

SCIENTIFIC REPORTS



OPEN

The temporal distribution of new H7N9 avian influenza infections based on laboratory-confirmed cases in Mainland China, 2013–2017

Zuiyuan Guo¹, Dan Xiao², Dongli Li¹, Yayu Wang¹, Tiecheng Yan¹, Botao Dai³ & Xiuhong Wang¹

In this study, estimates of the growth rate of new infections, based on the growth rate of new laboratory-confirmed cases, were used to provide a statistical basis for in-depth research into the epidemiological patterns of H7N9 epidemics. The incubation period, interval from onset to laboratory confirmation, and confirmation time for all laboratory-confirmed cases of H7N9 avian influenza in Mainland China, occurring between January 2013 and June 2017, were used as the statistical data. Stochastic processes theory and maximum likelihood were used to calculate the growth rate of new infections. Time-series analysis was then performed to assess correlations between the time series of new infections and new laboratory-confirmed cases. The rate of new infections showed significant seasonal fluctuation. Laboratory confirmation was delayed by a period of time longer than that of the infection (average delay, 13 days; standard deviation, 6.8 days). At the lags of -7.5 and -15 days, respectively, the time-series of new infections and new confirmed cases were significantly correlated; the cross correlation coefficients (CCFs) were 0.61 and 0.16, respectively. The temporal distribution characteristics of new infections and new laboratory-confirmed cases were similar and strongly correlated.

In March 2013, a new type of avian influenza virus, H7N9, was isolated for the first time in China¹. In the spring of that year, an outbreak of H7N9 influenza occurred in 10 provinces in central and eastern China². Until June 30, 2017, the H7N9 virus has caused 5 nationwide outbreaks: January 2013 to September 2013 (134 cases, 45 deaths), October 2013 to September 2014 (306 cases, 131 deaths), October 2014 to September 2015 (219 cases, 102 deaths), October 2015 to September 2016 (114 cases, 47 deaths), and October 2016 to June 2017 (1058 cases, 221 deaths)^{3,4}. The average case fatality risk of H7N9 infection is approximately 40%³, which is higher than the case fatality risk of SARS (11%)⁵ and lower than that of H5N1 (70%)². The fifth epidemic was the most severe; cases were diagnosed in 20 provinces and the total number of cases exceeded the sum of the previous 4 epidemics³. These 5 epidemics occurred mostly in China's Yangtze and Pearl River Deltas³. The genetic characteristics of the viral strains were similar in all 5 epidemics^{6,7}. The main sites of infection were live poultry markets^{8,9}. The main route of transmission was human contact with infected poultry or virus-contaminated environments; no connections were demonstrated between patients⁷, and only limited, non-sustained human-to-human transmission within individual regions was observed^{10,11}.

A typical confirmed case of H7N9 avian influenza, from infection to diagnosis, progresses through the following 3 stages: (1) From infection by exposure to the pathogen to symptom presentation, after an incubation period; (2) from symptom presentation to hospital admission and treatment; and (3) from hospitalization to laboratory confirmation by a state-certified network laboratory. The state-certified network laboratory reports each case, via the pneumonia of unknown etiology (PUE) surveillance system, to the Chinese Center for Disease Control and Prevention (China CDC)¹². A study conducted by Cowling *et al.*² showed that the mean incubation period of H7N9 is 3.1 days and follows a Weibull distribution, and that the median interval from onset to laboratory

¹Department of Disease Control, Center for Disease Control and Prevention of Shenyang Military Region, Shenyang, China. ²China National Clinical Research Center for Neurological Diseases, Beijing TianTan Hospital, Beijing, China. ³The Emergency Center for Animals Disease of Liaoning Province, Shenyang, China. Dongli Li, Yayu Wang, Tiecheng Yan and Botao Dai contributed equally to this work. Correspondence and requests for materials should be addressed to X.W. (email: 153712130@qq.com)

confirmation is 8.3 days and follows a log-normal distribution. These study results provide an important basis for our further investigations.

Because infection occurs before the diagnosis is made, assessing the epidemic trend using the growth rate of confirmed cases is imprecise; we consider the growth rate of new infections an important indicator of the seriousness of an epidemic. However, the exact time that a patient becomes infected is difficult to ascertain, mainly because patients themselves do not know exactly when they became infected; most can give only a rough estimate, and some are unclear even about this. Because the exact time that infection occurs cannot be obtained through epidemiological surveys, the use of mathematical models to arrive at a reasonable estimate is a good choice.

A number of researchers have used mathematical models to provide a quantitative description of the epidemiological characteristics of H7N9 epidemics^{13–19}. For example, Zhang *et al.*¹³ established a dynamic model including migratory birds, resident birds, domestic poultry, and human populations; they concluded that migrant birds were most likely the original source of infection. Lin *et al.*¹⁴ modelled chicken-to-chicken transmission and found that environmental transmission via viral shedding by infected chickens contributed to the spread of the virus. These studies help us to understand the prevalence of the virus among poultry and the mechanisms of its transmission from poultry to humans. According to the transmission route of H7N9 avian influenza, each infected individual is relatively independent, the dynamic model of infectious disease spread cannot be used. Quantitative analysis of epidemic spread within human communities is crucial for a deeper understanding of the mechanisms by which epidemics spread. However, no studies have yet been published on the evolutionary mechanism of the entire process from infection to symptom onset to diagnosis.

The present study used statistical data on the temporal distribution of laboratory-confirmed cases and applied the theory of stochastic processes to reveal the transmission mechanisms of this epidemic among human communities. In the study, we performed a quantitative analysis of the temporal distribution pattern for the expected values and 95% confidence intervals (CIs) of the growth rate of new infections. Therefore, this study has important implications for a deeper understanding of the onset and progression of H7N9 avian influenza epidemics and for a more precise description of its temporal distribution patterns. Its results can be used for the timely assessment of the effects of a series of government-instituted interventions and for exploring causal relationships between the time series of new infections and new confirmed cases, and provide a statistical basis for in-depth analysis of the impact of human and natural factors on the epidemic.

Methods

Data sources. The confirmation dates of all new H7N9 confirmed cases between January 2013 and January 2017 in Mainland China were obtained from the China CDC; data between February and June 2017 were obtained from the EMPES-i georeferenced disease data repository compiled by the Food and Agricultural Organization⁴. All cases were confirmed as H7N9 by local and/or provincial influenza network laboratories. H7N9 avian influenza is a category B infectious disease according to China's notifiable infectious diseases classification; hence, cases confirmed by network laboratories need to be reported immediately using the PUE surveillance system¹².

Temporal distribution of the number of new infections. Our goal was to treat all laboratory-confirmed cases between January 2013 and June 2017 as a sample with number N , and to use this sample to estimate the temporal distribution for the growth rate of new infections.

First, we used Bayes theorem to analyze the probability of infection and subsequent diagnosis given that each confirmed case experiences 2 stages between infection and diagnosis: The first is between infection and onset (the incubation period), the second is between onset and laboratory confirmation. The incubation period follows a Weibull distribution², with a probability density function expressed as $f(\cdot)$. The onset-to-laboratory confirmation interval follows a log-normal distribution², with a probability density function expressed as $g(\cdot)$. These 2 time periods are mutually independent. The conditional probability, $\delta(T|t_i)$, indicating the probability of an infected individual receiving laboratory-confirmation after a duration of T under the condition that infection occurred at time point t_i , was expressed as convolution equation (1),

$$\delta(T|t_i) = \int_0^T f(s)g(T-s)ds \quad 0 < s < T \quad (1)$$

where s denotes the duration of the incubation period and $T-s$ indicates the onset-to-laboratory-confirmation duration. As laboratory confirmation could happen at any time of the day, T includes a range of values, $[T_0, T_1]$, where T_0 and T_1 denote the beginning and the end of the day, respectively. The conditional probability $\tilde{\delta}(T|t_i)$ was used to indicate the probability of an infected individual having received laboratory-confirmation during the period $T_0 \sim T_1$ given that infection occurred at time point t_i (equation 2),

$$\tilde{\delta}(T|t_i) = \int_{T_0}^{T_1} \delta(T|t_i)dT \quad (2)$$

where $p(t_i)$ denotes the probability that a susceptible person is infected at time t_i and $P(t_i, T_i)$ denotes the probability of receiving laboratory confirmation during day T_i given that a susceptible person is infected at time t_i and after a period of T_i . Using Bayes theorem, $P(t_i, T_i)$ was expressed as equation (3),

$$P(t_i, T_i) = p(t_i)\tilde{\delta}(T_i|t_i) \quad (3)$$

Thus, we could construct a likelihood function $L(\cdot)$:

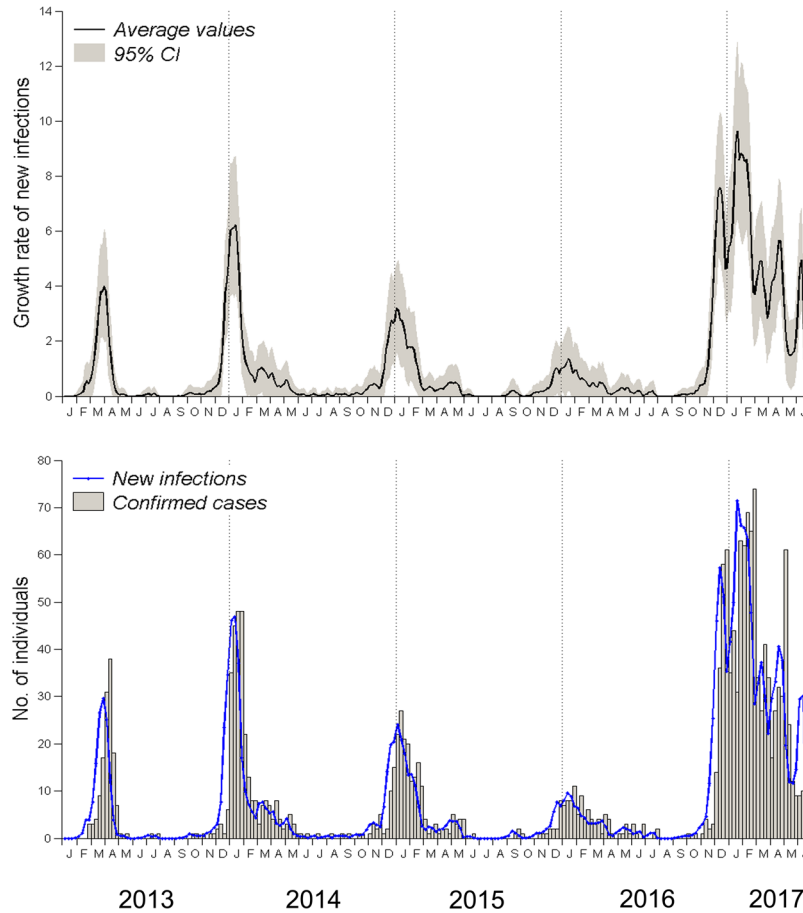


Figure 1. Temporal distribution of new infections and new confirmed cases of avian influenza A (H7N9) virus infection. **(A)** Temporal distribution of the growth rates of new infections. The black solid line represents the mean values, the grey area represents 95% confidence intervals. **(B)** Comparison of the number of new infections and new confirmed cases. The histograms show the number of new reported confirmed cases; the blue line represents the expected number of new infections.

$$\begin{aligned}
 L(X_1, \dots, X_N) &= \prod_{i=1}^N P(t_i, T_i) \\
 &= \prod_{i=1}^N p(t_i) \delta(T_i | t_i) \\
 &= \prod_{i=1}^N p(t_i) \prod_{i=1}^N \delta(T_i | t_i)
 \end{aligned} \tag{4}$$

Where X_i ($i = 1, \dots, N$) denotes the i^{th} infection that occurs. We then calculated the growth rate of new infections by analyzing $\prod_{i=1}^N p(t_i)$ using stochastic process theory.

The main transmission route of H7N9 is human contact with infected poultry or virus-contaminated environments; the virus is not continually spread among people, rather, susceptible individuals are randomly infected. The frequency of contact is variable at different times of the year, so the growth rate of new infections is also variable, being lower in summer than that in winter and spring (Fig. 1). We assumed that continuous emergence of infections represents a series of independent random events.

$\{N(\tau), \tau \geq 0\}$ denotes the process of counting H7N9 infections within a short time period τ . Based on the analysis above, $N(\tau)$ can be regarded as a Poisson process with rate λ , and the time interval of any 2 consecutive infections will follow independent exponential distribution²⁰. For a time interval of length τ , the probability $P(X_1, \dots, X_m; \lambda, \tau)$ that m infections will occur consecutively at time points s_1, \dots, s_m can be expressed as:

$$\begin{aligned}
 P(X_1, \dots, X_m; \lambda, \tau) &= P\{m \text{ infections in } [0, s_m]\}P\{\text{none in } (s_m, \tau]\} \\
 &= \lambda e^{-\lambda s_1} \lambda e^{-\lambda(s_2-s_1)} \dots \lambda e^{-\lambda(s_m-s_{m-1})} e^{-\lambda(\tau-s_m)} \\
 &= \lambda^m e^{-\lambda\tau}
 \end{aligned} \tag{5}$$

Because the growth rate of new infections varied during the period 2013–2017, and for reasons of convenience in the calculations, we assumed a constant growth rate over a short period of time. Hence, the rate changed over the course of a few consecutive time intervals of moderate length (if the interval is too long the result will be inaccurate, and if it is too short it will increase the amount of calculation required). We used intervals of an eighth of a month. There were 54 months between January 2013 and June 2017, giving 432 intervals. The growth rate of new infections in each interval was expressed as λ_k ($k = 1, \dots, 432$), and the length was a constant value, τ . The entire process of an infection appearing can be viewed as a continuous-time Markov chain with different rates and as a pure birth process²⁰ with states of X_0, X_1, \dots, X_N . Hence, $p(t_i)$ can be converted to the probability of when the system transitions from state X_{i-1} to state X_i at time t_i , as follows:

$$p(t_i) = p\{X(t_i) = i | X(t_{i-1}) \leq t < t_i = i - 1\} \tag{6}$$

$$p(t_i) = \begin{cases} \lambda_k e^{-\lambda_k(t_i - t_{i-1})} & \lambda_k: [t_{i-1}, t_i] \\ e^{-\lambda_{k-1}(t' - t_{i-1})} \lambda_k e^{-\lambda_k(t_i - t')} & \lambda_{k-1}: [t_{i-1}, t'], \lambda_k: [t', t_i] \end{cases} \tag{7}$$

Based on the above analysis, $\prod_{i=1}^N p(t_i)$ is associated with λ_k (equation 8).

$$\begin{aligned} \prod_{i=1}^N p(t_i) &= \prod_{k=1}^{432} P(X_1, \dots, X_{m_k}; \lambda_k, \tau) \left(\sum_{k=1}^{432} m_k = N \right) \\ &= \prod_{k=1}^{432} \lambda_k^{m_k} e^{-\lambda_k \tau} \end{aligned} \tag{8}$$

We calculated the value of λ_k using the maximum likelihood; equation (9) was derived by incorporating equation (8) into equation (4).

$$\begin{aligned} L(\hat{\lambda}, \tau) &= \max_{\lambda, \tau > 0} L(\lambda, \tau) \\ &= \max_{\lambda, \tau > 0} \prod_{i=1}^N p(t_i) \prod_{i=1}^N \delta(T_i | t_i) \\ &= \max_{\lambda, \tau > 0} \prod_{k=1}^{432} \lambda_k^{m_k} e^{-\lambda_k \tau} \prod_{i=1}^N \delta(T_i | t_i) \end{aligned} \tag{9}$$

After logarithmic transformation of equation (9), a very concise result could be obtained by taking the partial derivative of λ_k :

$$\lambda_k = \frac{m_k}{\tau} \tag{10}$$

Where λ_k is the growth rate of new infections.

On the basis of the above analysis, we calculated expected values and 95% CIs for λ_k as follows: For each confirmed case, according to a uniform distribution, we randomly selected a time point on the day of laboratory confirmation as the confirmation time. Then, the Markov chain Monte Carlo method was applied to generate random numbers that fitted the probability density function $\delta(\cdot)$ ²¹; these numbers represent the lengths of time from infection to laboratory confirmation. As the time of laboratory confirmation is known, the estimated time of infection can be calculated. After computing the infection times of all confirmed cases, counting was performed on 432 time intervals with length τ to sequentially obtain m_k , and hence to work out the growth rate of new infections λ_k of one calculation. The bootstrap method was used to repeat this step 1000 times to obtain the mean values and 95% CIs for λ_k . The above method was implemented using MATLAB and Statistics Toolbox Release 2012a (The MathWorks, Inc., Natick, Massachusetts, USA).

Correlation test of time series for new infections and new confirmed cases. Based on the above calculations, the expected number of new infections in quarters of a month can be obtained. These numerical values constitute the time series $X = \{X(t), t \geq 0\}$, and new confirmed cases will constitute the time series $Y = \{Y(t), t \geq 0\}$. X was taken as the covariant variable and Y as the dependent variable. The aim was to test the CCF at a lag of k , $r_k(X, Y)$, to verify the correlation between the 2 and to analyze whether the former was causing fluctuations in the latter.

$$r_k(X, Y) = \frac{\sum(X_t - \bar{X})(Y_{t-k} - \bar{Y})}{\sqrt{\sum(X_t - \bar{X})^2} \sqrt{\sum(Y_t - \bar{Y})^2}} \quad k = 1, 2, \dots \tag{11}$$

As these 2 time-series have strong autocorrelation (Fig. 1), if the correlation of the 2 processes is evaluated directly by calculating the CCFs, a false conclusion may be obtained (spurious correlation)²². Hence, the autocorrelation should be extracted from their respective series—pre-whitening is a precise tool for achieving this aim²². As the 2 series showed significant seasonal fluctuations, seasonal differencing was performed. After differencing, the series remained non-stationary; hence, first order differencing was performed to obtain 2 stationary

Years	Jan	Feb	Mar	Apr	May	Jun
2013	0	9	80	42	1	1
2014	147	27	27	15	8	1
2015	76	34	8	10	9	1
2016	33	19	13	5	8	4
2017	229	205	122	141	58	63
Years	Jul	Aug	Sep	Oct	Nov	Dec
2013	2	0	1	3	5	69
2014	1	2	2	4	11	61
2015	0	0	4	1	5	23
2016	3	0	1	4	45	190

Table 1. The average numbers of new infections per month during the period 2013~2017.

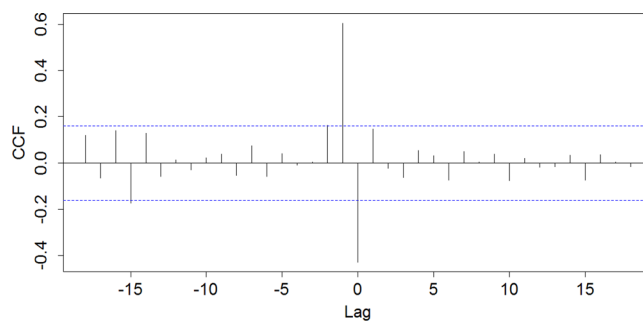


Figure 2. The cross-correlation graph for differenced and pre-whitened time series of new infections and new confirmed cases. The black vertical lines represent the cross-correlation coefficients; the two blue dotted lines represent 95% confidence interval.

time-series with a mean value of 0. Then, the same filter, based on the first series (an autoregressive model), was applied to pre-whiten the 2 time-series, followed by calculating the CCFs of the pre-whitened series²². The above method was implemented using R, version 3.1.0 (<http://www.R-project.org/>).

Results

Temporal distribution of new infections. Given that the incubation period of H7N9 fits a Weibull distribution (mean 3.1 days, standard deviation 1.4 days)², we could work out the shape (2.4) and scale (3.5) parameters. As the onset-to-laboratory-confirmation interval fits a log normal distribution (median 8.3 days, 95% CI 7.3–9.5 days)², we could work out the mean (2.1 days), standard deviation (0.6). Calculations were performed according to the methods described by JD Chen²³.

The temporal distributions of the expected number of new infections and new confirmed cases showed a similar patterns of variation. Both displayed significant seasonal fluctuations and autocorrelation. Figure 1 shows that the first cases of infection appeared in Mainland China in February 2013. The number of infections increased drastically, reached a peak in March 2013, and then decreased drastically. From May to September 2013, infections were sporadic. The number of infections gradually increased from October through November 2013, then markedly increased in December, reaching a peak in January 2014, thereafter drastically decreasing. The epidemic stabilized from March through September 2014. Thereafter, the pattern of temporal distribution was essentially similar to what was observed in 2014, i.e. a high incidence in winter and spring with a low incidence in summer and autumn. The fifth epidemic was the most serious, with the number of infections increasing drastically in October 2016 and reaching a peak in January 2017; an estimated 229 new infections occurred in January 2017. Thereafter, the number of cases gradually decreased, but still fluctuated at a high level. From 2013 through May 2017, the expected numbers of infections annually were 212, 305, 171, 324, and 755. The average numbers of new infections per month are presented in Table 1.

Correlation test of time series for new infections and new confirmed cases. Seasonal and first order differencing were performed on the time series of new infections and new confirmed cases, respectively. Then, pre-whitening (a nineteenth-order-autoregressive model was selected according to the Akaike information criteria) was performed, followed by computation of the CCFs. Figure 2 shows that 2 positive CCFs were obtained: When the lags were -2 and -1 , the CCFs were 0.16 and 0.61, respectively. When the lag was 0, the CCF obtained was -0.43 . In addition, when the lag was -15 , there was a marginally significant negative CCF, which we believe to be a false alarm. This is because when calculating the CCFs for a total of 37 samples, we expect there to be an average of $37 \times 0.05 = 1.85$ false alarms. Based on the mean value of 1 interval (7.5 days), the mean duration from infection to laboratory confirmation, and the standard deviation, the 2 significant positive CCFs are at -2 and -1 time units.

Discussion

Since the outbreak of the H7N9 avian influenza epidemic in multiple provinces across China in March 2013, researchers began conducting scientific research from different epidemiological perspectives. For example, Yuan *et al.*²⁴ conducted sampling in live-poultry markets; detection of viral RNA demonstrated that live-poultry markets were the main sites at which H7N9 infection. Lin *et al.*¹³ and Zhang *et al.*¹⁴ established mathematical models to explore viral transmission mechanisms within different populations. Cowling *et al.*², Wang *et al.*³, and Bui *et al.*¹⁸ provided detailed descriptions on the temporal, spatial, and population distributions of the epidemics. These studies laid the foundation for investigating the epidemiological characteristics of the H7N9 epidemics. Compared with these published studies, the innovative points lie in our analysis of the transmission mechanisms in the human population from a stochastic perspective, according to the transmission characteristics of H7N9. The occurrence of infection was treated as a random event produced by a Poisson process, and the rate of infection changed with time. Therefore, the entire epidemic process can be regarded as the connection of multiple Poisson processes occurring at different rates. We established a relationship between the probability of infection and a continuous time Markov chain; the growth rate of new infections was then obtained.

The initial stage of an infectious disease is effective contact with the pathogen. The growth rate of new infections reflects the frequency at which susceptible individuals come into contact with the virus: The more frequent the contact, the faster the rate of new infections increases. Therefore, we believe that this rate is the earliest and most accurate reflection of the severity of the epidemic. Although changes in the number of confirmed cases can reflect the temporal distribution features of the epidemic, laboratory confirmation of the diagnosis is delayed by a period of time greater than that of infection; in our study, the average delay was 13 days (standard deviation, 6.8 days). Therefore, evaluating the epidemic's development trend using the temporal distribution of confirmed cases is imprecise. This study aimed to analyze the epidemic trend more accurately and to establish a more precise statistical foundation for analyzing the impact of various prevention and control measures adopted by China, and meteorological and environmental factors, on the epidemic. For example, some influencing factors—such as closing and regular disinfection of live poultry markets, atmospheric temperature, rainfall, and the number of migratory birds—can be evaluated more timely and accurately by judging the growth rate of new infections.

From Fig. 1B, we intuitively observe that there is a relatively strong correlation between the time series of infections and confirmed cases. However, this might not be a true correlation; a spurious correlation may exist. By calculating the CCFs after pre-whitening, we found significant positive correlation between the 2 time series, with the lags of -1 and -2 time units. Based on this, we can conclude that the occurrence of infections will lead to the occurrence of confirmed cases. Furthermore, the results verified that the establishment of random models is consistent with objective facts. With a lag of 0, a significant negative CCF is obtained (Fig. 2); this does not imply a true negative correlation between the 2 time series. Figure 1B shows that the time series of infections was smoother than that of confirmed cases, with opposite trends of change at certain time points.

Predicting epidemic trends through a mathematical model can help us to adopt more effective preventive measures to control potential epidemics. However, our model can only estimate the growth rate of infections in the past based on existing confirmed cases, i.e., it can only perform a retrospective analysis of epidemics; it is not predictive. More in-depth research on the transmission mechanisms of epidemics is required to establish new mathematical models that may require the inclusion of diverse factors such as immunization, closure of live-poultry markets, regular environmental disinfection, and other human interventions.

References

- Gao, R. *et al.* Human infection with a novel avian-origin influenza A (H7N9) virus. *N Engl J Med.* **368**, 1888–97 (2013).
- Cowling, B. J. *et al.* Comparative epidemiology of human infections with avian influenza A H7N9 and H5N1 viruses in China: a population-based study of laboratory-confirmed cases. *Lancet.* **382**, 129–37 (2013).
- Wang, X. L. *et al.* Epidemiology of avian influenza A H7N9 virus in human beings across five epidemics in mainland China, 2013–17: an epidemiological study of laboratory-confirmed case series. *Lancet Infect Dis.* **8**, 822–32 (2017).
- EMPRES. Disease Events. <http://empres-i.fao.org/eipws3g/>. (Accessed 1 July 2017).
- Yeung, M. C. & Xu, R. H. SARS: epidemiology. *Respirology.* **8**, suppl s9–s14 (2003).
- Zhang, Q. Y. *et al.* H7N9 influenza viruses are transmissible in ferrets by respiratory droplet. *Science.* **341**, 410–4 (2013).
- Zhou, L. *et al.* Sudden increase in human infection with avian influenza A (H7N9) virus in China, September–December 2016. *Western Pac Surveill Response J.* **8**, 6–14 (2017).
- Peiris, J. S. M. *et al.* Interventions to reduce zoonotic and pandemic risks from avian influenza in Asia. *Lancet Infect Dis.* **16**, 252–8 (2016).
- Wang, C. M. *et al.* Relationship between domestic and wild birds in live poultry market and a novel human H7N9 virus in China. *J Infect Dis.* **209**, 34–7 (2013).
- Chowell, G., Simonsen, L., Towers, S., Miller, M. A. & Viboud, C. Transmission potential of influenza A/H7N9, February to May 2013, China. *BMC Med.* **11**, 214 (2013).
- Wu, Z. Q. *et al.* Epidemiological and virological differences in human clustered and sporadic infections with avian influenza A H7N9. *Int J Infect Dis.* **49**, 9–17 (2016).
- Xiang, N. J. *et al.* Use of national pneumonia surveillance to describe influenza A(H7N9) virus epidemiology, China, 2004–2013. *Emerg Infect Dis.* **19**, 1784–90 (2013).
- Zhang, J. *et al.* Determination of original infection source of H7N9 avian influenza by dynamical model. *Sci Rep.* **4**, 4846 (2014).
- Lin, Q. Y., Lin, Z. G., Chiu, A. P. Y. & He, D. Seasonality of influenza A (H7N9) virus in China-Fitting simple epidemic models to human cases. *PLoS One.* **11**, e0151333 (2016).
- Chen, Y. X. & Wen, Y. X. Global dynamic analysis of a H7N9 avian-human influenza model in an outbreak region. *J Theor Biol.* **367**, 180–8 (2014).
- Virlogeux, V. *et al.* Estimating the distribution of the incubation periods of human avian influenza A (H7N9) virus infections. *Am J Epidemiol.* **182**, 723–9 (2015).
- Virlogeux, V. *et al.* Association between the severity of influenza A (H7N9) virus infections and length of the incubation period. *PLoS One.* **11**, e0148506 (2016).
- Bui, C. M., Gardner, L., MacIntyre, R. & Sarker, S. Influenza A H5N1 and H7N9 in China: A spatial risk analysis. *PLoS One.* **12**, e0174980 (2017).
- Li, X. L. *et al.* Risk distribution of human infections with avian influenza H7N9 and H5N1 virus in China. *Sci Rep.* **5**, 18610 (2015).

20. Ross, S. M. *Introduction to probability models*. (Elsevier, 2011).
21. Ye, F. *Markov chain Monte Carlo method and implemented with R*. (Nanjing University, 2014).
22. Cryer, J. D. & Chan, K. S. *Time series analysis with applications in R*. (Springer, 2008).
23. Chen, J. D. *Survival analysis and reliability*. (Peking University, 2005).
24. Yuan, J. *et al.* Effect of live poultry market closure on avian influenza A (H7N9) virus activity in Guangzhou, China, 2014. *Emerg Infect Dis.* **21**, 1784–93 (2015).

Acknowledgements

We would like to thank Editage [www.editage.cn] for English language editing. The study is supported by National Science and Technology Major Project (2018ZX10713003).

Author Contributions

Guo Z.Y. designed the study, analyzed the data, and wrote most of the manuscript, including quality assurance and control; Xiao D. contributed to writing the manuscript; Li D.L., Wang Y.Y., Yan T.C. and Bao T.D. helped to conduct the literature review; Wang X.H. organized and coordinated all parties to ensure that the study was accomplished.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018