73	-0.28	-0.50	3.06
L3. MA	-0.36	0.28	-	-0.13	-1.84
L4. MA	-	-	-	-	0.47
AIC	32.14	27.41	30.97	30.91	25.72
BIC	52.63	46.21	49.18	51.40	43.35

AR (autoregressive), MA (moving average), AIC (Akaike information criteria), BIC (Bayesian information criterion), L (Lag). "-" not applicable.

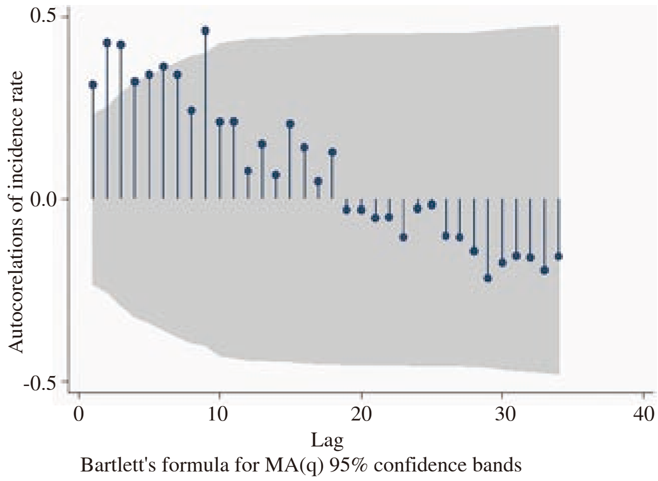

Figure 2. Autocorrelation functions (ACF) correlogram plot of the trend in human brucellosis incidence.

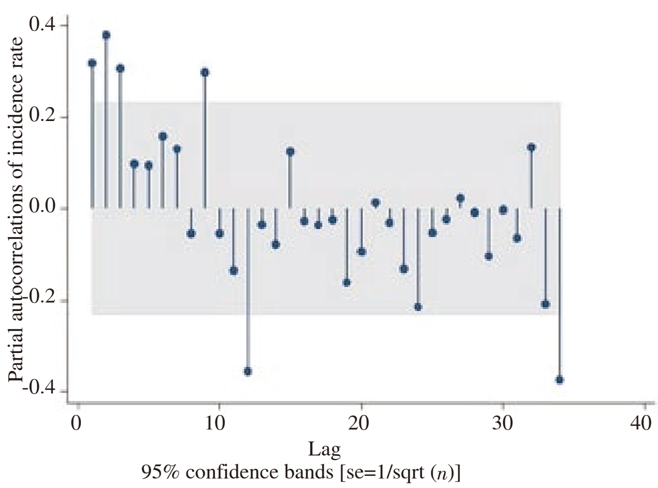

Figure 3. Partial autocorrelation functions (PACF) plots of the trend in human brucellosis incidence.

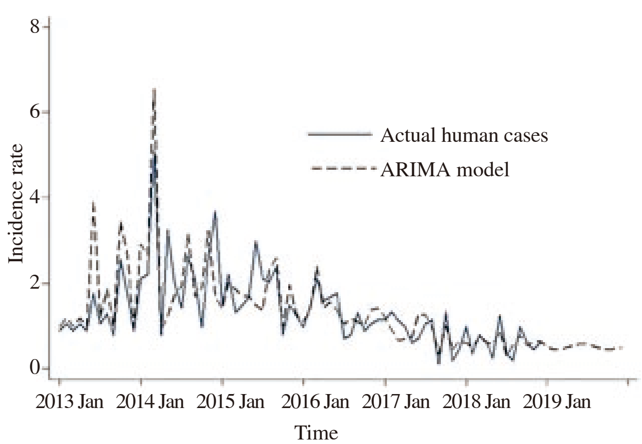

Figure 4. Observed reported cases of human brucellosis and predicted values based on final selected ARIMA model in central Iran.

4. Discussion

In this study, monthly cumulative incidences of brucellosis in Yazd province were forecast by an established ARIMA (3, 0, 4) model. The core benefit of this model is that it takes into attention seasonal differences, which might be valuable in predicting future monthly incidences[19,21]. Human brucellosis cases, were most common in the months of March, April, May, and June. It was observed that the disease incidence had a periodic and seasonal trend with the peak in March and the lowest point in December. In Yazd, goat and sheep are considered as the primary source of brucellosis. It seems that the transition from winter to spring and summer, and the subsequent increase in the temperature and reduced moisture content leads to increased cases of human brucellosis by affecting pasture conditions in terms of animal husbandry, cattle breeding and increased contact between humans and animals. Other studies reported seasonal distribution of human brucellosis[4,22], which is consistent with this study.

Public health surveillance is an instrument used for assessing the health conditions and manner of the population, administrated by ministries of health. The key objective of surveillance is to provide information for action and enable policymakers to lead and manage more efficiently by providing up-to-date, supportive evidence[23]. As a customary public health activity, epidemiological surveillance of transmissible infections is a standard procedure and model predictions provide further application to public health surveillance data sources[24]. Time analysis of infections data may serve as a number of objectives. Frequently, the chief applications consist of forecasting and the progress of alert systems to discover periods or locations where contagion exceeds some thresholds[25]. In this research, the ARIMA model had closely fitted incidences of brucellosis, giving it high accuracy. Therefore, the ARIMA (3, 0, 4) model can be used to forecast the next twelve months' brucellosis incidence in Yazd province. According to the findings of this study, a decreasing incidence in Yazd province in 2019 is forecast. Reductions in the incidence of human brucellosis may be associated with an increasing public awareness of the ways of transmission and prevention through the existence of a brucellosis national surveillance system and the implementation of more control programs, such as vaccination of livestock by the Yazd Veterinary Organization over recent years. The results might be attributed to public health planning and resource allocation.

ARIMA models have been usually used in econometrics; though, their usage is increasing in health fields. Anwar *et al.* applied the ARIMA (4, 1, 1) (1, 0, 1) 12 model to predict future trends in malaria incidence in Afghanistan[26]. Zhong *et al.* indicated ARIMA (2, 1, 0) models to predicting the incidence of pneumoconiosis in Nanjing[27]. Liu *et al.* used ARIMA (0, 3, 1) to predict hemorrhagic fever incidence in China[24]. Yang *et al.* found that the ARIMA (0, 2, 1) could forecast incidence of human brucellosis within a short-term in China[14]. Lee *et al.* evaluated relationship using the ARIMA model and negative binomial regression (NBR) model in forecasting bovine and human brucellosis in Korea[28]. The strengths of this study are that using statistical modeling creates a predictive instrument that can be applied to forecast brucellosis cases at a national level according to information from a passive surveillance system that can be useful for public health projecting and resource allocation.

The limitation of this study is that we assumed that size populations were stable on an annual basis, while not completely accurate. On the other hand, this can lead to a non-differential bias since the comparatively large denominators had negligible effect on the monthly cumulative incidence as a covariate in the model.

In summary, the study showed that the ARIMA (3, 0, 4) model can be used to predict human brucellosis patterns in Yazd province, complementing current surveillance systems, and may be better for health policy-makers and planners. The model provides a means to better knowledge of human brucellosis dynamics in a resource-limited context with minimal data input, yielding predictions that can be applied for public health scheming nationwide. Time-series analysis with ARIMA models are appropriate for short-term forecasting, so we suggest in future studies, other time-series models that do not have this limitation can be used for long-term prediction of time series in brucellosis.

Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgement

This study was derived from MPH thesis from Health Policy Research Center, Institute of Health, Shiraz University of Medical Sciences, Shiraz, Iran. It is necessary to appreciate all who worked with us in this study, especially staff who worked in the brucellosis surveillance system and the Health Deputy of Yazd University of Medical Sciences. Furthermore, the authors hereby would like to express their gratitude and appreciation to Dr. Behnam Honarvar, Associated Professor of Ccommunity Mmedicine at Shiraz University of Medical Sciences, who advised and collaborated on implementing this study in different stages and sections. This study was conducted using existing data from the primary health care system and did not impose additional costs.

Authors' contributions

VR, and SB conceived and designed the study. VR, KR, and SH were responsible for literature search and screening. VR, ATF, and SH were responsible for data collection and analyses. VR, SB, KR, SH and ATF, contributed to data interpretation. VR,SH, ATF drafted the manuscript and SB, KR, critically revised the manuscript.

References

[1] Adetunji SA, Ramirez G, Foster MJ, Arenas-Gamboa AM. A systematic review and meta-analysis of the prevalence of osteoarticular brucellosis. *PLoS Negl Trop Dis* 2019; **13**(1): 1-21.

[2] Franc K, Krecek R, Häsler B, Arenas-Gamboa A. Brucellosis remains a neglected disease in the developing world: A call for interdisciplinary action. *BMC Public Health* 2018; **18**(1): 125-133.

[3] Patra S, Ke V, Tellapragada C, Mukhopadhyay C. Human brucellosis: an experience from a tertiary care hospital in southern India. *Trop Doct* 2018; **48**(4): 368-372.

[4] Rahmanian V, Rahmanian K, Mansoorian E, Jahromi AS, Madani A. Epidemiological characteristics and temporal trend of human and bovine brucellosis cases, Southern Iran, 2009-2016. *Pak J Med Health Sci* 2018; **12**(1): 488-495.

[5] Alsaif M, Dabelah K, Girim H, Featherstone R, Robinson JL. Congenital brucellosis: A systematic review of the literature. *Vector Borne Zoonotic Dis* 2018; **18**(8): 393-403.

[6] Zhao M, Huang F, Zhang A, Zhang B, Zeng L, Xu J, et al. Congenital brucellosis in a Chinese preterm neonate: A case report. *J Int Med Res* 2019; **47**(5): 2296-301.

[7] Cash-Goldwasser S, Maze MJ, Rubach MP, Biggs HM, Stoddard RA, Sharples KJ, et al. Risk factors for human brucellosis in northern Tanzania. *Am J Trop Med Hyg* 2018; **98**(2): 598-606.

[8] Jiang W, Chen J, Li Q, Jiang L, Huang Y, Lan Y, et al. Epidemiological characteristics, clinical manifestations and laboratory findings in 850 patients with brucellosis in Heilongjiang Province, China. *BMC Infect Dis* 2019; **19**(1): 439.

[9] Marvi A, Asadi-Aliabadi M, Darabi M, Abedi G, Siamian H, Rostami-Maskopaee F. Trend Analysis and Affecting Components of Human Brucellosis Incidence During 2006 to 2016. *Med Arch* 2018; **72**(1): 17-21.

[10]Piroozi B, Moradi G, Safari H, Mohamadi P, Alinia C, Shirzadi MR, et al. Incidence, mortality, and burden of human brucellosis and its geographical distribution in Iran during 2009-2015. *Iran J Public Health* 2019; **48**(1): 20-27.

[11]Madut NA, Nasinyama GW, Muma JB, Sube KL, Ocan M, Muwonge A, Godfroid J, et al. Prevalence of brucellosis among patients attending Wau Hospital, South Sudan. *PloS One* 2018; **13**(6): e0199315.

[12]Zeinali M, Shirzadi M, Haj rasuliha H. *National guideline for brucellosis*

control. 2nd ed. Tehran: Raz Nahan Publishing; 2012, p. 9-11.

[13]Nabavi M, Hatami H, Jamaliarand H. Epidemiological, risk factors, clinical, and laboratory features of brucellosis in the Southwest of Iran within 2009–2015. *Int J Prev Med* 2019; **10**(1): 108.

[14]Yang L, Bi ZW, Kou ZQ, Li XJ, Zhang M, Wang M, et al. Time series analysis on human brucellosis during 2004–2013 in Shandong Province, China. *Zoonoses & Public Health* 2015; **62**(3): 228-235.

[15]Wu W, An SY, Guan P, Huang DS, Zhou BS. Time series analysis of human brucellosis in mainland China by using Elman and Jordan recurrent neural networks. *BMC Infect Dis* 2019; **19**(1): 414.

[16]Wang H, Tian C, Wang W, Luo X. Time-series analysis of tuberculosis from 2005 to 2017 in China. *Epidemiol Infect* 2018; **146**(8): 935-939.

[17]Zhao Y, Ge L, Zhou Y, Sun Z, Zheng E, Wang X, et al. A new Seasonal Difference Space-Time Autoregressive Integrated Moving Average (SD-STARIMA) model and spatiotemporal trend prediction analysis for hemorrhagic fever with renal syndrome (HFRS). *PloS One* 2018; **13**(11): 1-20.

[18]Cortes F, Martelli CMT, de Alencar Ximenes RA, Montarroyos UR, Junior JBS, Cruz OG, et al. Time series analysis of dengue surveillance data in two Brazilian cities. *Acta Trop* 2018; **182**: 190-197.

[19]Ren JP, Chen ZP, Sun JM, Chen EF, Shi XG, Zhang R, et al. The time-series analysis of human rabies in China. *Zhongguo Ren Shou Gong Huan Bing Za Zhi* 2018; **34**(3): 239-242.

[20]Dohoo IR, Martin W, Stryhn H. *Veterinary epidemiologic research*. 2nd ed. Charlottetown: AVC Incorporated; 2009, p. 637-651.

[21]Silawan T, Singhasivanon P, Kaewkungwal J, Nimmanitya S, Suwonkerd W. Temporal patterns and forecast of dengue infection in Northeastern Thailand. *Southeast Asian J Trop Med Public Health* 2008; **39**(1): 90.

[22]Li YJ, Li XL, Liang S, Fang LQ, Cao WC. Epidemiological features and risk factors associated with the spatial and temporal distribution of human brucellosis in China. *BMC Infect Dis* 2013; **13**(1): 547.

[23]Nsubuga P, White ME, Thacker SB, Anderson MA, Blount SB, Broome CV, et al. Public health surveillance: A tool for targeting and monitoring interventions. In: *Disease control priorities in developing countries (English)*. New York, Oxford University Press; 2006, p. 997-1018.

[24]Liu Q, Liu X, Jiang B, Yang W. Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. *BMC Infect Dis* 2011; **11**(1): 218.

[25]Chen CW, Khamthong K, Lee S. Markov switching integer-valued generalized auto-regressive conditional heteroscedastic models for dengue counts. *J R Stat Soc Ser C Appl Stat* 2019; **68**(4): 963-983.

[26] Anwar MY, Lewnard JA, Parikh S, Pitzer VE. Time series analysis of malaria in Afghanistan: using ARIMA models to predict future trends in incidence. *Malar J* 2016; **15**(1): 566.

[27]Zhong Q, Sui Y, Pang Y, Song H. ARIMA models to predict new-diagnosing cases of pneumoconiosis in Nanjing. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi* 2014; **32**(3): 211-213.

[28]Lee HS, Her M, Levine M, Moore GE. Time series analysis of human and bovine brucellosis in South Korea from 2005 to 2010. *Prev Vet Med* 2013; **110**(2): 190-197.