

Original Contribution

Persistence of Highly Pathogenic Avian Influenza H5N1 Virus Defined by Agro-Ecological Niche

Lenny Hogerwerf,^{1,2} Rob G. Wallace,³ Daniela Ottaviani,⁴ Jan Slingenbergh,⁴ Diann Prosser,⁵ Luc Bergmann,⁶ and Marius Gilbert^{1,7}

¹Biological Control and Spatial Ecology, Université Libre de Bruxelles CP160/12, Av FD Roosevelt 50, 1050 Brussels, Belgium

²Division of Epidemiology, Department of Farm Animal Health, Utrecht University, Utrecht, The Netherlands

³Institute for Global Studies, University of Minnesota, Minneapolis, MN

⁴Food and Agriculture Organization of the United Nations (FAO), Viale delle Terme di Caracalla, 00100 Rome, Italy

⁵USGS Patuxent Wildlife Research Center, 10300 Baltimore Avenue, Beltsville, MD 20705

⁶Department of Geography, University of Minnesota, Minneapolis, MN

⁷Fonds National de la Recherche Scientifique, rue d'Egmont 5, 1000 Brussels, Belgium

Abstract: The highly pathogenic avian influenza (HPAI) H5N1 virus has spread across Eurasia and into Africa. Its persistence in a number of countries continues to disrupt poultry production, impairs smallholder livelihoods, and raises the risk a genotype adapted to human-to-human transmission may emerge. While previous studies identified domestic duck reservoirs as a primary risk factor associated with HPAI H5N1 persistence in poultry in Southeast Asia, little is known of such factors in countries with different agro-ecological conditions, and no study has investigated the impact of such conditions on HPAI H5N1 epidemiology at the global scale. This study explores the patterns of HPAI H5N1 persistence worldwide, and for China, Indonesia, and India includes individual provinces that have reported HPAI H5N1 presence during the 2004–2008 period. Multivariate analysis of a set of 14 agricultural, environmental, climatic, and socio-economic factors demonstrates in quantitative terms that a combination of six variables discriminates the areas with human cases and persistence: agricultural population density, duck density, duck by chicken density, chicken density, the product of agricultural population density and chicken output/input ratio, and purchasing power per capita. The analysis identifies five agro-ecological clusters, or niches, representing varying degrees of disease persistence. The agro-ecological distances of all study areas to the medoid of the niche with the greatest number of human cases are used to map HPAI H5N1 risk globally. The results indicate that few countries remain where HPAI H5N1 would likely persist should it be introduced.

Keywords: global ecology, highly pathogenic avian influenza, agro-ecology, epidemiology

Electronic supplementary material: The online version of this article (doi: [10.1007/s10393-010-0324-z](https://doi.org/10.1007/s10393-010-0324-z)) contains supplementary material, which is available to authorized users.

Published online: June 29, 2010

Correspondence to: Marius Gilbert, e-mail: mgilbert@ulb.ac.be

INTRODUCTION AND PURPOSE

Each new introduction of the highly pathogenic avian influenza H5N1 virus (HPAI H5N1 hereafter) is a novel natural experiment providing clues about conditions that

promote the virus's establishment and persistence. Despite the large literature on the HPAI H5N1 panzootic, surprisingly few studies have so far investigated the spatial factors associated with disease presence. Much of this work has focused on a limited number of countries: Thailand, Vietnam, China, Romania, and Nigeria (Gilbert et al., 2006, 2007; Oyana et al., 2006; Pfeiffer et al., 2007; Cecchi et al., 2008; Ward et al., 2008). These studies identified a limited number of key agro-ecological risk factors. Free-grazing domestic ducks (Gilbert et al., 2006, 2007; Pfeiffer et al., 2007) and floodplain agriculture or wetland-related risk factors (Oyana et al., 2006; Gilbert et al., 2007; Cecchi et al., 2008; Ward et al., 2008) appear to be fundamental drivers for H5N1 persistence, and have been used to produce regional risk maps (Gilbert et al., 2008). Regional ecological niche modeling using parameters of surface reflectance and landform has been done for West Africa (Williams et al., 2008), the Middle East and northeastern Africa (Williams and Peterson, 2009), and the Indian subcontinent (Adhikari et al., 2009). These three papers all found that HPAI H5N1 cases were occurring under predictable ecological conditions, and found that, in most cases, predicted disease presence coincided with areas that had marked seasonal variation. However, to date, there have been few attempts at analyzing global patterns of HPAI H5N1 distribution, with the exception of work by Kilpatrick et al. (2006), which characterized the global risk of virus introduction based on large-distance poultry trade and migratory birds' movement patterns.

Our aim here was to identify where the virus would stand the best chance of becoming established and promoting secondary spread, should it be introduced. The approach requires quantitatively testing for matches between HPAI H5N1 persistence and the agro-ecological conditions that the virus experienced during each new invasion. These conditions may be defined in terms of poultry population composition, production structure, veterinary resources, or other factors that influence persistence and viral evolution (Peng, 1994; Abo et al., 1998). Areas with large populations of free-grazing ducks, high densities of backyard poultry, and large numbers of poultry farms (e.g., Vietnam) are likely to present selection regimes that differ from those in countries with heavily industrialized poultry production (e.g., Japan), or from countries with negligible poultry densities (e.g., Sudan). Not only would such analysis test the importance of risk factors previously identified, but it may also reveal the importance of factors that are broadly homogeneous within countries but may vary across coun-

tries. For example, an indicator such as minimum temperature may not so much contrast different regions at risk within Laos, but it would do so at the global scale. In addition, identifying and characterizing conditions of persistence opens a new line of study in evolutionary epidemiology, namely tracking viral evolutionary responses under different agro-ecological conditions.

There are several obstacles to overcome for such a global approach. First, quantifying HPAI H5N1 persistence across countries is difficult given major discrepancies in surveillance and reporting. The absence of HPAI in countries that fail to report H5N1 outbreaks cannot be reasonably inferred for those countries lacking veterinary resources and surveillance capacity. Second, a global data set by definition compiles data from countries characterized by very different epidemiological conditions, because disease persistence entails different processes under different sets of circumstances. For example, China has for several years adopted mass vaccination in poultry countrywide, and it appears difficult to differentiate if reported HPAI H5N1 cases reflect incomplete vaccination coverage or a field virus evading the vaccines applied. These uncertainties may confound honest attempts at mapping risk. For example, Fang et al. (2008) predict a high HPAI risk for the Tibetan highlands and the half-desert ecology of Xinjiang province and only a moderate high-risk level for the Guangdong and Fujian provinces where HPAI genetic diversity appears quite large (Wallace et al., 2007).

This article attempts to overcome these obstacles to analyze the agro-ecological conditions associated with HPAI H5N1 globally. First, the indicators that presumably best relate to different measures of HPAI H5N1 persistence are inventoried and validated. Second, the countries with similar agro-ecological conditions are statistically clustered into "niches." Finally, the agro-ecological distances between countries worldwide and the niche with the greatest human H5N1 risk are mapped.

METHODS

Data

The analyses included all countries and Chinese, Indonesian, and Indian provinces where HPAI H5N1 has been registered between January 2004 and December 2008 ($n = 61$ for the all-countries set, $n = 124$ for the all-countries and Chinese, Indonesian, and Indian provinces set; full list in Supporting Information). China, Indonesia, and

India were included at the province-level because of the countries' sizes and importance to HPAI H5N1 epidemiology (Vijaykrishna et al., 2008; Wallace and Fitch, 2008; Chakrabarti et al., 2009). We therefore had two sets of spatial units for our analyses: one with only national level data ($n = 61$), and another one with 58 countries and 66 provinces from China, India, and Indonesia.

In order to account for heterogeneities in reporting and epidemiological situation, three different indices of HPAI H5N1 persistence were considered (Table 1). First, the number of seasons (winter, spring, summer, and autumn) with HPAI H5N1 reports either in poultry, wild birds, or mammals, including humans, expressed as a proportion of total season number since the recorded introduction of HPAI H5N1 (PrS, numeric, from 0 to 1). Second, the presence of human cases for more than 1 year over the study period (HcP, binary). Third, although it may represent a transient epidemic, i.e., not true persistence, the presence of human cases (Hc, binary) was also considered as a complementary index. This was based on the assumption that, in most instances, human cases were indicative of a relatively extensive virus circulation, i.e., agro-ecological conditions permitting an easy transmission.

Data on human cases of HPAI H5N1 were compiled from WHO (2009). The presence of HPAI H5N1 in poultry and wild birds between autumn 2003 and autumn 2008 was established by season from the OIE WAHID (World Organization for Animal Health, 2009) and FAO EMPRES-I databases (FAO, 2009a).

Fourteen explanatory variables were collected for all spatial units (Table 1). The first series of indicators described the poultry setting:

- i) *Chicken density (ChDn, heads/km²)*: Although this variable was not found significantly related to HPAI H5N1 persistence in previous work (Gilbert et al., 2008; Ward et al., 2008), chicken production and trade appear to drive the expansion phase of H5N1 invasion (Slingenbergh and Gilbert, 2008);
- ii) *Chicken production output/input (OICh, meat kg/head)*: Chicken density alone is a necessary but insufficient marker of the environment in which poultry-associated HPAI may spread. How that density is organized is, in itself, another characteristic. The more industrialized the production, the more kg of meat is produced per poultry head in stock on an annual basis;

Table 1. List of Abbreviations for Three Indices of HPAI H5N1 Persistence and 14 Explanatory Variables

Code	Definition	Unit
Hc	Presence of human cases	Binary (0 or 1)
HcP	Presence of human cases for more than 1 year over the study period	Binary (0 or 1)
PrS	No. of seasons (winter, spring, summer, and autumn) with HPAI H5N1 reports either in poultry, wild birds, or mammals, including humans, expressed as a proportion of total season no. since the recorded introduction of HPAI H5N1	Proportion (between 0 and 1)
AgPopDn	Agricultural population density	People/km ²
AgPopDnOI	Agricultural population density by output/input chicken	Sqrt(AgPopDn × OICh)
ChDn	Chicken density	Heads/km ²
ChMtDn	Chicken meat density	Kg/km ²
DuChDn	Duck density by chicken density	Sqrt(DuDn × ChDn)
DuDn	Duck density	Heads/km ²
OICh	Chicken production output/input	Meat kg/head
PopDn	Total human population density	People/km ²
Pppc	Purchasing power per capita	USD/person
Tmax	Maximum temperature	°C
Tmin	Minimum temperature	°C
Trng	Range of temperature	°C
UrPopDn	Urban population density	People/km ²
UrRt	Proportion of urban population	Proportion

- iii) *Duck density (DuDn, heads/km²)*: Duck density was previously shown to be a significant risk factor for HPAI H5N1 (Pfeiffer et al., 2007; Gilbert et al., 2008) and is a marker of potential disease reservoirs;
- iv) *Duck density by chicken density (DuChDn; estimated as $\sqrt{DuDn \times ChDn}$)*: The product identifies areas where synergistic interactions between chicken and duck populations may take place;
- v) *Chicken meat density (ChMtDn, kg/km²)*: The variable is an indicator of the density of broiler industry plants.

In addition to the above, a number of socio-demographic indicators were included:

- i) *Agricultural population density (AgPopDn, people/km²)*: Human population density was previously found to be a significant risk factor (Pfeiffer et al., 2007; Gilbert et al., 2008), with agricultural population density indicative of the density of smallholders;
- ii) *Urban population density (UrPopDn, people/km²) and proportion of urban population (UrRt)*: Expansions in poultry production in periurban areas worldwide have been reported to play a role in HPAI epidemiology (Kapan et al., 2006; Fasina et al., 2007). Urban centers with live bird markets attract viruses from different poultry settings, acting as both sink and source for viral infestation;
- iii) *Total human population density (PopDn, people/km²)*: Total density was included to make allowance for the possibility that total human density alone may inform us about the relative importance of the avian–human interface;
- iv) *Purchasing power per capita (Pppc, USD/person)*: Purchasing power is a general indicator of poverty/wealth, and thereby the degree to which poultry production practices and hygiene measures are up to standard;
- v) *Agricultural population density by output/input chicken (AgPopDnOI; estimated as $\sqrt{AgPopDn \times OICh}$)*: This product provides a measure for the extent to which poultry industrialization has replaced smallholder production. Coexistence of old and new forms of poultry production may enhance viral circulation by increasing the mix of host types and farming sectors across local landscapes (Wallace et al., in press). An increase in such diversity increases the kinds of ecological refugia in which livestock influenza can avoid extirpation.

Finally, we also included indicators of climate (Table 1):

- i) *Minimum, maximum, and range of temperature (Tmin, Tmax, Trng, °C)*: Temperature may be indicative of virus persistence in the environment (Brown et al., 2007; Li et al., 2004). It also marks latitude and the role migratory birds play in H5N1 dispersal. Temperature range is a proxy for climatic seasonality.

National level agricultural and human demography indicators were collected from FAOSTAT data from 2007 (FAO, 2009b), purchasing power per capita was collected from the International Monetary Fund (2009), and temperature data were collected from the Tyndall Center (Mitchell et al., 2004). Subnational data for China were collected from animal husbandry yearbook data for Chinese provinces (Editorial Department of China Animal Husbandry Yearbook, 2005), and subnational data for India and Indonesia were obtained from Glipha (FAO, 2009c), Statistics Indonesia of the Republic of Indonesia (Badan Pusat Statistik, 2009), and Office of the Registrar General and Census Commissioner, India (2009). Countries worldwide were included in the database so that the persistence models developed could be extrapolated to areas where HPAI H5N1 has not yet spread. Duck data were not available for many countries. Because of the importance of ducks in HPAI H5N1 epidemiology (Gilbert et al., 2006), we developed a model to predict national level duck numbers for countries for which this variable was missing (see Supplementary Information).

For mapping purposes, we masked areas where poultry production is considered to be impossible, i.e., desert regions, using a suitability mask derived from the gridded livestock of the world database (Wint and Robinson, 2007).

Analysis

Exploratory data analysis consisted of testing all explanatory indicators for normality, log-transforming those variables requiring so (ChDn, DuDn, DuChDn, ChMtDn, AgPopDn, UrPopDn, PopDn, and AgPopDnOI). We aimed to quantify the statistical associations between the 14 predictors and the three different measures of HPAI H5N1 persistence (Hc, HcP, and PrS). Linear models (standard multiple linear regression for PrS, and multiple logistic regressions for Hc and HcP) were chosen because: i) they allow testing the statistical relationship between a set of independent variables and a quantitative (PrS), or binary independent variable (Hc and HcP) using the same modeling framework; ii) they allow testing the significance of

the identified relationships; iii) their calculation is fast, which allows testing multiple alternative models easily; and iv) they belong to a family of parametric statistics most commonly used by other scientists, hence allowing easy replication of the results. However, one difficulty with this type of analysis is the multi-colinearity between predictors, as the level of significance of a given variable in a model may be affected by another correlated variable in such a way as to confuse epidemiological interpretation. In addition, given the limited number of spatial units in the final data set, it was difficult to simultaneously fit more than three variables in any model.

Hence, we first explored the model space by running all possible models one can build for combinations of three explanatory variables out of the set of 14. Hc and HcP were modeled using multiple logistic regressions, and we used two indicators to assess these models' goodness of fit. The Akaike's Information Criteria (AIC) is a measure of the goodness of fit of an estimated statistical model, and is estimated as $AIC = 2k - 2\ln(L)$, where k is the number of parameters in the model, and L is the maximized value of the likelihood function for the estimated model. In addition, we also computed the area under curve (AUC) or the receiver operating characteristic curve (ROC). The ROC curve is a plot that present the fraction of true positives as a function of the fraction of false positives for changing values of a discrimination criterion. The AUC quantifies the area under the ROC curve, and represents the overall ability of the predictions to discriminate between positives and negatives. Useless predictions (no better at identifying positives than flipping a coin) have an AUC of 0.5, while perfect predictions (zero false-positives and zero false-negatives) have an AUC of 1.00. The AUC is increasingly used to quantify the goodness of fit of models with binary outcomes because it is relatively independent to the rate of positives (Fielding and Bell, 1997). Concurrently, PrS was modeled using standard multiple regression models, with goodness of fit estimated by AIC and R^2 . In total, we estimated the AIC, AUC, and R^2 of 364 different models from a round robin of three explanatory variables. This analysis was repeated using two sets of data: all-countries ($n = 61$), with China, Indonesia, and India each as a single unit, and all-countries plus the individual provinces of China, Indonesia, and India ($n = 124$). Second, we aimed to select a subset of explanatory variables adequate to discriminate high H5N1 persistence assumed by all three indicators (Hc, HcP, and PrS). We computed the median AUC and R^2 of all three-variable models and ranked the

variables by their capacity to generate a high median AUC score in models predicting Hc and HcP combined with a high median R^2 in models predicting PrS. This yielded a restricted number of variables with superior predictive capacity to model HPAI H5N1 human cases and seasonal H5N1 persistence in poultry, wild birds, and/or humans.

The above subset of variables was subject to partitioning around a medoids method (Kaufman and Rousseeuw, 1990) to identify five agro-ecological clusters or niches. This method is a clustering algorithm which aims to break a multivariate data set into groups. The algorithm identifies a set of medoids, that can be defined as the point in cluster, whose average dissimilarity to all other points in the cluster is minimal. After finding the set of medoids, each point of the data set is assigned to the nearest medoid in the multivariate space. The method is comparable to the K-mean clustering algorithm but is considered to be more robust to outlying values.

The Euclidian distances to the medoid of the niche with the greatest number of human cases was estimated in the space defined by the subset of variables and mapped as an index of dissimilarity.

RESULTS

The outcome of the 364 models differed depending on the persistence variable modeled and the set of areas incorporated (Fig. 1). A broad coherence was obtained when plotting the 364 data pairs representing the Akaike's Information Criteria (AIC) for the Hc and PrS models, and using the all-countries set (Fig. 1, top left). The coherence disappeared when the province level data from China, Indonesia, and India were included (Fig. 1, top right). In contrast, AIC values for Hc and HcP models yielded a coherent cloud of values for the two sets of territories (Fig. 1, bottom left and right). On the basis of these results, we used the set with all-countries in the follow-up analyses to quantify PrS versus Hc models, and the set of all-countries plus China, Indonesia, and India provinces to quantify HcP versus Hc models.

Some of the explanatory variables provided consistently better results than others, but these variables were not the same across persistence measures (Fig. 2 and variable full names therein). Using the set of all-countries, variables with the greatest median R^2 for PrS models were AgPopDn, AgPopDnOI, ToPopDn, while variables with the greatest median AUC for Hc models were Pppc, AgPopDn,

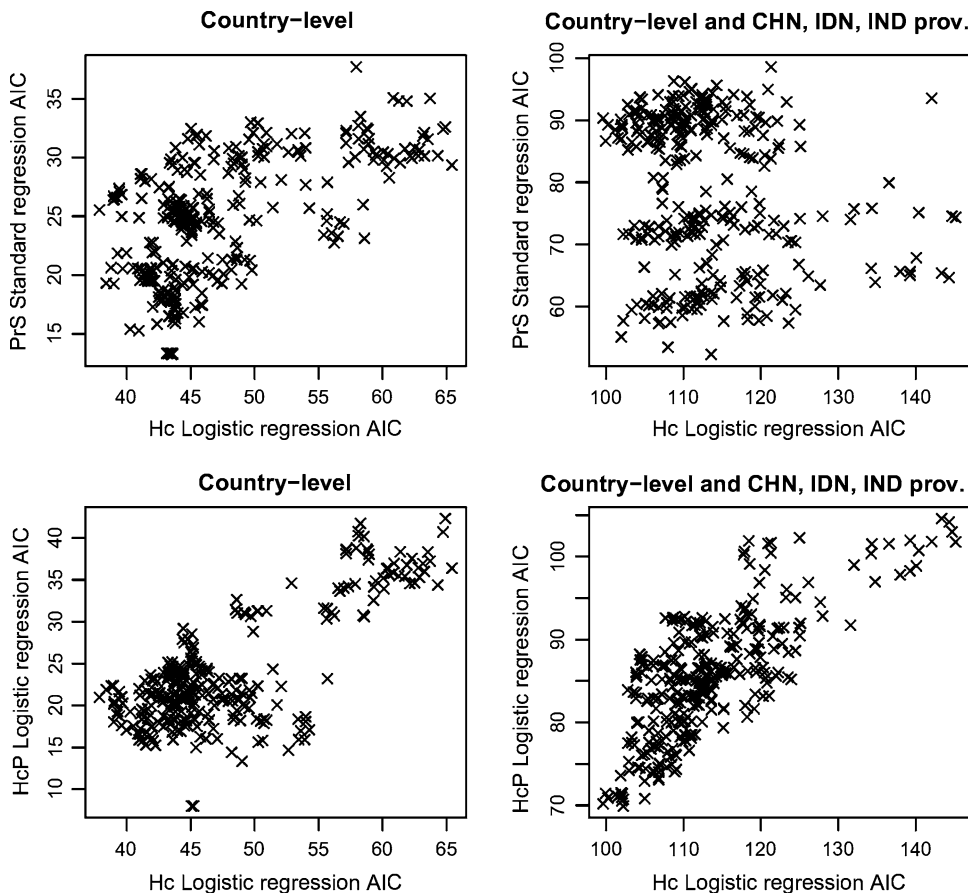


Figure 1. Akaike's information criteria of HPAI H5N1 model groups. Standard regression model AIC of HPAI H5N1 seasonal persistence (PrS) as function of logistic regression model AIC of HPAI H5N1 human case presence (Hc) for all countries (*top left*), and all countries and Chinese (CHN), Indonesian (IDN), and Indian (IND) provinces (*top right*). Logistic regression model AIC of HPAI H5N1 human cases persistence (HcP) as function of logistic regression model AIC of HPAI H5N1 human case presence (Hc) for all countries (*bottom left*), and all countries and Chinese (CHN), Indonesian (IDN), and Indian (IND) provinces (*bottom right*).

AgPopDnOI, and Tmax. For all countries plus China, Indonesia, and India provinces, the variables with the greatest median AUC for HcP models comprised DuChDn, ChDn, Tmin, Pppc, and variables with greatest median AUC for Hc models were DuChDn, Pppc, AgPopDn, DuDn, and ChDn. Based on these results, we selected a subset of six variables that provided best models along the two sets of spatial units and measures of persistence, i.e., located in the upper right quadrant of each plot: AgPopDn (agricultural population density; log-scale), AgPopDnOI (product of agricultural population density and output/input chicken, log-scale), DuChDn (duck by chicken density, log-scale), Pppc (purchasing power per capita), ChDn (chicken density, log-scale), and DuDn (duck density, log-scale). With this subset, we identified five agro-ecological niches by partitioning around medoids (see cluster plot, Fig. 3). Niches 4 and 1 included the greatest number of human cases per million people (Fig. 4, top left). In absolute numbers, niches 1–5 hosted 169, 0, 46, 164, and 0 human cases, and included 14, 0, 10, 12, and 0 spatial units with human cases, respectively. The medoid of the niche with the greatest number of human cases (niche 4) was the southern Chinese province of Henan.

The five niches differed in the patterns of main variables defining their structure (Figs. 3 and 4). Niche 5 was comprised largely of western European countries, Japan, and South Korea, and characterized by the highest purchasing power per capita and chicken production productivity, the lowest density of agricultural population, and an intermediate density of domestic ducks and chickens. Niche 2, with African countries, Russia, western India provinces, and a low number of Chinese and Indonesia provinces, included areas with the lowest duck and chicken densities, a low purchasing power per capita and chicken production productivity, and an intermediate density of agricultural population. Niche 3, with Laos, Cambodia, Nigeria, Turkey, Ukraine, and Romania, was defined by slightly higher purchasing power per capita, and intermediate densities of ducks, chickens, and agricultural population. Niche 4 (Egypt, the provinces of Java, and several Chinese provinces including Guangdong and Henan) had the greatest density of domestic ducks, chickens, and agricultural population density, and at the same time an intermediate productivity of chicken production and purchasing power, higher than niches 1, 2, and 3. Finally, niche 1 supported a high density of agricultural population,

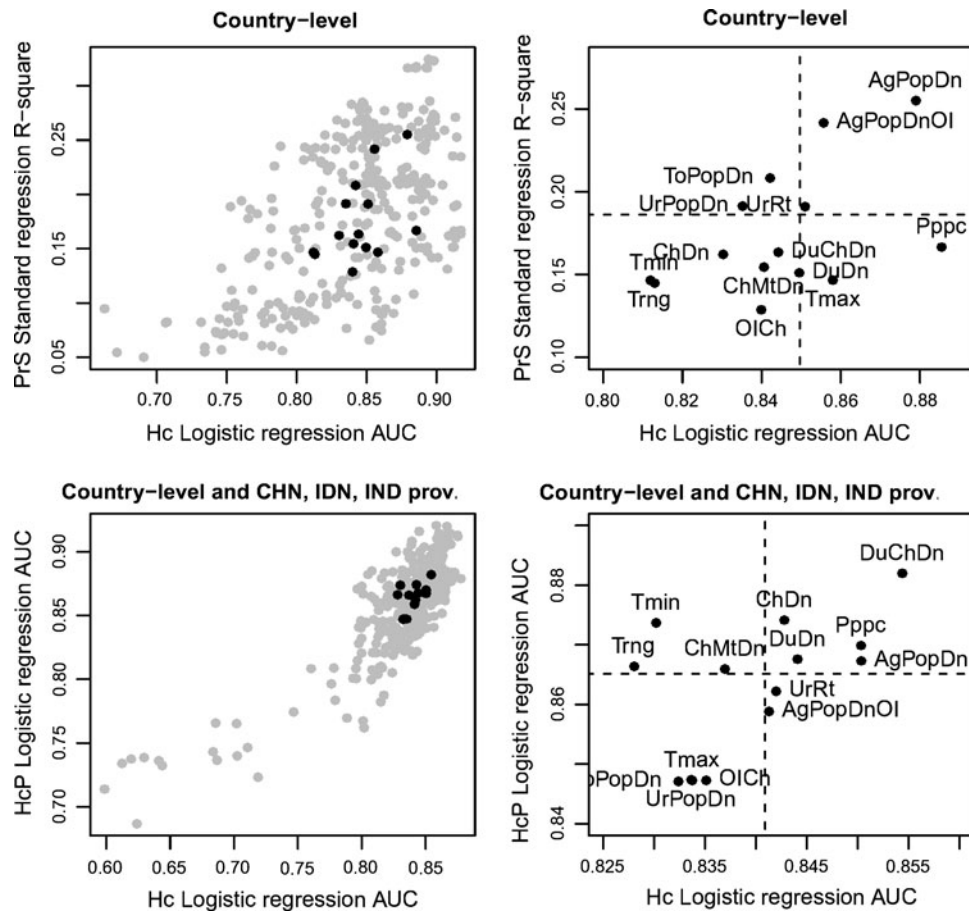


Figure 2. Goodness of fit indicators of HPAI H5N1 model groups for all countries (*top left and right*), and all countries and Chinese (CHN), Indonesian (IDN), and Indian (IND) provinces (*bottom left and right*). All-countries models (*top left and right*) are represented by the R^2 of the seasonal persistence standard regression model (PrS) as a function of the area under curve (AUC) of the human case presence logistic regression model (Hc). All-countries and province models (*bottom left and right*) are presented by the AUC of the human case persistence logistic regression model (HcP) as a function of the AUC of the human case presence logistic regression model (Hc). Left plots show all individual models (*gray dots*) and the median R^2 or AUC for

each collection of models containing a variable (*black dots*). The right plots show only R^2 or AUC median values for collection of models containing each variable (the *dotted lines* are the median R^2 or AUC estimated over all models). Variable abbreviations are as follows: *AgPopDn* agricultural population density, *AgPopDnOI* agricultural population density by output/input chicken, *ChDn* chicken density, *ChMtDn* chicken meat density, *DuDn* duck density, *DuChDn* duck by chicken density, *OICh* output/input chicken, *Pppc* purchasing power per capita, *ToPopDn* total population density, *Tmin*, *Tmax*, *Trng* temperature min, max, and range, *UrPopDn* urban population density, *UrRt* urban to total population ratio.

ducks, and chickens, but lower than in niche 4, and a low-to-intermediate chicken productivity and purchasing power per capita. Niche 1 included countries from Asia, such as Thailand, Vietnam, and several provinces of China and eastern India.

We mapped the environmental distance to Henan in the space defined by our set of six agro-ecological variables for all countries worldwide with the relevant data available (Fig. 5). The map highlights that some countries geographically distant from Henan share agro-ecological conditions: several China provinces, Java province, Bangladesh and some eastern India provinces, and Egypt in Africa. Several spatial units with agro-ecologies somewhat close to

that of Henan included Montenegro, Poland, and Hungary in Europe; Nigeria and Tunisia in Africa; Haiti, Guatemala, and El Salvador in the Americas; and Yemen and Pakistan in western and southern Asia. One should stress that the environmental distance to Henan is thought to be an indicator of persistence upon introduction, i.e., not the risk of introduction itself.

DISCUSSION

The epidemiology of HPAI H5N1 is characterized by a fundamental contradiction. On the one hand, HPAI in

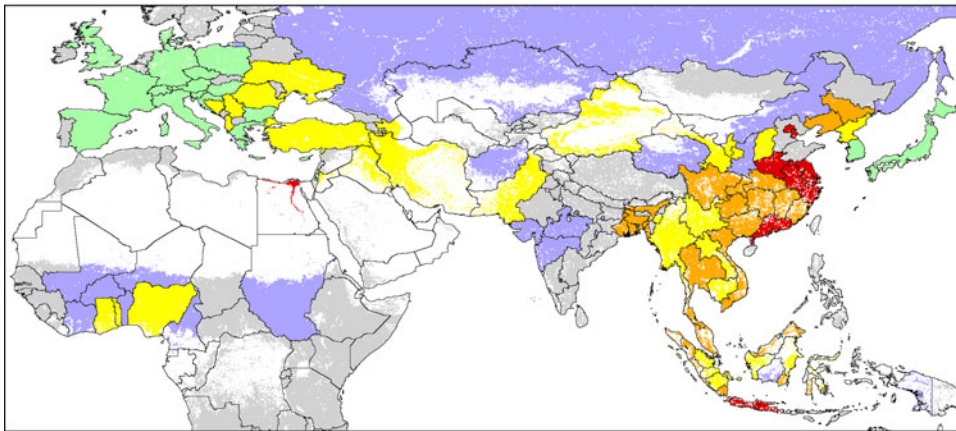
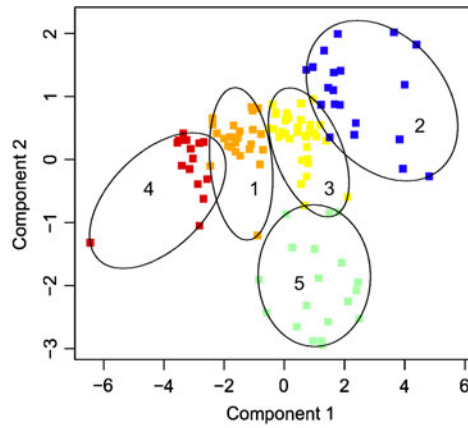


Figure 3. Cluster plot of the partitioning around medoids presented in the two principal component space (*top*), and map of the five agro-ecological clusters or niche in countries and Chinese, Indonesian, and Indian provinces (*bottom*) where H5N1 was reported (areas unsuitable to poultry production are masked in white).

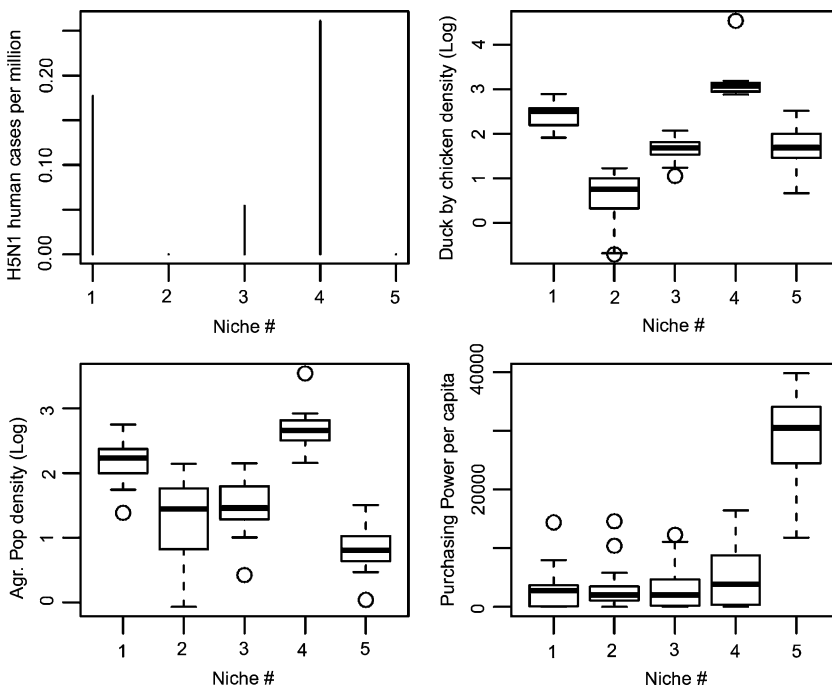


Figure 4. Characterization of each niche by the number of H5N1 human cases per million people (*top left*), the product of duck by chicken density (DuChDn, log scale [*top right*]), agricultural population density (AgPopDn, log-scale [*bottom left*]), and purchasing power per capita (Pppc [*bottom right*]).

poultry is highly infectious, spans a short generation time, tends toward boom-and-bust population dynamics (Villarreal et al., 2000), and would be expected to burn

faster through the supply of susceptible hosts than low pathogenic variants (Dieckmann et al., 2002; Ebert and Bull, 2008). On the other hand, the HPAI H5N1 epizootic,

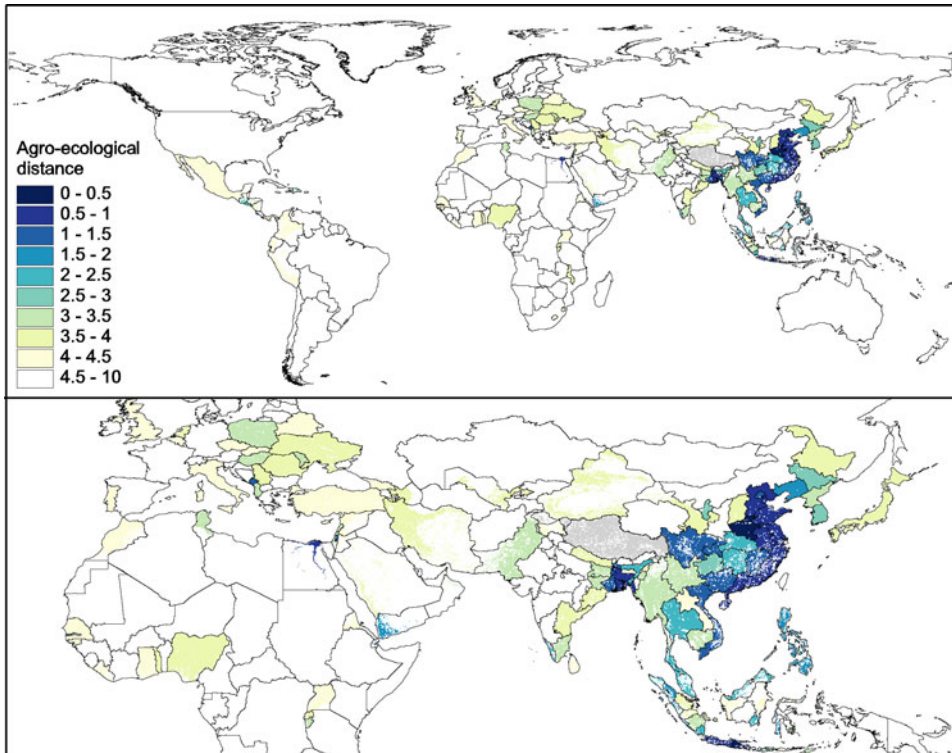


Figure 5. Distance to Henan (China) in the space defined by the variables agricultural population density (log-scale), duck density (log-scale), duck by chicken density (log-scale), chicken density (log-scale), the product of agricultural population density and chicken output/input ratio (log-scale), and purchasing power per capita (areas unsuitable to poultry production are masked in white).

now in its 12th year, remains the longest recorded highly pathogenic influenza outbreak in poultry.

A number of explanations have already been offered. First, the virus rotates across multiple host types (Olsen et al., 2006; Barrett et al., 2008; Wallace and Fitch, 2008), providing alternate hosts when one host species becomes too small to maintain transmission. Second, the virus experiences shifts in life history across host types, including a reduced virulence (and increased persistence) in duck reservoirs (Keawcharoen et al., 2008). Third, virulence can be an evolutionary stable strategy in situations where an ample supply of susceptibles is available, especially if such populations are distributed over wide geographical areas, as is the case with HPAI H5N1. Fourth, persistence in the physical environment may contribute to sapronotic transmission (Brown et al., 2007).

The results presented here suggest that some agro-ecologies are more capable of supporting HPAI H5N1 persistence than others, perhaps by better integrating the aforementioned mechanisms. The multivariate analysis discriminated localities with human cases and high persistence, and identified five agro-ecological niches, which, although derived independently of HPAI H5N1 data, vary by HPAI H5N1 persistence (Fig. 3).

How might the agro-ecological context support different levels of persistence? The models developed here,

using human cases as the indicator of persistence, provided fairly good predictive power (with AUC values of 0.88 and 0.89) and highlight a series of variables that, although correlated among themselves, differentiated HPAI's epidemiological environments. Two of these variables were previously identified as risk factors in country-level studies (Gilbert et al., 2008): the density of ducks, known to play a critical role in HPAI H5N1 epidemiology and evolution (Hulse-Post et al. 2005; Gilbert et al., 2006), and the density of agricultural population. The density of chickens influences disease flare up (Slingenbergh and Gilbert, 2008). The product of chicken by duck density appears most important in areas where both hosts act synergistically. Importantly, the current study also revealed an as yet little-addressed third dimension: purchasing power per capita and associated poultry production intensification (combined with agricultural population density).

A number of underlying processes may contribute to this result. Output/input chicken is a measure of poultry productivity and corresponding investment in biosecurity, shielding off invasive pathogens. Similarly, higher purchasing power per capita generally implies a restricted interface between people and their poultry, with reduced exposure to any sick or dead birds. Biosecurity, hygiene standards, and vaccination are common practice in modern poultry food chains, with a progressively diminishing reli-

ance on live bird marketing. However, it is important to note that the niches with the greatest number of human cases were not those with the lowest levels of poultry productivity. Niches 1 and 4, with the greater number of human cases, were characterized by the largest duck and agricultural population densities, but not by the lowest output/input level, nor by the lowest purchasing power per capita. In particular, niche 4, with the highest number of human cases per capita, mainly included countries or provinces with intermediate levels of productivity and low-to-medium purchasing power per capita (e.g., Egypt, Guangdong and Anhui provinces). These regions, in the midst of an economic transition, with intermediate levels of poultry production and purchasing power, and a geographic mosaic of old and new modes of production, may offer the virus the array of micro-niches needed to spread and evolutionarily radiate.

There is, in addition, a historical component (Wallace et al., in press). Prevalent modes of agriculture have deep historical roots. In China, rice cultivation marked the transition between Mesolithic foragers and the surplus food-producing economies of the Neolithic (Zong et al., 2007). Ducks were domesticated at least 3000 to 5000 years ago (Simoons, 1991; Cherry and Morris, 2008) and, by the middle of the Ming dynasty, were deployed for pest control in the rice paddies of the Pearl River Delta (Peng, 1994). The intensification of chicken and duck production, which appears to have driven the evolution of virulent influenza strains worldwide (Li et al., 2004; Vijaykrishna et al., 2008), was instituted in Asia mostly during the economic liberalization of the past 30 years (Wallace, 2009). The H5N1 duck–rice–poultry niche in Asia has resulted from a series of agro-ecological changes, with changes in agricultural practices, ancient (rice), modern (ducks), and present-day (poultry intensification), melding in a way that supports the present evolution of multiple virulent influenzas.

At the global level, the H5N1 niches defined here are clearly structured by geography, with the most vulnerable niches arrayed across South and East Asia, including along the Chinese lowlands and coastline into the river basins of Indochina and, further south, Indonesia. At the same time, other important regions are scattered across H5N1's range. For instance, Egypt and wetland areas in northern Nigeria support agro-ecologies similar to those that allow H5N1 to persist so well in China. The niches' geographic distribution, then, may offer another mechanism by which HPAI H5N1 persists. In acting as something of a western persistence hub, Egypt may assist seeding outbreaks into sub-Saharan Africa.

As is often the case, the details introduce several complications. First, understanding the distribution of HPAI H5N1 in relation to agro-ecological conditions depends heavily on the epidemiological records and conditions underlying the analyses. One way to circumvent the heterogeneity in disease report and epidemiological conditions was here to separately test multiple indicators of virus circulation, analyze these in relation to agro-ecological conditions, and assess whether they converge on a risk prediction. We found broadly coherent results at the global scale and country level, a coherence that disappeared when sub-national data on China, Indonesia, and India were introduced (Fig. 1 top).

The disparity arises to a large extent because the distribution of human cases in China's provinces does not match that of outbreaks reported in domestic poultry. At least two possibilities may explain this pattern: Human cases may truly be geographically disconnected from HPAI H5N1 circulation in poultry, or they are geographically correlated but HPAI H5N1 circulation remains undetected. When we consider the models based on human cases and human case persistence, we find a comparable scatter of models when comparing both (i) countries and (ii) all countries plus Chinese, Indonesian, and Indian provinces (Fig. 1 bottom). This indicates that the risk factors influencing the distribution of human cases outside China are broadly similar to those found within China. Given that the human cases outside China correlate also with the highest levels of seasonal persistence (e.g., Indonesia, Egypt, Vietnam), this suggests that the same risk factors should underpin distribution of seasonal persistence in China. Hence, there is a real possibility that HPAI H5N1 virus circulation in China's poultry remains, in part, undetected. In concordance with this possibility, the virus is detected at high prevalence for poultry sampled in live bird markets across southern Chinese provinces (Smith et al., 2006).

Second, in contrast to the models using human cases as a marker for persistent human infections, the models of seasonal persistence (all-countries set) offered a low level of predictability. The disparity is likely to stem from the difficulty to disentangle new introductions from local persistence and from the very high variability in disease surveillance in poultry and wild birds. For instance, Germany, a high-income country with a high chicken productivity, and comparatively abundant resources devoted to disease surveillance and biosecurity, repeatedly reported HPAI H5N1 cases several consecutive years, making it a country of intermediate persistence, comparable in our

data set to Myanmar, where disease surveillance and veterinary resources are sparse. The epidemiological situations of the two countries, however, are completely different. Germany was exposed to repeated introductions, most likely by wild birds on winter escape (Starick et al., 2008; Ottaviani et al., 2010). Clinically silent infections did apparently occur in industrial duck-fattening holdings in Germany (Harder et al., 2009), and small pockets of persistence in resident wild birds cannot be fully ruled out (Globig et al., 2009). However, it can be realistically assumed that continuous viral circulation could hardly have gone undetected in Germany, while the lack of veterinary infrastructure in Myanmar is such that the true level of virus circulation is difficult to assess.

Third, the kind of ecological niche modeling conducted here carries methodological caveats. First, it is important to avoid mistaking the agro-ecological niches identified here as deterministic in their effects on HPAI H5N1. The virus evolved during the study period, undergoing shifts in its transmission mode, virulence level, and host range. One may assume these changes are the means by which the virus adapts to the opportunities offered within each niche. A key test of the assumptions underlying niche modeling, and by extension a step toward modeling the pathway or trajectory of HPAI emergence, would be to establish if these evolutionary changes exhibit repeated patterns. Do localities with similar agro-ecological conditions select for HPAI H5N1 variants with similar phenotypes? The identification of different agro-ecological niches for HPAI H5N1 persistence is a first and necessary step toward identifying what could be niche-specific evolutionary signatures.

There are likely spatial signatures as well. In a situation where the training set consists of data pooled from heterogeneous sources representing different levels of detection capacity, a prediction of risk less dependent on the spatial distribution of the positive cases is helpful as it may highlight regions where agro-ecological conditions are more suitable; that is, where the disease has not spread but may have the potential to persist, or regions to where it may have spread but where the introduction events may have gone undetected. That is why distance-based metrics from a region of assumed risk were estimated here, and were preferred to the probabilities for HPAI presence arising from the more classically defined predictive models.

Two steps should be undertaken in any follow-up. First, with more phylogenetic data becoming available, it

should be possible to separate true persistence in a country from re-introduction, e.g., Chaichoune et al. (2009), which would enhance our capacity to identify the areas most susceptible to persistence. Second, we should more finely estimate risk indices at sub-national levels, as country-level data often hide considerable within-country variation. For example, duck production in Nigeria is concentrated in the central-northern part of the country, one of two sub-national sources Fusaro et al. (2010) identify, where outbreaks persisted more so than elsewhere in the country. With sub-national data, one may find Nigerian regions of considerable agro-ecological distance from niche 1 or 4, while the central-northern region may be closer than suggested by the national level data (Fig. 5). It is anticipated that many possible hotspots may be better identified with sub-national data, in particular in large countries that include considerable variation in agro-ecological conditions. Our Indonesian results (Fig. 3) point the way. All of Java's comparatively small provinces are estimated here to be members of the niche of greatest risk for persistence. Still-finer landscape- and even pixel-level analyses are clearly possible.

Importantly, agro-ecological conditions are dynamic. Expanding temporal coverage may help identify trajectories over which some countries, undergoing economic transitions, develop greater risk for HPAI persistence. Integrating such data sets with HPAI genetic sequences will help to more fully describe the agro-ecological space in which influenza evolves.

ACKNOWLEDGMENTS

The Food and Agricultural Organization thanks the studied countries for sharing their information. This study was supported by the Food and Agricultural Organization and National Institutes of Health, Fogarty International Center Grant R01TW00786901 (through the National Institutes of Health, National Science Foundation, Ecology of Infectious Diseases Program).

OPEN ACCESS

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

REFERENCES

- Abo M, Sy A, Alegbejo M (1998) Rice yellow mottle virus (RYMV) in Africa: evolution, distribution, economic significance on sustainable rice production and management strategies. *Journal of Sustainable Agriculture* 11:85–111
- Adhikari D, Chettri A, Barik SK (2009) Modelling the ecology and distribution of highly pathogenic avian influenza (H5N1) in the Indian subcontinent. *Current Science* 97:73–78
- Badan Pusat Statistik (2009) *Indonesia Statistics*. Available: <http://www.bps.go.id/eng/>
- Barrett L, Thrall P, Burdon J, Linde C (2008) Life history determines genetic structure and evolutionary potential of host-parasite interactions. *Trends in Ecology and Evolution* 23:678–685
- Brown J, Swayne D, Cooper R, Burns R, Stallknecht D (2007) Persistence of H5 and H7 avian influenza viruses in water. *Avian Diseases* 51:285–289
- Cecchi G, Ilemobade A, Le Brun Y, Hogerwerf L, Slingenbergh J (2008) Agro-ecological features of the introduction and spread of the highly pathogenic avian influenza (HPAI) H5N1 in northern Nigeria. *Geospatial Health* 3:7–16
- Chaichoune K, Wiriyarat W, Thitithanyanont A, Phonarknguen R, Sariya L, Suwanpakdee S, et al. (2009) Indigenous sources of 2007–2008 H5N1 avian influenza outbreaks in Thailand. *Journal of General Virology* 90:216
- Chakrabarti AK, Pawar SD, Cherian SS, Koratkar SS, Jadhav SM, Pal B, et al. (2009) Characterization of the influenza A H5N1 viruses of the 2008–09 outbreaks in India reveals a third introduction and possible endemicity. *PLoS ONE* 4:e7846. doi:10.1371/journal.pone.0007846
- Cherry P, Morris TR (2008) *Domestic Duck Production: Science and Practice*, Oxford, UK: CABI
- Dieckmann U, Metz J, Sabelis M, Sigmund K (2002) *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management*, Cambridge, UK: Cambridge University Press
- Ebert D, Bull J (2008) The evolution and expression of virulence. In: *Evolution in Health and Disease*, Sterans SC (editor), Oxford, UK: Oxford University Press
- Editorial Department of China Animal Husbandry Yearbook (2005) *China Animal Husbandry Yearbook*, Beijing, China: China Agricultural Press
- Fang LQ, de Vlas SJ, Liang S, Looman CW, Gong P, Xu B, et al. (2008) Environmental factors contributing to the spread of H5N1 avian influenza in mainland China. *PLoS ONE* 3:e2268
- FAO (2009a) *EMPRES-I Database*. Available: <http://empres-i.fao.org/empres-i/>
- FAO (2009b) *FAO Statistical Database*. Available: <http://faostat.fao.org/>
- FAO (2009c) *Global Livestock Production and Health Atlas (GLI-PHA)*. Available: <http://kids.fao.org/glipha/>
- Fasina F, Bisschop S, Webster R (2007) Avian influenza H5N1 in Africa: an epidemiological twist. *Lancet Infectious Diseases* 7:696–697
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24:38–49
- Fusaro A, Nelson MI, Joannis T, Bertolotti L, Monne I, Salviato A, et al. (2010) Evolutionary dynamics of multiple sublineages of H5N1 influenza viruses in Nigeria, 2006–2008. *Journal of Virology* 84:3239–3247
- Gilbert M, Chaitaweesub P, Parakamawongsa T, Premashtira S, Tiensin T, Kalpravidh W, et al. (2006) Free-grazing ducks and highly pathogenic avian influenza, Thailand. *Emerging Infectious Diseases* 12:227–234
- Gilbert M, Xiao X, Chaitaweesub P, Kalpravidh W, Premashtira S, Boles S, et al. (2007) Avian influenza, domestic ducks and rice agriculture in Thailand. *Agriculture Ecosystems and Environment* 119:409–415
- Gilbert M, Xiao X, Pfeiffer D, Epprecht M, Boles S, Czarnecki C, et al. (2008) Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia. *Proceedings of the National Academy of Sciences of the United States of America* 105:4769–4774
- Globig A, Baumer A, Revilla-Fernández S, Beer M, Wodak E, Fink M, et al. (2009) Ducks as sentinels for avian influenza in wild birds. *Emerging Infectious Diseases* 15:1633–1636
- Harder TC, Teuffert J, Starick E, Gethmann J, Grund C, Fereidouni S, et al. (2009) Highly pathogenic influenza virus (H5N1) in frozen duck carcasses, Germany, 2007. *Emerging Infectious Diseases* 15:272–279
- Hulse-Post D, Sturm-Ramirez K, Humberd J, Seiler P, Govorkova E, Krauss S, et al. (2005) Role of domestic ducks in the propagation and biological evolution of highly pathogenic H5N1 influenza viruses in Asia. *Proceedings of the National Academy of Sciences of the United States of America* 102:10682–10687
- International Monetary Fund (IMF) (2009) *IMF Data and Statistics*. Available: <http://www.imf.org/external/data.htm>
- Kapan D, Bennett S, Ellis B, Fox J, Lewis N, Spencer J, et al. (2006) Avian influenza (H5N1) and the evolutionary and social ecology of infectious disease emergence. *EcoHealth* 3:187–194
- Kaufman L, Rousseeuw P (1990) *Finding Groups in Data: An Introduction to Cluster Analysis*, New York: Wiley
- Keawcharoen J, van Riel D, van Amerongen G, Bestebroer T, Beyer W, van Lavieren R, et al. (2008) Wild ducks as long-distance vectors of highly pathogenic avian influenza virus (H5N1). *Emerging Infectious Diseases* 14:600–607
- Kilpatrick A, Chmura A, Gibbons D, Fleischer R, Marra P, Daszak P (2006) Predicting the global spread of H5N1 avian influenza. *Proceedings of the National Academy of Sciences of the United States of America* 103:19368–19373
- Li K, Guan Y, Wang J, Smith G, Xu K, Duan L, et al. (2004) Genesis of a highly pathogenic and potentially pandemic H5N1 influenza virus in eastern Asia. *Nature* 430:209–213
- Mitchell TD, Carter TR, Jones PD, Hukme M, New M (2004) *A Comprehensive Set of High-resolution Grids of Monthly Climate for Europe and the Globe: The Observed Record (1901–2000) and 16 Scenarios (2001–2100)*, Norwich, UK: Tyndall Center
- Office of the Registrar General and Census Commissioner, India (2009) *Census of India*. Available: <http://www.censusindia.gov.in/>
- Olsen B, Munster V, Wallensten A, Waldenstrom J, Osterhaus A, Fouchier R (2006) Global patterns of influenza A virus in wild birds. *Science* 312:384–388
- Ottaviani D, de la Rocque S, Khomenko S, Gilbert M, Newman SH, Roche B, et al. (2010) The cold European winter of 2005–2006 assisted the spread and persistence of H5N1 influenza virus in wild birds. *EcoHealth*. doi:10.1007/s10393-010-0316-z
- Oyana T, Dai D, Scott K (2006) Spatiotemporal distributions of reported cases of the avian influenza H5N1 (bird flu) in southern China in early 2004. *Avian Diseases* 50:508–515

- Peng S (1994) Brief discussion on the rise and decline of the Zheng Gao rice growing system in Zhujiang Delta. *History of Chinese Agriculture* 13:88–93
- Pfeiffer D, Minh P, Martin V, Epprecht M, Otte J (2007) An analysis of the spatial and temporal patterns of highly pathogenic avian influenza occurrence in Vietnam using national surveillance data. *The Veterinary Journal* 174:302–309
- Simoons FJ (1991) *Food in China: A Cultural and Historical Inquiry*, Boca Raton, FL: CRC Press
- Slingenbergh J, Gilbert M (2008) Do old and new forms of poultry go together? In: *Proceedings of the International Conference on Poultry in the 21st Century*, Bangkok, Thailand, November 2007, Rome: Food and Agriculture Organization
- Smith G, Fan X, Wang J, Li K, Qin K, Zhang J, et al. (2006) Emergence and predominance of an H5N1 influenza variant in China. *Proceedings of the National Academy of Sciences of the United States of America* 103:16936–16941
- Starick E, Beer M, Hoffmann B, Staubach C, Werner O, Globig A, et al. (2008) Phylogenetic analyses of highly pathogenic avian influenza virus isolates from Germany in 2006 and 2007 suggest at least three separate introductions of H5N1 virus. *Veterinary Microbiology* 128:243–252
- Vijaykrishna D, Bahl J, Riley S, Duan L, Zhang J, Chen H, et al. (2008) Evolutionary dynamics and emergence of panzootic H5N1 influenza viruses. *PLoS Pathogens* 4:e1000161. doi:10.1371/journal.ppat.1000161
- Villarreal L, DeFilippis V, Gottlieb K (2000) Acute and persistent viral life strategies and their relationship to emerging diseases. *Virology* 272:1–6
- Wallace R, Bergmann L, Hogerwerf L, Gilbert M (in press) Are influenzas in southern China byproducts of the region's globalising historical present? In: *Influenza and Public Health: Learning from Past Pandemics*, Craddock S, Giles-Vernick T, Gunn J (editors), London: Earthscan Press
- Wallace R, Hodac H, Lathrop R, Fitch W (2007) A statistical phylogeography of influenza A H5N1. *Proceedings of the National Academy of Sciences of the United States of America* 104:4473–4478
- Wallace RG (2009) Breeding influenza: the political virology of offshore farming. *Antipode* 41:916–951
- Wallace RG, Fitch WM (2008) Influenza A H5N1 immigration is filtered out at some international borders. *PLoS ONE* 3:e1697
- Ward M, Maftai D, Apostu C, Suru A (2008) Environmental and anthropogenic risk factors for highly pathogenic avian influenza subtype H5N1 outbreaks in Romania, 2005–2006. *Veterinary Research Communications* 32:627–634
- WHO (2009) *Confirmed Human Cases of Avian Influenza A(H5N1)*, Geneva: World Health Organization
- Williams RAJ, Fasina FO, Peterson AT (2008) Predictable ecology and geography of avian influenza (H5N1) transmission in Nigeria and West Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 102:471–479
- Williams RAJ, Peterson AT (2009) Ecology and geography of avian influenza (HPAI H5N1) transmission in the Middle East and northeastern Africa. *International Journal of Health Geographics* 8:47. doi:10.1186/1476-072X-8-47
- Wint W, Robinson T (2007) *Gridded Livestock of the World*, Rome: Food and Agricultural Organisation
- World Organization for Animal Health (2009) *World Animal Health Information Database (WAHID)*. Available: <http://www.oie.int/wahis/public.php?page=home>
- Zong Y, Chen Z, Innes J, Chen C, Wang Z, Wang H (2007) Fire and flood management of coastal swamp enabled first rice paddy cultivation in east China. *Nature* 449:459–462