**Changing geographic patterns and risk factors ofavian influenza A(H7N9) human cases in China**

Jean Artois1,\*, Hui Jiang2,\*, Xiling Wang3, Ying Qin2, Morgan Pearcy1, Shengjie Lai2,3,4, Yujing Shi2, Juanjuan Zhang3, Zhibin Peng2, Jiandong Zheng2, Yangni He3, Madhur S Dhingra1,5, Sophie von Dobschuetz6, Fusheng Guo7, Vincent Martin8, Wantanee Kalpravidh7, Filip Claes7, Timothy Robinson6, Simon I. Hay9,10, Xiangming Xiao11,12, Luzhao Feng2, Marius Gilbert1,11,†, Hongjie Yu2,3†

1 Spatial epidemiology Lab (SpELL), Université Libre de Bruxelles, Brussels, Belgium.

2 Key Laboratory of Surveillance and Early-warning on Infectious Disease, Division of Infectious Disease, Chinese Center for Disease Control and Prevention, Beijing, China

3 School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China.

4 WorldPop, Department of Geography and Environment, University of Southampton, Southampton, UK

5 Food and Agriculture Organization of the United Nations, Rome, Italy

6 Livestock Information, Sector Analysis and Policy Branch (AGAL), Food and Agriculture Organization of the United Nations, Rome, Italy

7 Food and Agriculture Organization of the United Nations, Regional Office for Asia and the Pacific, Bangkok, Thailand

8 Food and Agriculture Organization of the United Nations, China Office, Beijing, China

9 Institute for Health Metrics and Evaluation, University of Washington, Seattle, USA

10 Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK

11 Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73019, USA

12 Key Laboratory of Biodiversity Science and Ecological Engineering, Institute of Biodiversity, Fudan University, Shanghai 200433, China

13 Fonds National de la Recherche Scientifique, Brussels, Belgium.

\* Contributed equally to this work

† Corresponding authors: HY (cfetpyhj@vip.sina.com), MG (mgilbert@ulb.ac.be)

**Article Summary Line**

 The fifth epidemic wave of human cases of H7N9 virus showed a geographical range expansion from the highly populated southeastern coast toward mainland China further east and north, with a trend towards a stronger association with poultry distribution risk factors than in any previous epidemic waves.

**Abstract (161 !! <150)**

 The 5th epidemic wave of avian influenza A(H7N9) virus in China in 2016-2017 demonstrated a geographical range expansion and caused more human cases than any previous wave. The factors that may explain the recent range expansion and surge in incidence remain unknown. We investigated the effect of anthropogenic, poultry and wetland information and of market closures on all epidemic waves (1-5). Poultry predictor variables became much more important in the last two epidemic waves than before, supporting the assumption of much wider H7N9 transmission in the chicken reservoir. We show that the future range expansion of H7N9 to northern China may translate into a higher risk of coinciding peaks with those of seasonal influenza, leading to a higher risk of reassortments than before, although the risk is still low so far.

 4652 words!!!

 The third and fourth epidemic waves of avian influenza A (H7N9) human infections in China showed an apparent reduction in incidence compared to spring 2013 and winter 2013/2014 epidemic waves. However, during the winter of 2016/2017, the incidence rose, growing to levels never observed before and reaffirming concerns of a pandemic threat posed by the H7N9 virus (1–3). Since 2013, more than 1582 human cases of H7N9 virus have been reported, mostly located in eastern China, with a case fatality rate ranging between 30% and 40% (4–6).

 The H7N9 virus that caused the first epidemic wave in March 2013 originated from multiple reassortment events of avian influenza viruses from domestic poultry and wild birds (10). Mainly restricted to Yangtze River Delta in eastern China including urban areas of Shanghai, Jiangsu and Zhejiang provinces in the first wave, the spatial range of H7N9 human cases increased during the second wave along the coast into Guangdong province in southern China (11). Over time, phylogeographic inference suggested that H7N9 had become established in separate parts of China during the second and third waves, reassorting with local avian influenza viruses (12,13).

 Humans are not a natural reservoir but an occasional spillover host of H7N9; human cases act as an indicator, presumably reflecting the circulation of H7N9 in poultry (12) and could represents the most effective way to study the spatial distribution of H7N9 virus. Indeed, surveillance in poultry is difficult as the virus thus far had a low pathogenicity in chickens (14,15) and the absence of clinical signs requires active and targeted sampling, which made the characterization of the spatial distribution of the virus reservoir inconclusive. This may change in the future due to the recent evolution of a highly pathogenic strain of H7N9 (16–18).

Three sets of factors may have a significant influence on spatial variation in H7N9 incidence, and were considered in this study.

 First, we include live-poultry market (LPM) density and human population density. Visits to LPMs are the main known risk factor of H7N9 infection at the human case level (19–21) and LPMs represent a key interface between human and poultry. At a higher level, LPM networks may also support the spread and persistence of H7N9 virus through the network of LPM and poultry farms linked by trade (22). In previous studies, we showed that a high density of LPM in some specific areas could regionally increase the risk of H7N9 infection at the market level (23), which translates into higher risk at the county level measured in several studies (24–26). So, the first set of spatial risk variables, termed “anthropogenic variables”, included the distribution of LPMs and human population density. The latter was included as it may be a good surrogate for some surveillance bias or for some anthropogenic transmission mechanisms.

 Second, we included two variables describing chicken and duck density. 69%-80% of H7N9 human cases of the five epidemic waves reported exposure to live poultry prior to infection, including LPM (52%–60%) and backyard poultry (13%-40%), and these figures remained fairly stable with time (1). Whilst the majority of those exposures may correspond to LPM visits, other opportunities for contact with poultry along the production and value chain could take place. For example, poultry workers in Beijing were shown to be at a higher risk of H7N9 infection than the remaining population of the city (27). Poultry may become a reservoir when the circulation of avian influenza viruses through the production and value-chain cannot be prevented, and poultry-related variables were found to be significant predictors of H7N9 risk in several previously published studies (8,23,26,31). During the 5th wave, outbreaks in farms started to be reported in higher numbers, thus we included a second set of predictor variables, termed “poultry” including the density of chickens and ducks, as these may regionally influence the risk of H7N9 virus transmission to humans.

Third, we included two indicator variables of proximity and abundance of water. Although the most conservative hypothesis remains that human infections would be linked to the circulation of H7N9 in domestic chicken reservoirs with occasional human exposure in LPMs, one can not exclude that wild birds may have taken part in the transmission. The virus precursors of the H7N9 virus in China were found in a wide variety of bird species, wild and domestic (10), and avian influenza viruses circulating in wild bird represent a gene pool that may recombine with H7N9 viruses and allow better adaptation and persistence. There is little information on the wild host specificity of H7N9, and data on the distribution of wild bird species is generally fairly coarse, with populations varying strongly according to the season and proxy variables indicative of proximity or abundance of water and wetlands are thus required.

 The spatial variation of H7N9 incidence in the human population during the 5 epidemic waves was studied in relation to these three sets of spatial factors. More specifically, we aimed to compare the association with these spatial factors in the five epidemic waves, to investigate the spatial distribution of repeated re-occurrences, and the year-to-year variation in predictability of H7N9 infections.

**Material and methods**

Data

*H7N9 human cases and seasonal Influenza*

 All confirmed H7N9 human cases during the period from March 2013 to 18th April 2017 were analysed. Information on laboratory-confirmed H7N9 human cases was collated by collecting data from the WHO Monthly Risk Assessment Summary report, websites of national and provincial Health and Family Planning Commission of China, FluTrackers, HealthMap, and avian influenza reports from Centre of Health Protection of HK. The WHO report prevailed in case there was inconsistent information. A detailed description of case definitions, surveillance for identification of cases, and laboratory testing for A H7N9 virus have been provided elsewhere (4,37,38). For each case, the information about place of residence and date of onset of symptoms were used and 6.5 days were subtracted from the date of onset of symptoms to estimate the dates of first contact with the virus, as estimated elswhere (39). To compare the seasonality of H7N9 human cases with that of human seasonal influenza A in space and time, we extracted influenza sentinel surveillance data for the period January 2013 – March 2017 from Influenza Weekly Reports managed by Chinese National Influenza Centre (http://www.chinaivdc.cn/cnic/zyzx/lgzb/). More information on the sentinel network supporting these data can be found in Yu et al. (34).

*Live poultry markets and permanent closure measures*

 A database recording the location of LPMs was assembled from multiple sources by the authors and described in Text 1 – Supplementary information and included the location of 8943 retail and wholesale markets. In addition, a database recording the market closure measures implemented since the first wave was compiled by the authors, with starting and end date of each measure.

*Spatial predictors*

 The first set of predictor variables included the LPM density (LPM/km2) and human population density (people / km2). Some counties do not have LPMs but the people living there may easily go to surrounding counties. LPMs may also act at a higher level by providing a network of markets through which the disease could spread and persist. So, an LPM spatial density was computed by means of a Gaussian smoothing kernel function with the optimal bandwidth found by Gilbert et al. (23). The data on permanent market closures was used to update the yearly distribution of LPM density, by annually removing the permanently closed markets out of the full LPM database prior to the Gaussian smoothing, hence resulting in a different LPM density distribution per epidemic wave. For human population, we used the human population density from the 2010 census (China Data Center).

 The second set of predictor variables included chicken and domestic duck density from a new data set produced using the Gridded Livestock of the World methodology applied to an extensively improved data set compiled by the authors with the 2010 reference year (40,41). A high correlation was noted between duck and chicken density at the county level and in order to reduce colinearity and make the results more easily interpretable, we derived two more independent predictor variables, namely the poultry density (chicken + duck heads / km2) and the chicken to duck ratio (chicken heads / duck heads).

 The last set of predictors indicative of water bird habitat included two variables. The distance to the largest lakes and reservoirs (km), measuring the distance between the county centroids and the nearest lakes (area ≥ 50 km²) or reservoirs (storage capacity ≥ 0.5 km³) (Lehner & Döll 2007), and the proportion (%) of county covered by wetlands, derived from the hybrid wetland map for China (42).

*Analyses*

 The analyses involved the development of Poisson Boosted regression tree (BRT) models to predict the daily incidence rate (DIR) of H7N9 virus in human population as a function of the set of six predictor variables. A brief description of BRT models and a list of model parameters is provided in Text 2 – Supplementary information. The models were developed using the number of human cases as dependent variable with an offset term corresponding to the product of human population by the duration of the epidemic. The duration of each epidemic was set as the period separating the 5th from the 95th percentiles of the days of onset of illness in each wave. One model per epidemic wave was built in to compare the effect of predictor variables, and to assess the predictive capacity from one wave to another The contribution of each predictor variable to the model was quantified by its relative contribution (RC), a measure of its overall importance in the model (45) and of its partial dependence plots, or BRT profiles, which provide a graphical description of its effect on the DIR after accounting for the average effects of all other predictor variables in the model (46). The presence of spatial autocorrelation in the model residuals was tested using spline correlograms (47) and the approach of Crase, Liedloff, and Wintle (48) was used when autocorrelation was present in the model residuals. We also converted the predicted incidence rate into a probability of having at least one human case in the county using a Binomial distribution as following:

P(X > 0) = 1 - (1 - *p*)*nd*

where *nd* is the population times the number of days in the epidemic duration; and *p* is the incidence rate predicted by the Poisson BRT model. This allowed testing the capacity of the models to discriminate between the presence and the absence of human cases at the county level. Finally, the analysis was replicated with Generalized Linear Models (GLMs) as BRT models do not explicitly allow formally testing the significance of individual factors.

**Results**

 Table 1 presents the relative contribution (RC) of the predictor variable in the different epidemic waves. The RC of anthropogenic predictor variables were generally high (w1 = 40.61%; w2 = 50.12%; w3 = 39.26%; w4 = 17.61%; w5 = 17.94%) but decreased strongly after the third epidemic wave. In parallel, the RC of poultry predictors increased and was greatest in the last epidemic wave (w1 = 10.47%; w2 = 5.83%; w3 = 2.64%; w4 = 28.54%; w5 = 41.83%). In this last epidemic wave, the most important predictor variables were by decreasing order of RC the Chicken to Duck ratio (27.28%), the LPM density (16.04%), the poultry density (14.55%) and the distance to open lakes and reservoirs (6.16%). Fig. 1 presents the BRT profiles of these four predictor variables in the different epidemic waves (the other profiles are provided as Fig. 1 – Supplementary information). The chicken to duck ratio had a significant RC only in waves 4 and 5, when it showed a positive association with incidence up to a ratio of approximately 30. The LPM density profile of wave 5 also showed a positive association with the LPM density, levelling-off at a density of 0.01, and with a relatively similar profile to the other epidemic waves. The 5th wave tended to associate lower incidence with the highest densities (> 0.03), in contrast to previous epidemic waves. The poultry density profile changed gradually over time, with an increasing RC, and the incidence rate in wave 5 is predicted to increase strongly in counties with a very high density of poultry (> 60,000 heads/km²). Finally, the profile of the distance to lakes showed a decreasing association, which in the range 0 – 100 km.

 The assessment of the BRT models goodness of fit is presented in Table 2, and with the exception of the 4th epidemic waves, the predictability of the models were moderate with cross-validation correlation coefficients within a range from 0.42 to 0.55. In presence/absence term, the models had a good discriminatory capacity with AUC ranging from 0.78 to 0.92 but this decreased over the years (w1 = 0.92; w2 = 0.85; w3 = 0.83; w4 = 0.86; w5 = 0.78). This highlights that it was apparently easier to predict the presence or absence of a human case (good discrimination capacity and AUC values) than it was to predict their number (moderate predictability and correlation coefficients). The results obtained with the same risk factors and dependent variable from the GLM models are provided in Text 3 – Supplementary information, and highlight a similar pattern, i.e. a shift between live poultry market and poultry variables noticeable after the 4th epidemic wave (Supplementary information).

 Figure 2 shows the distribution of the top-three predictor variables (Live-poultry market density, poultry density and chicken to duck ratio) in relation to the distribution of the past and last epidemic wave. The RGB composite plot (Fig. 2a) highlights areas where all three predictor variables were high and where H7N9 persisted over time (Fig. 2b). The figure highlights a large areas to the east of the Taihu Lake on the urban areas of Wuxi, Suzhou and Shanghai where high live-poultry market density is found and that includes several small hotspots of high poultry production. The RGB composite plot highlights three additional urban areas with high LPM density and high poultry density: the Guangdong province, the Tianjin and the Beijing urban areas and the Chongqing urban area. These different areas highlighted in the RGB maps visually correspond to areas of high H7N9 re-occurrence displayed in Fig. 2b, which contrasts counties with repeated reoccurrences from those with sporadic infections. These areas include southern Jiangsu, Shanghai and northern Zhejiang provinces, as well as Guangdong counties located around Hong Kong but to a lesser degree than the areas in and around Shanghai. Fig. 2c highlights that the spatial pattern of wave 5 showed a marked geographic expansion from these previous hotspots of persistence, with a 90 counties reporting H7N9 for the first time (50.85% of the total number of counties infected in wave 5). One can also visually measure why live-poultry density was a lower predictor in wave 5 than in previous waves, as these newly infected counties no longer match green areas depicted in Fig. 2a. The heat maps presented in Fig. 3 show that until now, the majority of H7N9 human cases have taken place in February to March (Fig. 3B) with a latitudinal gradient. The seasonality of common influenza A infection shows different levels of seasonality in China (Fig. 3C), with the province north of 34.1 degrees showing a much stronger annual winter seasonality of infection than more southern provinces, with a peak in December – February. The comparison of figure 3B to 3C highlights that the peaks of H7N9 and seasonal influenza have so far not been strongly coinciding in space and time. However, a geographic range expansion of H7N9 infections in the northern provinces, keeping its current seasonality, would bring the H7N9 and seasonal influenza incidence peaks to coincide much more extensively.

**Discussion**

 The results of our spatial models demonstrate a significant shift over time from anthropogenic to poultry predictor variables linked to H7N9 human cases, apparent in waves 4 and confirmed in wave 5. This shift was present in both BRT and GLM models. More specifically, the predictive power of poultry variables increased over time and was greatest in the last epidemic, pointing to areas with very high chicken densities and high chicken to duck ratios. A recent study on H7N9 human cases showed an increase in semi-urban and rural cases in the last wave, and a comparatively higher number of middle-aged cases (1). However, apart from the overall increase in cases, the study did not suggest any other major epidemiological differences, and other authors made similar observations when comparing waves 1-4 (1,5,11). Our results do not contradict the observation of a higher number of human cases in peri-urban and rural areas, because high poultry production regions are typically located in peri-urban and rural settings. But they strongly support the hypothesis that the H7N9 virus may have spread in the chicken reservoir much more extensively in the last two epidemic waves than previously, with a particularly marked geographical range expansion in the last epidemic wave. This observation based on human case can be linked to the emergence of HPAI H7N9 that was reported early 2017 in southern China (16). Recently published results showed that human cases of HPAI H7N9 were already found beyond Guangdong, in Hunan and Guangxi in early 2017 (18). In parallel, there was a comparatively higher number of reports of H7N9 positives found in poultry farms this year in comparison to previous epidemic waves, including reports of HPAI H7N9 in northern China, in Tianjin (32). The precise role of the gain in pathogenicity on the range expansion of H7N9 is yet unclear, as of the main mechanisms of transmission along the poultry production and value chain networks. However, the fact that such a range expansion took place in parallel to the emergence of a highly pathogenic variant can hardly be coincidental.

 It should be borne in mind that the measure of predictor weights in the model is relative, i.e. the sum of relative contribution equals to 1, so if poultry variables become better predictors of H7N9 incidence in human, the RC of other variable would decrease, even if their effect on the predicted incidence remained fairly constant, as could be the case for the LPM variable. The contribution of LPMs may have remained high, but its combination with increasing transmission along the poultry production and value chains may be responsible for the geographical range expansion and higher incidence of the 5th wave.

 Although some of the highest incidences were observed along Taihu Lake, the predictive capacity of water bird-related predictor variables appeared to have a much lower influence on the predicted incidence than anthropogenic and poultry variables. Many interfaces combining wetlands, intensive poultry farming and rice paddy fields are present in south-eastern China and may have played a role in the initial emergence of the H7N9 virus in the Shanghai area (33). But as the virus spread in the following epidemic wave, the contribution of wild birds to overall disease circulation may be fairly low, which is reflected by the low relative contribution of the water bird habitat proxy variables.

 The predictive capacity of the incidence models was only moderate, as these spatial models did not account for the variability in incidence linked to market closure measures. In contrast, the predictions of presence/absence were generally better, because presence cannot be influenced by market closure measures, as such measures usually followed human cases rather than preceding them, and few counties implemented market closure measures in the absence of human cases.

The geographical range expansion and increase in incidence of human cases in the 5th wave of H7N9 brings serious human health concerns. First, repeated human infection by avian influenza viruses increases the chances of virus recombination and/or mutation possibly leading to human to human transmission. Second, the provinces affected by earlier H7N9 epidemic waves do not have a strong seasonal influenza A peak in January and February (34) that matches the peak of H7N9 cases (Fig. 3). However, if the H7N9 virus continues to expand its range northward, in areas with a strong influenza A peak in January and February, there will be a higher chance of local coincidence of peaks of incidence between human cases of H7N9 and seasonal influenza A virus. This may enhance the chances of co-infections that could lead to the emergence of reassortants with the capacity to easily transmit between humans. Third, the extent of the geographical range of the expansion is not yet fully known and in the absence of new measures, it may spread further within China, and internationally through poultry value-chains.

 An important limitation of the study is that there may be a significant underreporting of milder symptomatic infections (Yu et al. 2013) and the effect and geaographical distribution of this bias is unknown. Another aspect is that the poultry dataset underlying our analyses is of uneven quality, with better and more detailed data along the east than in western parts of the country, as shown in Artois et al. (2016). Finally, although all efforts were made to compile the most comprehensive LPM dataset as possible, many LPMs may have opened and closed, including illegal ones, which may add further uncertainty. Finally, we were not able to integrate poultry movement and trade data into this analysis (legal or illegal) due to the lack of centralized data, and this may be an important path of future progresses.

In conclusions, an important shift was observed between the predictor variables of H7N9 human cases during the course of the five epidemic waves, with poultry predictor variables becoming more important than anthropogenic predictor variables, particularly in waves 4 and 5. These results support the hypothesis of a recent and significant geographical expansion of H7N9 viruses in the poultry reservoir, possibly linked to the higher pathogenicity and increased frequency of reports in poultry farms. The current range expansion of H7N9 to more northerly latitudes may increase the chances of H7N9 peaks coinciding in both space and time with those of seasonal influenza A infection, leading to higher risk of co-infections and reassortments than before, although the risk is still low so far.

**Acknowledgements**

We thank Shuanbao Yu from Chinese Center for Disease Control and Prevention and Wanqi Yang from Wuhan University for assistance in data collection. The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the Food and Agriculture Organization of the United Nations.

**Funding**

This study was funded by grants from the National Science Fund for Distinguished Young Scholars (grant no. 81525023), the US National Institutes of Health (Comprehensive International Program for Research on AIDS grant U19 AI51915 and grant number 1R01AI101028-02A1), China CDC’s Key Laboratory of Surveillance and Early-warning on Infectious Disease, and the National Natural Science Foundation of China (grant no. 81602936). The authors would like to thank the USAID Emerging Pandemic Threat programme (EPT) for their continued support.

**References**

1. Wang X, Jiang H, Wu P, Uyeki TM, Feng L, Lai S, 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 [Internet]. 2017 Jun 2 [cited 2017 Jun 4];0(0). Available from: http://www.thelancet.com/journals/laninf/article/PIIS1473-3099(17)30323-7/abstract

2. Zhou L, Ren R, Yang L, Bao C, Wu J, Wang D, et al. Sudden increase in human infection with avian influenza A (H7N9) virus in China, September-December 2016. West Pac Surveill Response [Internet]. 2017 [cited 2017 Apr 3];8(1). Available from: http://ojs.wpro.who.int/ojs/index.php/wpsar/article/view/521

3. Uyeki TM, Katz JM, Jernigan DB. Novel influenza A viruses and pandemic threats. The Lancet. 2017 Jun 3;389(10085):2172–4.

4. Yu H, Cowling BJ, Feng L, Lau EH, Liao Q, Tsang TK, et al. Human infection with avian influenza A H7N9 virus: an assessment of clinical severity. The Lancet. 2013 Jul 19;382(9887):138–45.

5. Xiang N. Assessing Change in Avian Influenza A (H7N9) Virus Infections During the Fourth Epidemic—China, September 2015–August 2016. MMWR Morb Mortal Wkly Rep [Internet]. 2016 [cited 2017 Apr 3];65. Available from: https://espanol.cdc.gov/enes/mmwr/volumes/65/wr/mm6549a2.htm?s\_cid=mm6549a2\_w

6. Wu Z-Q, Zhang Y, Zhao N, Yu Z, Pan H, Chan T-C, et al. Comparative Epidemiology of Human Fatal Infections with Novel, High (H5N6 and H5N1) and Low (H7N9 and H9N2) Pathogenicity Avian Influenza A Viruses. Int J Environ Res Public Health. 2017 Mar 4;14(3):263.

7. Lai S, Qin Y, Cowling BJ, Ren X, Wardrop NA, Gilbert M, et al. Global epidemiology of avian influenza A H5N1 virus infection in humans, 1997–2015: a systematic review of individual case data. Lancet Infect Dis [Internet]. 2016 May [cited 2016 Jun 7]; Available from: http://linkinghub.elsevier.com/retrieve/pii/S1473309916001535

8. Artois J, Lai S, Feng L, Jiang H, Zhou H, Li X, et al. H7N9 and H5N1 avian influenza suitability models for China: accounting for new poultry and live-poultry markets distribution data. Stoch Environ Res Risk Assess. 2016 Dec 5;1–10.

9. Bui CM, Gardner L, MacIntyre R, Sarkar S. Influenza A H5N1 and H7N9 in China: A spatial risk analysis. PLOS ONE. 2017 Apr 4;12(4):e0174980.

10. Lam TT-Y, Wang J, Shen Y, Zhou B, Duan L, Cheung C-L, et al. The genesis and source of the H7N9 influenza viruses causing human infections in China. Nature. 2013 Oct 10;502(7470):241–4.

11. Xiang N, Iuliano AD, Zhang Y, Ren R, Geng X, Ye B, et al. Comparison of the first three waves of avian influenza A(H7N9) virus circulation in the mainland of the People’s Republic of China. BMC Infect Dis [Internet]. 2016 Dec [cited 2017 Apr 3];16(1). Available from: http://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-016-2049-2

12. Zhu H, Lam TT-Y, Smith DK, Guan Y. Emergence and development of H7N9 influenza viruses in China. Curr Opin Virol. 2016 Feb;16:106–13.

13. Lam TT-Y, Zhou B, Wang J, Chai Y, Shen Y, Chen X, et al. Dissemination, divergence and establishment of H7N9 influenza viruses in China. Nature. 2015 Mar 11;522(7554):102–5.

14. Pantin-Jackwood MJ, Miller PJ, Spackman E, Swayne DE, Susta L, Costa-Hurtado M, et al. Role of Poultry in the Spread of Novel H7N9 Influenza Virus in China. J Virol. 2014 May 15;88(10):5381–90.

15. Kalthoff D, Bogs J, Grund C, Tauscher K, Teifke JP, Starick E, et al. Avian influenza H7N9/13 and H7N7/13: A comparative virulence study in chickens, pigeons, and ferrets. J Virol. 2014 Jun 4;JVI.01241–14.

16. Zhu W, Zhou J, Li Z, Yang L, Li X, Huang W, et al. Biological characterisation of the emerged highly pathogenic avian influenza (HPAI) A(H7N9) viruses in humans, in mainland China, 2016 to 2017. Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull. 2017 May 11;22(19).

17. Ke C, Mok CKP, Zhu W, Zhou H, He J, Guan W, et al. Human Infection with Highly Pathogenic Avian Influenza A(H7N9) Virus, China. Emerg Infect Dis [Internet]. 2017 Jul [cited 2017 Jun 4];23(8). Available from: http://wwwnc.cdc.gov/eid/article/23/8/17-0600\_article.htm

18. Zhou L, Tan Y, Kang M, Liu F, Ren R, Wang Y, et al. Preliminary Epidemiology of Human Infections with Highly Pathogenic Avian Influenza A(H7N9) Virus, China, 2017. Emerg Infect Dis [Internet]. 2017 Aug [cited 2017 Jun 4];23(8). Available from: http://wwwnc.cdc.gov/eid/article/23/8/17-0640\_article.htm

19. Yu H, Wu JT, Cowling BJ, Liao Q, Fang VJ, Zhou S, et al. Effect of closure of live poultry markets on poultry-to-person transmission of avian influenza A H7N9 virus: an ecological study. The Lancet. 2014 Feb;383(9916):541–8.

20. Yuan J, Lau EHY, Li K, Leung YHC, Yang Z, Xie C, et al. Effect of Live Poultry Market Closure on Avian Influenza A(H7N9) Virus Activity in Guangzhou, China, 2014. Emerg Infect Dis. 2015 Oct;21(10):1784–93.

21. Wu J, Lu J, Faria NR, Zeng X, Song Y, Zou L, et al. Effect of Live Poultry Market Interventions on Influenza A(H7N9) Virus, Guangdong, China. Emerg Infect Dis. 2016 Dec;22(12):2104–12.

22. Zhou X, Li Y, Wang Y, Edwards J, Guo F, Clements ACA, et al. The role of live poultry movement and live bird market biosecurity in the epidemiology of influenza A (H7N9): A cross-sectional observational study in four eastern China provinces. J Infect. 2015 Oct;71(4):470–9.

23. Gilbert M, Golding N, Zhou H, Wint GRW, Robinson TP, Tatem AJ, et al. Predicting the risk of avian influenza A H7N9 infection in live-poultry markets across Asia. Nat Commun. 2014 Jun 17;5:4116.

24. Fang L-Q, Li X-L, Liu K, Li Y-J, Yao H-W, Liang S, et al. Mapping Spread and Risk of Avian Influenza A (H7N9) in China. Sci Rep [Internet]. 2013 Sep 26 [cited 2013 Dec 19];3. Available from: http://www.nature.com/srep/2013/130926/srep02722/full/srep02722.html

25. Fuller T, Havers F, Xu C, Fang L-Q, Cao W-C, Shu Y, et al. Identifying areas with a high risk of human infection with the avian influenza A (H7N9) virus in East Asia. J Infect. 2014 Aug;69(2):174–81.

26. Li X-L, Yang Y, Sun Y, Chen W-J, Sun R-X, Liu K, et al. Risk Distribution of Human Infections with Avian Influenza H7N9 and H5N1 virus in China. Sci Rep. 2015 Dec 22;5:18610.

27. Yang P, Ma C, Cui S, Zhang D, Shi W, Pan Y, et al. Avian influenza A(H7N9) and (H5N1) infections among poultry and swine workers and the general population in Beijing, China, 2013–2015. Sci Rep [Internet]. 2016 Dec [cited 2017 Apr 3];6(1). Available from: http://www.nature.com/articles/srep33877

28. Kapan DD, Bennett SN, Ellis BN, Fox J, Lewis ND, Spencer JH, et al. Avian Influenza (H5N1) and the Evolutionary and Social Ecology of Infectious Disease Emergence. EcoHealth. 2006 Sep 22;3(3):187–94.

29. Peiris JSM, Cowling BJ, Wu JT, Feng L, Guan Y, Yu H, et al. Interventions to reduce zoonotic and pandemic risks from avian influenza in Asia. Lancet Infect Dis. 2016 Feb;16(2):252–8.

30. Bahl J, Pham TT, Hill NJ, Hussein ITM, Ma EJ, Easterday BC, et al. Ecosystem Interactions Underlie the Spread of Avian Influenza A Viruses with Pandemic Potential. PLOS Pathog. 2016 May 11;12(5):e1005620.

31. Xu M, Cao C, Li Q, Jia P, Zhao J. Ecological Niche Modeling of Risk Factors for H7N9 Human Infection in China. Int J Environ Res Public Health. 2016 Jun 16;13(6).

32. FAO Empres. FAO H7N9 situation update - Avian Influenza A(H7N9) virus - FAO Emergency Prevention System for Animal Health (EMPRES-AH) [Internet]. 2017 [cited 2017 Jun 4]. Available from: http://www.fao.org/ag/againfo/programmes/en/empres/h7n9/situation\_update.html

33. Liu D, Shi W, Shi Y, Wang D, Xiao H, Li W, et al. Origin and diversity of novel avian influenza A H7N9 viruses causing human infection: phylogenetic, structural, and coalescent analyses. The Lancet. 2013 Jun;381(9881):1926–32.

34. Yu H, Alonso WJ, Feng L, Tan Y, Shu Y, Yang W, et al. Characterization of Regional Influenza Seasonality Patterns in China and Implications for Vaccination Strategies: Spatio-Temporal Modeling of Surveillance Data. PLOS Med. 2013 Nov 19;10(11):e1001552.

35. Martin V, Zhou X, Marshall E, Jia B, Fusheng G, FrancoDixon MA, et al. Risk-based surveillance for avian influenza control along poultry market chains in South China: The value of social network analysis. Prev Vet Med. 2011 Dec 1;102(3):196–205.

36. Soares Magalhães RJ, Zhou X, Jia B, Guo F, Pfeiffer DU, Martin V. Live poultry trade in Southern China provinces and HPAIV H5N1 infection in humans and poultry: the role of Chinese New Year festivities. PloS One. 2012;7(11):e49712.

37. Cowling BJ, Jin L, Lau EH, Liao Q, Wu P, Jiang H, 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. The Lancet. 2013;382(9887):129–37.

38. Qin Y, Horby PW, Tsang TK, Chen E, Gao L, Ou J, et al. Differences in the Epidemiology of Human Cases of Avian Influenza A(H7N9) and A(H5N1) Viruses Infection. Clin Infect Dis Off Publ Infect Dis Soc Am. 2015 Aug 15;61(4):563–71.

39. Virlogeux V, Li M, Tsang TK, Feng L, Fang VJ, Jiang H, et al. Estimating the Distribution of the Incubation Periods of Human Avian Influenza A(H7N9) Virus Infections. Am J Epidemiol. 2015 Oct 15;182(8):723–9.

40. Robinson TP, Wint GRW, Conchedda G, Van Boeckel TP, Ercoli V, Palamara E, et al. Mapping the Global Distribution of Livestock. PLoS ONE. 2014 May 29;9(5):e96084.

41. Nicolas G, Robinson TP, Wint GRW, Conchedda G, Cinardi G, Gilbert M. Using Random Forest to Improve the Downscaling of Global Livestock Census Data. PLOS ONE. 2016 Mar 15;11(3):e0150424.

42. Ma K, You L, Liu J, Zhang M. A Hybrid Wetland Map for China: A Synergistic Approach Using Census and Spatially Explicit Datasets. PLOS ONE. 2012 Oct 23;7(10):e47814.

43. Dhingra MS, Artois J, Robinson TP, Linard C, Chaiban C, Xenarios I, et al. Global mapping of highly pathogenic avian influenza H5N1 and H5Nx clade 2.3.4.4 viruses with spatial cross-validation. eLife. 2016 Nov 25;5:e19571.

44. Bretz F, Hothorn T, Westfall PH, others. Multiple comparisons using R. CRC Press Boca Raton; 2011.

45. Friedman JH, Meulman JJ. Multiple additive regression trees with application in epidemiology. Stat Med. 2003 May 15;22(9):1365–81.

46. Elith J, Leathwick JR, Hastie T. A working guide to boosted regression trees. J Anim Ecol. 2008 Jul;77(4):802–13.

47. BjØrnstad ON, Falck W. Nonparametric spatial covariance functions: Estimation and testing. Environ Ecol Stat. 2001 Mar 1;8(1):53–70.

48. Crase B, Liedloff AC, Wintle BA. A new method for dealing with residual spatial autocorrelation in species distribution models. Ecography. 2012 Oct 1;35(10):879–88.

Table 1 Relative contribution of the different BRT models across the different epidemic waves

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|   |   | Wave 1 | Wave 2 | Wave 3 | Wave 4 | Wave 5 |
| Anthropogenic (sum of RC) | 40.61 | 50.12 | 39.26 | 17.61 | 17.94 |
|  | LPM density  | 38.68 ± 0.76 | 49.48 ± 0.4 | 11.02 ± 0.89 | 13.59 ± 0.77 | 16.04 ± 0.16 |
|  | Human pop. density  | 1.93 ± 0.17 | 0.64 ± 0.03 | 28.24 ± 0.55 | 4.02 ± 0.24 | 1.9 ± 0.06 |
|  |  |  |  |  |  |  |
| Poultry (sum of RC) | 10.47 | 5.83 | 2.64 | 28.54 | 41.83 |
|  | Chicken to duck ratio  | 5.29 ± 0.15 | 4.34 ± 0.05 | 0.72 ± 0.11 | 19.96 ± 0.42 | 27.28 ± 0.34 |
|  | Poultry density  | 5.18 ± 0.16 | 1.49 ± 0.04 | 1.92 ± 0.17 | 8.58 ± 0.32 | 14.55 ± 0.11 |
|  |  |  |  |  |  |  |
| Water habitat (sum of RC) | 2.21 | 3.94 | 9.75 | 5.95 | 7.67 |
|  | Prop. of wetlands | 0.55 ± 0.05 | 1.34 ± 0.06 | 1.81 ± 0.12 | 0.8 ± 0.05 | 1.51 ± 0.05 |
|  | Distance to lakes | 1.66 ± 0.07 | 2.6 ± 0.09 | 7.94 ± 0.31 | 5.15 ± 0.19 | 6.16 ± 0.07 |
|  |  |  |  |  |  |  |
| Autoregressive term | 46.7 ± 0.6 | 40.11 ± 0.34 | 48.35 ± 0.77 | 47.9 ± 1.22 | 32.55 ± 0.26 |
|   |   |   |   |   |   |   |

Table 2 Goodness of fit metrics of the BRT models across the different epidemic waves

|  |  |  |
| --- | --- | --- |
|  | Pearson Corr. Coefficient | AUC |
|   | Training | Training\_Auto | CV | Training | Training\_Auto |
| wave 1 | 0.785 ± 0.012 | 0.552 ± 0.002 | 0.477 ± 0.015 | 0.923 ± 0.001 | 0.907 ± 0.002 |
| wave 2 | 0.753 ± 0.004 | 0.326 ± 0.007 | 0.548 ± 0.013 | 0.85 ± 0.001 | 0.849 ± 0 |
| wave 3 | 0.582 ± 0.011 | 0.495 ± 0.004 | 0.42 ± 0.013 | 0.831 ± 0.003 | 0.814 ± 0.001 |
| wave 4 | 0.44 ± 0.012 | 0.293 ± 0.006 | 0.259 ± 0.008 | 0.856 ± 0.001 | 0.832 ± 0.001 |
| wave 5 | 0.61 ± 0.002 | 0.527 ± 0.002 | 0.447 ± 0.009 | 0.778 ± 0.001 | 0.755 ± 0 |

**Figure Captions**

Figure 1 Marginal effect plots of the top-4 predictor variables on the predicted incidence rate, with the change in relative contribution over time indicated by the bars on the top of each plot, showing the increasing relative contribution of the poultry predictor variables. The smoothed line on the top left part of each plot is indicative of the distribution of each variable.

Figure 2 Distribution of predictor variables and H7N9 infections. A. Red-Green-Blue visualisation of poultry density (red), live-poultry market density (green) and chicken / duck ratio (blue), with dark areas corresponding to low values in all three predictors, and white areas to high values in all three predictors. B. Number of years with at least 1 human case per county. C Distribution of the 5th wave of human infections over previous ones.

Figure 3 Seasonality of H7N7 infections in comparison to seasonal influenza. A. epidemic curve for H7N9; B seasonality for H7N9; C seasonality for seasonal influenza.