INFLUENZA

Evolution-informed forecasting of seasonal influenza A (H3N2)

Xiangjun Du,[1] Aaron A. King,[2] Robert J. Woods,[3] Mercedes Pascual[1,4]*

Interpandemic or seasonal influenza A, currently subtypes H3N2 and H1N1, exacts an enormous annual burden both in terms of human health and economic impact. Incidence prediction ahead of season remains a challenge largely because of the virus' antigenic evolution. We propose a forecasting approach that incorporates evolutionary change into a mechanistic epidemiological model. The proposed models are simple enough that their parameters can be estimated from retrospective surveillance data. These models link amino acid sequences of hemagglutinin epitopes with a transmission model for seasonal H3N2 influenza, also informed by H1N1 levels. With a monthly time series of H3N2 incidence in the United States for more than 10 years, we demonstrate the feasibility of skillful prediction for total cases ahead of season, with a tendency to underpredict monthly peak epidemic size, and an accurate real-time forecast for the 2016/2017 influenza season.

INTRODUCTION

Interpandemic or seasonal influenza is responsible for an average of about 1 billion cases globally, including 3 to 5 million cases of severe illness and 250,000 to 500,000 deaths annually (1). Since its first occurrence in 1968, seasonal H3N2 influenza A has continually circulated in the human population and is currently the major cause of seasonal influenza morbidity and mortality (2). Subtype H1N1 and two major lineages of influenza B also contribute today to seasonal burden, with their importance varying with the particular season and region. The sustained "success" of influenza viruses responsible for seasonal outbreaks, especially H3N2, stems from their ability to evolve and escape the immune system by modifying their surface proteins (3). Phylogenetic trees depicting evolutionary changes in influenza (H3N2) viruses illustrate rapid drift with successive and punctuated replacement of one antigenic type by another (4, 5). The past decade has seen significant conceptual advances in the understanding of these phylogenetic patterns, enabled by computational and statistical advances at the interface of transmission dynamics and virus evolution (5–12). There is now a considerable interest in translating these conceptual advances into actual prediction at the population level that would inform the update of vaccines and epidemic preparedness. Because of virus evolution, the seasonal influenza vaccine varies in efficacy and requires frequent updates. The composition of the trivalent and quadrivalent vaccines, chosen from circulating H3N2, H1N1, and type B strains, is reviewed biannually (in February for the Northern Hemisphere), with changes proposed and coordinated by the World Health Organization (WHO) when strains considered most likely responsible for next year's outbreak differ sufficiently from those represented in the current vaccine. The decision results from a complex consultation process involving the WHO Global Surveillance Network and the Collaborating Centers for Reference and Research on Influenza.

The challenge of influenza prediction has progressed largely along two separate tracks. The first one seeks to predict evolutionary change with computational methods based on phylogenies and mutation patterns in the surface protein hemagglutinin (HA) (13–17). The result-

ing predictions of successful lineages and their relative frequencies for the future season do not, however, provide precise information on absolute incidence. The second track generates real-time incidence forecasts of influenza-like illness (ILI) on the basis of mathematical models describing the transmission dynamics of influenza viruses (18, 19). With data assimilation methods (20), these models must be fitted within each season because of season-to-season viral evolution (21–25). With such models, the fact that one needs to wait until the outbreak starts limits the lead time of epidemiological prediction.

Here, we bridge the gap between these two approaches and propose an epidemiological model specifically for seasonal H3N2 influenza that incorporates information on the evolutionary change of the virus. The resulting model is sufficiently parsimonious that parameter estimation based on retrospective surveillance records is possible. A key feature is its use of an evolutionary index of virus innovation constructed using readily available sequence data. The goal is to generate H3N2 incidence forecasts before the season begins, significantly earlier than what is currently possible. We illustrate two model formulations for H3N2 in the United States between 2002 and 2016 and produce a forecast for the upcoming 2016/2017 influenza season.

RESULTS

The monthly incidence variability from 2002 to 2016 for the whole United States is shown in Fig. 1A for reported cases of influenza type A, for subtypes H3N2 and H1N1 (including both seasonal H1N1 and pandemic H1N1), and type B. The time series are computed as the product of the ILI incidence rate, the proportion of ILI samples that tested positive for influenza, the subtype proportion, and the U.S. population size. Thus, incidence data are extrapolated to the U.S. population from outpatients in a network of health care providers, with untyped influenza specimens assigned to H3N2, H1N1, and B, respectively, on the basis of the proportions from the U.S. surveillance system (see Materials and Methods for details). The temporal variability of H3N2 exhibits seasonal outbreaks, whose size varies considerably from one year to the next. This interannual variability can result not only from epidemiological processes such as the loss of immunity to a specific variant of the virus (26) but also to a large degree, from the antigenic evolution of the virus (27) and the combined and complex interactions of the two (28, 29).
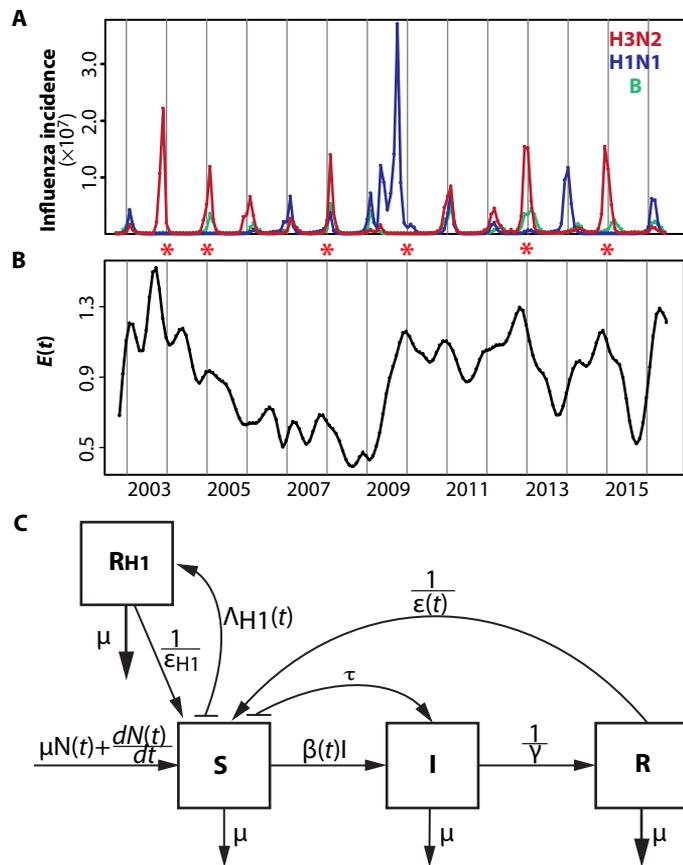
Fig. 1. Data and model. (**A**) Monthly influenza incidence data for the United States between October 2002 and June 2016. Red, blue, and green curves are for subtype H3N2, subtype H1N1, and type B, respectively. Seasons with an H3N2 antigenic cluster transition were marked with asterisks. (**B**) Monthly evolutionary change $E(t)$. $E(t)$ was calculated on the basis of epitope sites of HA, as a weighted sum of normalized amino acid distances (hamming distances) between strains in month t and previous strains. These distances were weighted by a decaying function back in time whose time scale was estimated as part of the model fitting. (**C**) Diagram for the epidemiological model. A classical SIRS epidemic model was used to represent the population dynamics of H3N2 incidence. Details are described in Materials and Methods.

We measure antigenic innovation or evolutionary change of the virus at a given time (relative to a window of time in the past) using a novel evolutionary index readily computed from available sequences for the HA1 domain. The idea is to use amino acid sequences to quantify the change in the virus' antigenic properties relative to those the human population has recently experienced. We therefore focus on sequences that encode epitopes known to be important for antibody binding and in which antigenic evolution is commonly observed (14, 30, 31). We define the index as a weighted average distance between the current viruses and their predecessors in the past. So that more recent sequences are weighted more heavily than older ones, the weights are taken to be a decaying function of the intersequence interval. Figure 1B illustrates the estimated index.

Before examining predictions of the "full" model that considers both epidemiology and the evolutionary index, we evaluate the ability of different models, encapsulating different degrees of complexity, to retrospectively explain the temporal patterns in

the data from 2002 to 2016. To establish a baseline against which to evaluate the full model, we begin with a simpler formulation for the population dynamics of H3N2 that includes only influenza epidemiology and the seasonality of transmission. This basic model follows the structure of the well-known compartmental susceptible-infected-recovered-susceptible (SIRS) formulation, which divides the population into classes for susceptible (nonimmune), infected, and recovered (immune) individuals (Fig. 1C). For the purpose of model comparisons, we rely initially on the whole temporal extent of the data (2002 to 2016) to fit the models and infer parameter values. Figure 2A shows that simulations of the basic model reproduce the average seasonality of incidence but fail to capture its interannual variation.

Several variants of the epidemiological model were considered next, starting with the effect of temporal variation in the incidence of H1N1. We find a significant negative correlation between annual incidence of H3N2 and H1N1 ($r = -0.60$, $P = 0.02$) but no clear relationship between annual incidence of type B and H3N2 or H1N1 ($r = 0.42$ and -0.33, $P = 0.13$ and 0.25, respectively). Given these findings and the observation that disease burden due to type B in humans is typically lower than that due to H3N2 or H1N1 in the United States (32, 33), we postulate an effect of H1N1, but not B, on the dynamics of H3N2 and do so as a covariate affecting the system as an external observed variable (in the "basic-H1" model, fig. S1A). We also allow measurement error to differ between the summer (1 April to 30 September) and winter (1 October to 31 March) seasons, reasoning that the reporting rate is more variable outside the winter (transmission) season (34). (Hereafter, we refer to the year/year + 1 influenza "season" as the period from 1 July of the current calendar year to 30 June of the next calendar year; for example, the 2008/2009 influenza season corresponds to the period between 1 July 2008 and 30 June 2009). Additionally, we relax the assumption of a linear dependence of the force of infection on the number of current infected individuals, allowing a nonlinear functional form and the potential for subexponential growth of the epidemic curve (in the "refined" model, fig. S1B). This functional form has been found effective in the modeling of a number of different infectious diseases, as a means of parameterizing processes operating at scales smaller than can be explicitly represented (35–37). Finally, we consider a version of the refined model in which the (nonparametric) periodic function in the transmission rate is replaced by a function of specific humidity (the "humidity" model, fig. S1C) (21, 38, 39). Although all these model variants improve the fit of the data (Table 1), they still fail to properly capture the interannual variability in incidence (fig. S1). Of these purely epidemiological models, we retain the best one, that is, the refined model, which includes H1N1, season-dependent measurement error, and subexponential epidemic growth, and turn next to whether the inclusion of an index of evolutionary change can improve on this foundation.

Because evolutionary novelty increases the probability that a virus escapes existing protective immunity and thereby achieves more efficient transmission, we incorporated the innovation index as an external driver of the SIRS model, by allowing it to modulate either the duration of immunity or the transmission rate, or both. The model that includes evolutionary change in the duration of immunity was best able to explain the data (the "continuous" model) (Table 1). Simulations with the estimated parameters further demonstrate the improved performance of this formulation, with the median incidence reproducing the main trends in the interannual variation (Fig. 2B). As a general result (Table 1), the models that incorporate evolutionary
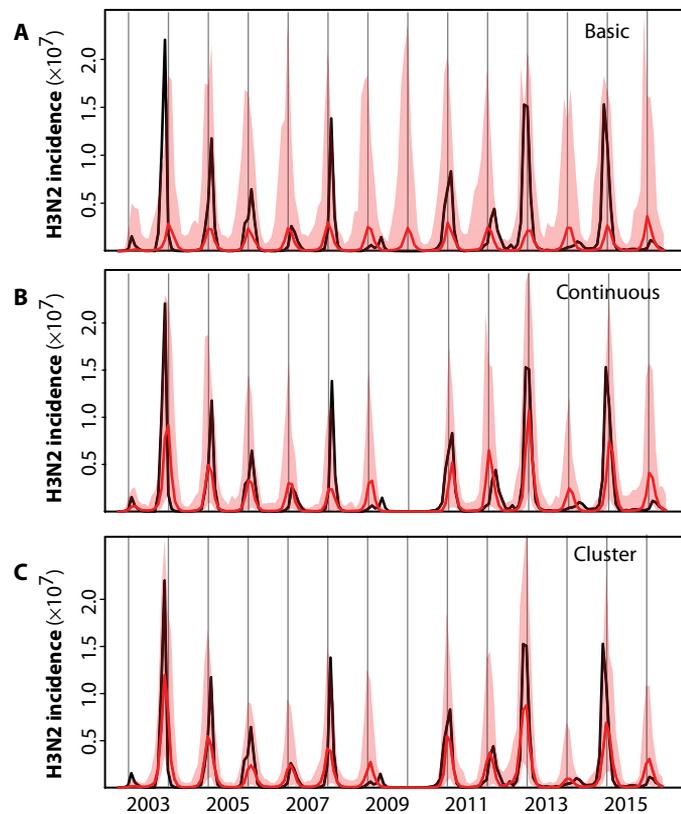
Fig. 2. Best model fits. Illustration of the best model fits for the (**A**) basic, (**B**) continuous, and (**C**) cluster models. See Table 1 and main text for the specification and statistical comparison of the different models considered. Here, results from 1000 monthly simulations starting with estimated initial conditions in October 2002 are shown for the respective models with the maximum likelihood estimates parameters, for the median (in red), and for the 95% uncertainty intervals (shaded red, for the 2.5 and 97.5% quantiles of the 1000 runs). For comparison, the observed monthly H3N2 incidence data for the United States are shown in black. The estimated reporting rate and its confidence intervals for both models including virus evolution are given in fig. S11.

information are significantly better than those that do not ($P < 0.05$, likelihood ratio test). Among the former, the models incorporating evolutionary change just in the immunity loss or in both this parameter and transmission (the "immunity loss/transmission" model) are comparable to each other but perform significantly better than that with an effect only in transmission (the "transmission" model). A degree of similarity between the two best models is to be expected in view of the fact that these two epidemiological parameters determine the overall infection rate, making it difficult to disentangle an effect on the number of susceptibles from one on the per-susceptible risk of infection. Our results here indicate that inclusion of the evolutionary driver as a modulator of the duration of immunity is sufficient.

Our best model so far included a smoothly changing measure of evolutionary change based purely on virus genotype. It is recognized, however, that the genotype-phenotype map for antigenic properties of the virus is discontinuous, such that virus strains cluster antigenically and switches between clusters affect the population dynamics and phylodynamics of H3N2 in punctuated fashion (4, 10). In particular, recorded antigenic cluster transitions are consistently followed by larger outbreaks (Fig. 1A). We found that higher values of our estimated index of evolutionary change tended to precede a winter season with an antigenic cluster transition (compare Fig. 1, A and B). This observation and what is known about the genotype-phenotype map of the virus (4, 40, 41) led us to a second evolutionary change index based on cluster transitions. In this "cluster" model, the effect of a cluster transition is punctuated and localized in time: The rate of immunity loss is varied only during the summer season that precedes the winter season with an antigenic cluster transition. Specifically, the rate of immunity loss only during that time becomes a function of the degree of evolutionary change; with this change now measured by comparing current virus sequences to those 2 years ago, a time scale characteristic of cluster transitions (4). This time scale is also consistent with our estimate of the effective evolutionary time (fig. S9), which is much shorter than the estimate of the basic latent period (fig. S10). The resulting model performs better than the best model with continuous evolutionary change (Table 1 and Fig. 2C), although the most significant difference is between purely epidemiological models and those that

Table 1. Model comparison. See Fig. 1 for a diagram of the general epidemiological model and the main text for a summary of the different models and their corresponding names used here.

Models	Epidemiology			Evolution		Number parameters	AIC
	H1N1	$\alpha \neq 1$ ρ_{winter} & ρ_{summer}	Seasonal forcing	Loss of immunity $w_\varepsilon \neq 0$	Transmission $w_\beta \neq 0$		
Basic	×	×	Periodic splines	×	×	14	4493
Basic-H1	√	×	Periodic splines	×	×	17	4478
Refined	√	√	Periodic splines	×	×	19	4471
Humidity	√	√	Specific humidity	×	×	15	4484
Immunity loss/transmission	√	√	Periodic splines	√	√	22	4450
Transmission	√	√	Periodic splines	×	√	21	4458
Immunity loss (continuous)	√	√	Periodic splines	√	×	21	4448
Immunity loss (cluster)	√	√	Periodic splines	√	×	20	4445

incorporate evolutionary information based on model comparisons with the Akaike information criterion (AIC) and the likelihood ratio test (Table 1).

With our best model, the cluster model, we now turn to the task of predicting H3N2 incidence before the influenza season begins. For this purpose, we divide the data into a "training" section (2002 to 2011) used to fit the models and an "out-of-fit" one (2011 to 2016) used to evaluate their prediction accuracy. The implementation of the model for forecasting purposes requires particular assumptions on the drivers in the system because these observed quantities will, by definition, not yet be available over the time windows we wish to predict. For H1N1 incidence, we make the simplifying assumption that monthly averages for this quantity over the training set provide a sufficient approximation. Additionally, for the cluster model, an upcoming season dominated by a novel antigenic cluster needs to be anticipated in the summer before the transmission season. For this, we developed and tested a rule based on a published genotype-phenotype map for prediction of new antigenic variants (40). Specifically, when the proportion of antigenic variants (PAVs) accumulated during the summer season exceeded a given threshold, we took this to be predictive of a cluster transition in the next winter season (figs. S2 and S3; see Materials and Methods for details). Figure 3A shows the resulting retrospective predictions together with observations for each of the last five influenza seasons from 2011/2012 to 2015/2016. Two criteria were used to quantify prediction accuracy. The first one compared the absolute monthly observed incidence to the medians of monthly predicted incidence from 1000 simulations. Predictions and observations are significantly correlated ($r = 0.87$ and $r^2 = 0.76$ for the monthly data; $r = 0.95$ and $r^2 = 0.91$ for the seasonal data; see Fig. 3 for the data that include the most recent 2016/2017 season and fig. S4). Moreover, the observations mostly fell within the 95% uncertainty intervals defined by the 2.5 and 97.5% quantiles of the simulations. Although the models tend to underpredict the absolute value of peak incidence, they do capture the overall interannual behavior of the trends reflected in both low and high seasons (Fig. 3B). The second criterion evaluates the model's ability to predict an outbreak season by computing the probability of surpassing a selected incidence value deemed high by public health practitioners. We are interested here in evaluating the risk of an anomalous "high" season relative to a typical season in the past and relative to a threshold level of cases of interest to public health. We consider first a threshold equal to the median of seasonal totals observed over the training data set. A given influenza season is forecasted as high or low risk level depending on the proportion of simulations that exceed the median, with the critical proportion that separates low and high levels chosen on the basis of receiver operating characteristic (ROC) curves (fig. S5). Specifically, we predict a high risk level when more than 40% of the 1000 simulations surpass the median. All five seasons were predicted accurately on the basis of this criterion (Table 2).

To further evaluate the prediction ability, we considered hindcast predictions, by removing one season at a time during the period from influenza season 2003/2004 to 2010/2011 and predicting its incidence, with the model parameters reestimated each time on the basis of the remaining data and exactly the same search strategy. Because there are multiannual correlations in the data, this test is less stringent and realistic than the one based on multiple sequential out-of-fit seasons at the end of the time series. Nevertheless, it allows us to extend prediction evaluation and demonstrates high prediction accuracy (Fig. 3 and table S1).

Encouraged by these results, we present a "real-time" forecast prepared before the 2016/2017 influenza season and based on the data

available up to June 2016 by the end of the 2015/2016 influenza season. Notable evolutionary change relative to previous viruses is indicated by our evolutionary index during the 2015/2016 influenza season (Fig. 1B), consistent with the observation that a number of antigenic variants were also accumulating over this period (fig. S2). Concurrently, H1N1 dominated the 2015/2016 influenza season, which would have resulted in an increased number of individuals susceptible to H3N2. The cluster model predicted a high risk level for seasonal H3N2 influenza in the 2016/2017 influenza season for the United States (Table 2), with a predicted annual incidence rate of 0.11 (95% confidence intervals, 0.07 and 0.15) (table S2), consistent with observations (Fig. 3 and Table 2).

Forecast results based on the continuous model also capture the interannual trend in the size of epidemics (fig. S6) and correctly predict risk levels above the 50% epidemic threshold (median) for the period between 2011 and 2017 (table S3). The quality of the forecasts is lower, however, than that obtained with our best model (the cluster model) (compare Fig. 3 and fig. S6). We also note that the 2016/2017 season is correctly forecasted as high risk but that its predicted peak size is overestimated and its predicted timing is earlier than that observed (fig. S6).

Another approach in the literature to predicting specific H3N2 incidence is based on its correlation with antigenic change, as measured by the hemagglutination inhibition assay (42). Our approach exhibits a higher correlation between seasonal observations and predictions (0.83 for the whole U.S. data set from 2002 to 2016 and 0.90 for the testing data set from 2011 to 2017 compared to 0.52 between antigenic change and H3N2 incidence for the period between 1998 and 2009) (Fig. 3B and fig. S7).

Finally, to further test the general approach, we applied it to one region of the United States—U.S. Department of Health and Human Services (HHS) region 3 (see Materials and Methods). This choice provides a smaller surveillance area for a fairly populous part of the country under moderate climate conditions. A robust result is that the models with evolutionary information are significantly better at capturing the dynamics of seasonal H3N2 influenza than those without it (table S4), although which particular evolutionary index is best can differ. Forecasts based on the best cluster and continuous models capture both the interannual variation of the outbreaks and disease risk for this U.S. region (fig. S8 and table S5).

DISCUSSION

Our results demonstrate the feasibility of improving epidemiological forecasts by incorporating information on evolutionary change into mechanistic models. Comparisons between models with and without this information show significant differences in their ability to capture the interannual variation in incidence data, which underscores the importance of evolutionary change in the epidemiological dynamics of seasonal influenza. Our best models are able to capture the temporal behavior of observed incidence for H3N2 in the recent past in the United States, and they provide the means to lengthen the lead time of prediction so that effective forecasts can be based in the summer, before the transmission season begins.

We predict interannual disease risk rather than finer-scale outbreak timing during the season; that is, we forecast whether the upcoming season will be anomalously large or small. Timing itself has been the target of existing within-season prediction efforts (22–25), which are better suited for this purpose and could be applied in tandem with our approach. Earlier forecasts of incidence dynamics further inform public health preparedness and can aid public health efforts by indicating
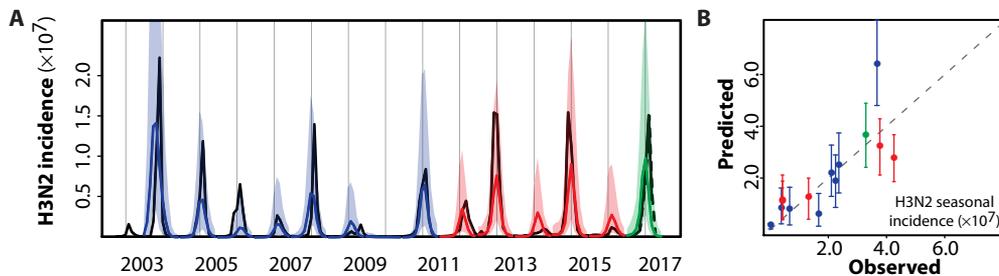
Fig. 3. H3N2 incidence forecasts based on the cluster model for the Unites States. (**A**) Monthly forecasts and (**B**) predictions versus observations of seasonal incidence totals. The different colors in both panels correspond to the following: leave-one-out cross-validation (for each influenza season from 2003/2004 to 2010/2011; in blue), retrospective forecasts (for each out-of-fit influenza season from 2011/2012 to 2015/2016; in red), and a real-time forecast for the 2016/2017 influenza season (in green). These predictions for cross-validations are simulated on a seasonal basis from estimated initial conditions starting in June and based on parameters estimated with all the data in the training data set with the target season excluded. The forecasts for the out-of-fit period are also simulated on a seasonal basis from estimated initial conditions starting in June but are based on parameters estimated with all the data up to that point in time. The average monthly H1N1 incidence from this training data set was used for forecasting purposes because the observation of this driver quantity would not be available. Similarly, the quantities specifying the evolutionary change of the virus were extrapolated as the sequences required for their computation would not be available (see Materials and Methods). The black curve is the monthly observed H3N2 incidence; the colored curve is the predicted monthly median incidence with shaded 95% uncertainty intervals (2.5 and 97.5% quantiles) from 1000 random simulations with the best models. The seasonal incidence is the sum of the monthly incidence for a specific influenza season. The observed incidence data for the 2016/2017 influenza season, which were not yet available when this study was conducted, are shown with the dotted line (and based on data from the weekly U.S. influenza surveillance report until week 14 ending on 8 April 2017). There is a significant correlation ($P < 0.05$) between observed incidence and median incidence, from the cross-validation forecasts for both monthly data ($r = 0.66$ and $r^2 = 0.44$) and seasonal data ($r = 0.88$ and $r^2 = 0.77$) and from both the retrospective and real forecasts ($r = 0.81$ and $r^2 = 0.67$ for monthly data, fig. S4; $r = 0.90$ and $r^2 = 0.81$ for seasonal data; Pearson's product moment correlations).

Table 2. H3N2 risk level forecasts for the United States based on the cluster model. Seasonal risk level for H3N2 influenza virus is defined as high or low for each season compared to a reference level defined as the median of the seasonal total H3N2 incidence cases in the corresponding training data set. We defined an observed season as H3N2 high risk, when the observed total H3N2 incidence surpasses the reference level; and an H3N2 low-risk season otherwise. For the forecasts, the percentage of 1000 simulations that exhibit an H3N2 high risk was obtained. When this percentage exceeded 40% (a level chosen based on fig. S5), we forecasted an H3N2 high-risk season. Otherwise, an H3N2 low-risk season was predicted.

Seasons	Observed	% High (1000 simulations)	Forecasts (>40% high)
2011/2012	Low	8.2	Low
2012/2013	High	99.6	High
2013/2014	Low	3.6	Low
2014/2015	High	99.9	High
2015/2016	Low	7.0	Low
2016/2017	High*	100.0	High

*On the basis of the updated data from the weekly U.S. influenza surveillance report until week 14 ending on 8 April 2017.

when to expect a surge in demand for health care resources and infrastructure. They can also contribute to the development of control strategies that take risk levels into consideration.

The use of evolutionary drivers in epidemiological dynamics follows an earlier study by Axelsen *et al.* (*43*) on long-term ILI incidence prediction in Tel Aviv, Israel. Their model incorporated the timing of known discrete antigenic changes in seasonal influenza and demonstrated the importance of considering these discrete antigenic jumps and their interaction with the waxing and waning of immunity levels in the population. Prediction of multiannual temporal patterns over multiple seasons was shown possible after the observation of such an event and as long as another one did not recur, which is an impediment to real-time forecasting. A number of more mechanistic models coupling evolution and transmission dynamics have also been developed to address theoretical questions on the phylodynamics of seasonal influenza (*6, 7, 10, 12*). Because these individual-based, stochastic formulations are high-dimensional and computationally expensive to work with, they are not well-suited for the repeated estimation of parameters from time series data on reported cases or for epidemiological prediction.

Here, we have shown that readily available sequences of the virus can be used to construct an evolutionary covariate in both continuous and discrete versions. In the continuous model, the time scale of the virus' antigenic evolution is relatively short, with an average effective time of about 16 months. At the same time, the estimated duration of homotypic immunity is relatively long—30 years or even longer. The importance of incorporating the short-term changes emphasizes the critical role of timely virological surveillance for identifying new emerging variants. Another intriguing observation related to the continuous measure of evolutionary change applied to our model is that the H1N1 pandemic of 2009 coincided with a valley in H3N2 fitness. This suggests that such times may provide a window of opportunity for the emergence of new types (including for cluster transitions of H3N2 itself). Thus, our evolutionary index, together with the proposed method for identifying and anticipating antigenic cluster transitions, could provide useful complements to the current surveillance system.

Our models made several simplifying assumptions, which can be investigated further and used to improve the approach in the future. The compartmental model did not include age structure (*44*) or social structure (*45*), proper inclusion of which might help to correct the underestimation of incidence peaks in this study. The evolutionary indices mainly consider antigenic change based on mutations in HA; these measures could be improved by further knowledge of antigenic phenotype (*46, 47*), including other viral segments like neuraminidase (*48*). Incorporating advances concerning antigenic prediction could be particularly relevant to further lengthen the lead time of forecasts, given the February timing of vaccine decisions for the next influenza season in the Northern Hemisphere. In addition, other factors affecting the fitness of the virus could be considered (*14, 49*), including receptor binding ability related to cell entry and transmission (*50*). Vaccination information could also be incorporated in the population dynamics.

Du *et al.*, *Sci. Transl. Med.* **9**, eaan5325 (2017) 25 October 2017

5 of 10

We decoupled the two-way interaction between subtypes H1N1 and H3N2 by including the effect of the former as a driver on the population dynamics of the latter. Although we observed that the dynamics of H3N2 was most strongly determined by its own evolutionary change, a more realistic model incorporating interactions between H3N2 and H1N1, and perhaps type B, could be considered (51). Our model can also be applied at a finer spatial resolution and to other regions, especially in Asia, where the likely source of evolutionary novelty for the seasonal influenza virus is to be found (5, 52). Preliminary investigation indicates that the general framework could be used in capturing and forecasting regional population dynamics of seasonal H3N2 influenza in the United States. Because immigration and emigration are also important processes in determining the local dynamics of seasonal influenza [for example, (53, 54)], a further step would consist of coupling regional dynamics to represent the effect of movement and the dependencies between adjacent regions. Similarly, at a larger scale, one could incorporate information on global influenza circulatory patterns into the model. Finally, ways to better extrapolate the evolutionary covariate itself beyond the summer should be addressed. Overall, the fact that incorporation of pathogen evolution into epidemiological models increases forecasting skill should embolden future efforts to further improve on the model presented here.

The limits to lead times in influenza prediction are not set by chaotic dynamics, as is the case for the weather system; they are determined instead by the stochastic nature of virus evolution. Thus, the increased availability of genetic sequence data from surveillance efforts around the globe can inform, in important ways, epidemiological prediction. One key limitation identified in our work is the low and variable number of sequences outside the transmission season, when this information would be most critical. Improvements to the general idea presented here will result from current efforts on purely evolutionary forecasting, which can provide better means to quantify antigenic change of the virus (13, 14, 17, 40, 46, 47) and to lengthen lead times further by concatenating evolutionary and evo-epidemiological prediction. Similarly, increased understanding of the virus' genotype-phenotype map will also further inform this kind of effort. Ultimately, in the same way that routine weather forecasting provided the impetus for much better sampling of the climate system, incidence prediction is computationally feasible but will ultimately depend on the quality, depth, and resolution of epidemiological and genetic surveillance.

MATERIALS AND METHODS
Study design
To predict H3N2 incidence ahead of the season for the United States at the national level, an epidemiological model was formulated that incorporates an effect of evolutionary change in the virus on selected parameters (waning immunity rate, transmission rate, or both of these). A series of formulations were compared via model selection to evaluate the importance of including such a change with two indices of evolutionary innovation based on genetic sequences, respectively, continuous and discrete in time. Model parameters were estimated with a sequential Monte Carlo method implementing an iterated filtering algorithm. The predictive ability of the best models is evaluated on the basis of both hindcasts and forecasts, by training the models with only a part of the data. Data from October 2002 to June 2016 were used in this study to focus on a period long enough to inform parameter inference and to avoid a considerably lower sampling effort of genetic sequences earlier on with no surveillance data for the summer season. The robustness of the proposed framework was tested by its further application to data from one U.S. region, the U.S. HHS region 3, including Delaware, the District of Columbia, Maryland, Pennsylvania, Virginia, and West Virginia.

Data
HA protein sequences of seasonal H3N2 influenza virus from the United States were downloaded from the Global Initiative on Sharing Avian Influenza Data (55). Sequences were then aligned with MUSCLE v3.7 using default settings (56). Undetermined amino acids were replaced by gaps, and only the HA1 domain was retained for further analysis. Outliers based on a reconstructed phylogenetic tree using FastTree 2 (57) with default settings were manually removed. Outpatient illness surveillance data and viral surveillance data were downloaded from FluView of the U.S. Centers for Disease Control and Prevention (CDC) (58). Seasonal and pandemic H1N1 influenza were combined together as seasonal H1N1 influenza, and the two lineages of seasonal B influenza were combined as seasonal B influenza. Untyped influenza-positive specimens were assigned to either H3N2, H1N1, or B, according to their proportions from typed specimens. The final weekly subtype-specific incidence was calculated as the product of ILI incidence rate, proportion of ILI samples that tested positive for influenza, subtype-specific proportion, and population size. Weekly incidence data were then aggregated to monthly data. Monthly national-level population estimates were downloaded from the U.S. Census Bureau (59). Specific humidity data for the United States were obtained from the National Land Data Assimilation System Phase 2 products (60). These primary measurements are provided on a 0.125° grid. National data were averaged over all grids for the monthly data. The epidemiological, virological, and population data for the regional study were downloaded from the same websites as for the national data. The regional data (ILI incidence rate, proportion of ILI samples that tested positive for influenza, and type/subtype-specific proportion) are sparser and therefore noisier, with frequent data missing during the low season. Thus, the incidence data were smoothed from peak (time point with the highest incidence during a winter season) to peak by local linear regression using a 4-week window.

Epidemiological model
A compartmental SIRS model was formulated that follows the flow of the population in susceptible, infected (and infectious), and recovered classes for seasonal H3N2 influenza, with the following equations:

$$\frac{dS}{dt} = \left(\mu N(t) + \frac{dN(t)}{dt}\right) - \beta(t)S\left(\frac{I}{N(t)}\right)^{\alpha} + \frac{R}{\varepsilon(t)} -$$

$$\tau - \mu S - \Lambda_{H1}(t) + \frac{R_{H1}}{\varepsilon_{H1}} \tag{1}$$

$$\frac{dI}{dt} = \beta(t)S\left(\frac{I}{N(t)}\right)^{\alpha} + \tau - \frac{I}{\gamma} - \mu I \tag{2}$$

$$\frac{dR}{dt} = \frac{I}{\gamma} - \frac{R}{\varepsilon(t)} - \mu R \tag{3}$$

Du *et al.*, *Sci. Transl. Med.* **9**, eaan5325 (2017)   25 October 2017

6 of 10

$$\frac{dR_{H1}}{dt} = \Lambda_{H1}(t) - \frac{R_{H1}}{\varepsilon_{H1}} - \mu R_{H1} \qquad (4)$$

S, I, and R denote the number of susceptible, infected, and recovered individuals in the population, respectively, and $N(t)$ is the population size at time t. The death rate μ was fixed to 0.015 per year (a life span of about 67 years). The total birth rate was quantified as $\left(\mu N(t) + \frac{dN(t)}{dt}\right)$ to reproduce the observed population increase over time. The exponent α is used to implement the nonlinear dependence of the force of infection on I and the resulting subexponential growth of seasonal epidemics. τ is the external importation rate of H3N2 influenza cases, which was fixed to 36.5 per year (one import per 10 days) (24). The contact rate $\beta(t)$ is given by:

$$\beta(t) = e^{\left(\sum_{i=1}^{6} w_i s_i(t) + w_\beta E(t)\right)} \frac{d\Gamma}{dt} \qquad (5)$$

which includes three components: (i) seasonality implemented through six b-splines $s_i(t)$ with coefficients w_i; (ii) evolutionary change $E(t)$ (see below for details) with coefficient w_β; and (iii) environmental noise simulated by a gamma distribution Γ (see Additional description of materials and methods in the Supplementary Materials for the model version where seasonality depends on humidity) (61). $\varepsilon(t)$ is the average latent period at time t, given by:

$$\varepsilon(t) = \varepsilon_0 e^{-w_\varepsilon E(t)} \qquad (6)$$

where ε_0 is the basic latent period, and w_ε is the scaling factor. γ is the average infectious period. An additional R_{H1} class was designed to track the reduction in susceptibles for H3N2 influenza due to cross-immunity and therefore the protected population due to infection by seasonal H1N1 influenza. The rate of susceptible individuals temporarily moved to the R_{H1} class was measured by:

$$\Lambda_{H1}(t) = \frac{C_{H1}(t)}{w_{H1}\varphi} \qquad (7)$$

where $C_{H1}(t)$ is the observed incidence due to seasonal H1N1 influenza, and φ is the reporting rate for H3N2, which is scaled here for H1N1 with the factor w_{H1}. The average latent period for individuals in the R_{H1} class returning to the susceptible class is denoted by ε_{H1}.

A measurement model is implemented that relies on a normal distribution and transforms the incidence in the transmission model to the actual observations by the reporting system (see Additional description of materials and methods in the Supplementary Materials). Specific modifications for the regional model are also described in this section of the Supplementary Materials.

Evolutionary index
Two quantities were formulated to incorporate information on virus evolution into the epidemiological model. The first one, which varies continuously in time, is described here; the second, which varies discontinuously to reflect the punctuated antigenic change of the virus, is described below under Antigenic cluster transitions and discrete evolutionary index.

An evolutionary index, $E(t)$, was used to measure the degree of evolutionary change of the virus at current time t (in months) compared to historical strains in the past. This index is therefore formulated as a weighted sum of normalized distance between current viruses and those in the past for amino acid sequences encoding epitopes of the HA protein. Its general expression is given by:

$$E(t) = \sum_{s=0}^{n} \widetilde{d(s,t)} e^{-\frac{s+1}{\theta}} \qquad (8)$$

where $\widetilde{d(s,t)}$ denotes a normalized distance between sequences at month t and those in previous season s back in time, and n is the total number of previous seasons including the current one for which we set $s = 0$. Here, $n = 19$ and $t = 1, \ldots l$ (where l is the length of the incidence time series in months starting from October 2002). In the formula for $E(t)$, changes relative to more recent viruses circulating in the population have a stronger weight than those relative to earlier viruses, and this weight decays exponentially back in time (scaled by θ in Eq. 8). The final computation requires a series of additional steps because of geographic and temporal biases in numbers of sampled sequences and the consideration of months and seasons in computing distances (see details and biological interpretation in Additional description of materials and methods in the Supplementary Materials).

Antigenic cluster transitions and discrete evolutionary index
Antigenic cluster transitions were identified on the basis of influenza season summary reports for seasons between 2003 and 2013 (62) and Morbidity and Mortality Weekly Reports for seasons between 2013 and 2016 (63–66) from U.S. CDC. An antigenic cluster transition is defined here as a season when more than half of the viruses were antigenically similar to the new vaccine strain. On the basis of this criterion, antigenic cluster transition seasons are 2003/2004, 2004/2005, 2007/2008, 2009/2010, 2012/2013, and 2014/2015. For the 2006/2007 influenza season, although there was a change in vaccine strain, most of the circulating strains were not antigenically similar to the new vaccine strain (62). As a result, no antigenic cluster transition was considered for the 2006/2007 influenza season. For the 2013/2014 influenza season, although there was a change in vaccine strain from A/Victoria/361/2011 to A/Texas/50/2012, the vaccine strains were antigenically similar (66). Thus, here too, this vaccine strain change was not considered an indication of an antigenic cluster transition.

For the model that incorporates evolutionary change through cluster transitions, cluster transitions are identified first on the basis of sequences, and when one such transition is identified, the degree of the change is then quantified. This change influences model parameters (rate of immunity loss here) in the form of a step function; that is, in the cluster model, an effect of evolutionary change in the virus is only applied to the epidemiological parameters when a cluster transition is identified (or predicted), and at this time, a measure of the magnitude of the change in the virus relative to the recent past is used as a covariate. In practice, when the model is implemented for actual forecasting, the emergence and establishment of a new cluster need to be predicted first. This first step is implemented with a cluster transition rule based on a genotype-phenotype map for H3N2 previously published for both identifying and predicting antigenic cluster transitions purely based on sequence data without having to rely on antigenic assay data (40). This

first step allows the prediction of an antigenic cluster transition ahead of the transmission season during the summer. After the identification (for retrospective data) or prediction (for out-of-fit data/forecasting purposes) of the emergence of a new antigenic cluster, a second step quantifies how much the new viruses differ from those in the recent past. For this purpose, virus sequences between September and the previous second season (our "reference" season) are compared for two reasons: First, the average effective period back in time estimated with parameter θ for the evolutionary index is about 2 years (fig. S9), and second, this period is close to the average time of cluster replacements (4).

The following quantity was defined to specifically measure evolutionary change (at the monthly scale) for the month of September:

$$
E_D(t) = \begin{cases} d_{2,\text{Sept.}} & \text{for } t \text{ within summer seasons that precede winter} \\ & \text{seasons with an antigenic cluster transition} \\ 0 & \text{otherwise} \end{cases}
$$

(9)

where $d_{2,\text{Sept.}}$ was calculated as a normalized distance based on epitope regions of HA1 between sequences for September of the current season and sequences of the previous second season (for example, $d_{2,\text{Sept.}}$ in 2000 was calculated on the basis of sequences from September of current 2000/2001 influenza season and 1998/1999 influenza season, that is $s = 2$ in the notation introduced above). Because there are typically a limited number of sequences for the summer season, distances calculated on the basis of those sequences can vary considerably depending on sampling effort. Therefore, a final value of $d_{2,\text{Sept.}}$ was obtained with a similar procedure than that used for calculating $\widetilde{d(s,t)}$ above. For the purpose of fitting the time series data, the resulting distance quantity $E_D(t)$ was introduced in the model as described in Eq. 6. When the goal is instead that of specifically forecasting ahead of the transmission season, cluster transitions need to be anticipated, and the value of $E_D(t)$ was computed accordingly. The approach taken for this purpose is described in the section below on Forecasts.

Forecasts

Forecasts of incidence dynamics with the cluster model were obtained for a given season through three steps. First, the cluster model was trained on the basis of the data set before the target season using exactly the same procedure described above. Second, the model with the best likelihood was chosen and used to estimate the initial conditions for the forecasts. These initial conditions require estimates of the "hidden" (unmeasured) variables in the model, S, I, R, and R_{H1}, in June of the year for which the predictions will be made. Maximum likelihood by iterated filtering (MIF) allows one to estimate these variables as the filtered states of the system at that time, given the observations of the cases. Third, forecasts were obtained through forward simulations for the target season with the selected model and starting from these initial conditions.

Because the system of equations in our models is nonautonomous, including external variables or drivers whose values must be specified independently from the dynamics, real forecasts (versus retrospective ones) require specifying assumptions for these drivers whose observation in the future is, by definition, unavailable. Specifically, information is required on both $C_{H1}(t)$ in $\Lambda_{H1}(t)$ and $d_{2,\text{Sept.}}$ in $E_D(t)$ for the simulations. First, for $C_{H1}(t)$, the average monthly value from the training data set was used. This is a simplification and a rough approximation but a reasonable choice in the absence of modeling the coupled dynamics of H1N1 and H3N2. This approximation is most likely to be sufficient,

when predictability of our models is evaluated, if the population dynamics of H3N2 is most strongly driven by its own evolutionary change rather than by the precise levels of H1N1. Second, the evolutionary change between new and old clusters is also needed, and this quantity can be extrapolated linearly as the value for September using the same procedure for calculating $d_{2,\text{Sept.}}$ above but based on sequences only from October of the previous year to June of the current year.

In addition, the weekly reports used to identify antigenic cluster transitions will not yet be available at the time of forecasting. Previously, a naïve Bayes model was developed as a genotype-phenotype to translate sequence changes in HA to antigenic changes and specifically calculate an odds ratio measuring antigenic similarity between a pair of strains, given their sequences (40). On the basis of this method, quarterly measures of the PAVs were calculated here. This quantity provides the proportion of pairs that are antigenically different (odds ratio < 1) among all pairs between sequences of a specific quarter and those in the previous year. Again, the same subsampling process was applied but on the basis of time points for quarters, not months.

To predict a cluster transition, rules were examined that combine a local increase in PAV with this quantity exceeding a threshold. This cutoff value was selected on the basis of the ROC curve (lower bound of best accuracy, see fig. S3). Again, because a new antigenic cluster would need to be established before the winter season, PAV was evaluated in the third quarter (1 July to 30 September) right before the coming winter season: PAV for the third quarter was linearly extrapolated on the basis of the data in the previous three quarters (first and second quarter of the current year and the fourth quarter in the previous year, that is, from 1 October to 30 June). Specifically, on the basis of data from the current season, if PAV increases from the first quarter to the second quarter and crosses the selected cutoff of 0.11 for the third quarter (whose value was extrapolated as explained above), an antigenic cluster transition was identified for the upcoming season. If PAV decreases from the first quarter to the second quarter but is still higher than the cutoff value of 0.11 for the third quarter (again, extrapolated), a cluster transition will also be assigned for the upcoming season, with the additional requirement that an antigenic cluster transition was not already assigned for the current season. Otherwise, no cluster transition will be anticipated for the upcoming season. With this rule and for the data between 2002 and 2012, all antigenic cluster transitions would have been correctly anticipated (true-positive rate equal to 1 and false-positive rate equal to 0) with the cutoff value of PAV ranging from 0.11 to 0.2 based on the ROC analysis. The lower bound of 0.11 was applied as the cutoff in this study to select for high sensitivity to the emergence of a different antigenic type. The specifics of forecasting with the continuous model are provided in the Additional description of materials and methods in the Supplementary Materials.

Cross-validation (hindcast)

To test the predicting ability further, and specifically for different patterns of alternating dominant subtypes in adjacent seasons, a cross-validation analysis was conducted for the training data set itself covering the period from 2002 to 2011. For each influenza season from 2003/2004 to 2010/2011, one at a time, the cluster model was fitted de novo by removing the target year and using the same search strategy than for the full data set before. With the resulting specific parameters, a prediction was generated using the same strategy described above and calculating the mean H1N1 incidence with all years except the target one. In practice, the fitting of the model is implemented in MIF, with

the parameters prevented from performing a random walk during the window of time that contains the target year (so that the corresponding data are not used in the filtering process) and with the likelihood evaluated by setting $L(t) = 0$ (in the log scale) in Eq. 10 below.

Statistical analyses

The resulting SIRS model was fitted to the data using MIF in the R package pomp (67, 68). Both parameters and initial conditions (for S, I, R, and R_{H1}) were estimated on the basis of the likelihood function:

$$L(t) = \mathrm{pnormal}\Big(C(t), \varphi I, \rho I\Big) \tag{10}$$

with $C(t) > 0$, where $C(t)$ is the reported number of cases from the model, ϕ is the reporting rate, and ρ defines the SD of the measurement error as proportional to the size of the infected population. If $C(t) = 0$, $L(t)$ was set to a very small value equal to $-10{,}000$ (log scale) as a penalty. The search of parameters and initial conditions was started with a grid of 10,000 random combinations sampled using the Latin hypercube sampling (69) from wide ranges. This step was followed with additional phases of increasingly localized searches. Confidence intervals were estimated separately for each parameter with the target parameter fixed at different values while allowing optimization of all other parameters also using MIF (67, 68).

AIC was used to measure goodness of a model (70). The AIC score takes into account model complexity and penalizes the likelihood based on the number of parameters. It was calculated here as AIC = $2k - 2\ln(L)$, where k is the number of parameters and L is the maximum likelihood. The likelihood ratio test was used for model selection for nested models (71). Pearson's product moment correlation was used for all correlation analyses. Throughout the manuscript, a significance level of 0.05 was applied.

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/9/413/eaan5325/DC1
Additional description of materials and methods
Fig. S1. Illustration of the best model fits for alternative models.
Fig. S2. Proportion of antigenic variants.
Fig. S3. ROC curve for predicting antigenic cluster transitions based on the training data set covering the period from 2002 to 2011 in United States.
Fig. S4. Comparison of monthly observations and forecasts generated with the cluster model for the out-of-fit period covering the period from 2011 to 2017.
Fig. S5. ROC curve for choosing the percentage cutoff applied to risk level prediction for the cluster model.
Fig. S6. Forecasts based on the continuous model for the Unites States.
Fig. S7. Scatter diagram for seasonal simulations and observations for the U.S. data set covering the period from 2002 to 2016 based on the cluster model.
Fig. S8. H3N2 incidence forecasts for the U.S. HHS region 3.
Fig. S9. Log-likelihood profiling of the average effective time θ.
Fig. S10. Log-likelihood profiling of the basic average latent time ε₀.
Fig. S11. Log-likelihood profiling of the reporting rate φ.
Fig. S12. ROC curve for choosing the percentage cutoff applied to risk level prediction for the continuous model.
Fig. S13. ROC curves for choosing the percentage cutoff applied to risk level prediction for the U.S. HHS region 3.
Table S1. H3N2 risk level forecasts based on leave-one-out cross-validation using the cluster model for the period between 2003 and 2011.
Table S2. Observed and predicted H3N2 seasonal incidence rate for the United States based on the cluster model for the period between 2011 and 2017.
Table S3. H3N2 risk level forecasts based on the continuous model for the United States.
Table S4. Model comparison based on data from the U.S. HHS region 3.
Table S5. H3N2 risk level forecasts for the U.S. HHS region 3.

REFERENCES AND NOTES

1. World Health Organization, Influenza (Seasonal), Fact Sheet No. 211, (2014); www.who.int/mediacentre/factsheets/fs211/en/.
2. W. W. Thompson, D. K. Shay, E. Weintraub, L. Brammer, N. Cox, L. J. Anderson, K. Fukuda, Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA 289, 179–186 (2003).
3. M. I. Nelson, E. C. Holmes, The evolution of epidemic influenza. Nat. Rev. Genet. 8, 196–205 (2007).
4. D. J. Smith, A. S. Lapedes, J. C. de Jong, T. M. Bestebroer, G. F. Rimmelzwaan, A. D. M. E. Osterhaus, R. A. M. Fouchier, Mapping the antigenic and genetic evolution of influenza virus. Science 305, 371–376 (2004).
5. T. Bedford, S. Riley, I. G. Barr, S. Broor, M. Chadha, N. J. Cox, R. S. Daniels, C. P. Gunasekaran, A. C. Hurt, A. Kelso, A. Klimov, N. S. Lewis, X. Li, J. W. McCauley, T. Odagiri, V. Potdar, A. Rambaut, Y. Shu, E. Skepner, D. J. Smith, M. A. Suchard, M. Tashiro, D. Wang, X. Xu, P. Lemey, C. A. Russell, Global circulation patterns of seasonal influenza viruses vary with antigenic drift. Nature 523, 217–220 (2015).
6. D. Zinder, T. Bedford, S. Gupta, M. Pascual, The roles of competition and mutation in shaping antigenic and genetic diversity in influenza. PLOS Pathog. 9, e1003104 (2013).
7. T. Bedford, A. Rambaut, M. Pascual, Canalization of the evolutionary trajectory of the human influenza virus. BMC Biol. 10, 38 (2012).
8. K. Koelle, P. Khatri, M. Kamradt, T. B. Kepler, A two-tiered model for simulating the ecological and evolutionary dynamics of rapidly evolving viruses, with an application to influenza. J. R. Soc. Interface 7, 1257–1274 (2010).
9. M. Recker, O. G. Pybus, S. Nee, S. Gupta, The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. Proc. Natl. Acad. Sci. U.S.A. 104, 7711–7716 (2007).
10. K. Koelle, S. Cobey, B. Grenfell, M. Pascual, Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. Science 314, 1898–1903 (2006).
11. N. M. Ferguson, D. A. T. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, D. S. Burke, Strategies for mitigating an influenza pandemic. Nature 442, 448–452 (2006).
12. N. M. Ferguson, A. P. Galvani, R. M. Bush, Ecological and immunological determinants of influenza evolution. Nature 422, 428–433 (2003).
13. R. A. Neher, C. A. Russell, B. I. Shraiman, Predicting evolution from the shape of genealogical trees. eLife 3, 10.7554/eLife.03568 (2014).
14. M. Łuksza, M. Lässig, A predictive fitness model for influenza. Nature 507, 57–61 (2014).
15. R. M. Bush, C. A. Bender, K. Subbarao, N. J. Cox, W. M. Fitch, Predicting the evolution of human influenza A. Science 286, 1921–1925 (1999).
16. I. A. Wilson, N. J. Cox, Structural basis of immune recognition of influenza virus hemagglutinin. Annu. Rev. Immunol. 8, 737–771 (1990).
17. M. Lässig, V. Mustonen, A. M. Walczak, Predicting evolution. Nat. Ecol. Evol. 1, 0077 (2017).
18. E. O. Nsoesie, J. S. Brownstein, N. Ramakrishnan, M. V. Marathe, A systematic review of studies on forecasting the dynamics of influenza outbreaks. Influenza Other Respir. Viruses 8, 309–316 (2014).
19. J.-P. Chretien, D. George, J. Shaman, R. A. Chitale, F. E. McKenzie, Influenza forecasting in human populations: A scoping review. PLOS ONE 9, e94130 (2014).
20. W. Yang, A. Karspeck, J. Shaman, Comparison of filtering methods for the modeling and retrospective forecasting of influenza epidemics. PLOS Comput. Biol. 10, e1003583 (2014).
21. W. Yang, M. Lipsitch, J. Shaman, Inference of seasonal and pandemic influenza transmission dynamics. Proc. Natl. Acad. Sci. U.S.A. 112, 2723–2728 (2015).
22. W. Yang, B. J. Cowling, E. H. Y. Lau, J. Shaman, Forecasting influenza epidemics in Hong Kong. PLOS Comput. Biol. 11, e1004383 (2015).
23. J. Shaman, A. Karspeck, W. Yang, J. Tamerius, M. Lipsitch, Real-time influenza forecasts during the 2012–2013 season. Nat. Commun. 4, 2837 (2013).
24. J. Shaman, A. Karspeck, Forecasting seasonal outbreaks of influenza. Proc. Natl. Acad. Sci. U.S.A. 109, 20425–20430 (2012).
25. W. Yang, D. R. Olson, J. Shaman, Forecasting influenza outbreaks in boroughs and neighborhoods of New York City. PLOS Comput. Biol. 12, e1005201 (2016).
26. L. Stone, R. Olinky, A. Huppert, Seasonal dynamics of recurrent epidemics. Nature 446, 533–536 (2007).
27. S. K. Greene, E. L. Ionides, M. L. Wilson, Patterns of influenza-associated mortality among US elderly by geographic region and virus subtype, 1968–1998. Am. J. Epidemiol. 163, 316–326 (2006).
28. B. T. Grenfell, O. G. Pybus, J. R. Gog, J. L. N. Wood, J. M. Daly, J. A. Mumford, E. C. Holmes, Unifying the epidemiological and evolutionary dynamics of pathogens. Science 303, 327–332 (2004).
29. S. Gandon, T. Day, C. J. E. Metcalf, B. T. Grenfell, Forecasting epidemiological and evolutionary dynamics of infectious diseases. Trends Ecol. Evol. 31, 776–788 (2016).
30. D. C. Wiley, I. A. Wilson, J. J. Skehel, Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature 289, 373–378 (1981).
31. M.-S. Lee, J. S.-E. Chen, Predicting antigenic variants of influenza A/H3N2 viruses. Emerg. Infect. Dis. 10, 1385–1390 (2004).

32. S. Caini, Q. S. Huang, M. A. Ciblak, G. Kusznierz, R. Owen, S. Wangchuk, C. M. P. Henriques, R. Njouom, R. A. Fasce, H. Yu, L. Feng, M. Zambon, A. W. Clara, H. Kosasih, S. Puzelli, H. A. Kadjo, G. Emukule, J.-M. Heraud, L. W. Ang, M. Venter, A. Mironenko, L. Brammer, T. Q. Mai Le, F. Schellevis, S. Plotkin, J. Paget; Global Influenza B Study, Epidemiological and virological characteristics of influenza B: Results of the Global Influenza B Study. *Influenza Other Respir. Viruses* **9** (suppl. 1), 3–12 (2015).

33. W. Paul Glezen, J. K. Schmier, C. M. Kuehn, K. J. Ryan, J. Oxford, The burden of influenza B: A structured literature review. *Am. J. Public Health* **103**, e43–e51 (2013).

34. R. Yaari, G. Katriel, A. Huppert, J. B. Axelsen, L. Stone, Modelling seasonal influenza: The role of weather and punctuated antigenic drift. *J. R. Soc. Interface* **10**, 20130298 (2013).

35. G. Chowell, L. Sattenspiel, S. Bansal, C. Viboud, Mathematical models to characterize early epidemic growth: A review. *Phys. Life Rev.* **18**, 66–97 (2016).

36. M. Roy, M. Pascual, On representing network heterogeneities in the incidence rate of simple epidemic models. *Ecol. Complex.* **3**, 80–90 (2006).

37. B. F. Finkenstädt, A. Morton, D. A. Rand, Modelling antigenic drift in weekly flu incidence. *Stat. Med.* **24**, 3447–3461 (2005).

38. M. Lipsitch, C. Viboud, Influenza seasonality: Lifting the fog. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3645–3646 (2009).

39. E. R. Deyle, M. C. Maher, R. D. Hernandez, S. Basu, G. Sugihara, Global environmental drivers of influenza. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13081–13086 (2016).

40. X. Du, L. Dong, Y. Lan, Y. Peng, A. Wu, Y. Zhang, W. Huang, D. Wang, M. Wang, Y. Guo, Y. Shu, T. Jiang, Mapping of H3N2 influenza antigenic evolution in China reveals a strategy for vaccine strain recommendation. *Nat. Commun.* **3**, 709 (2012).

41. J. B. Plotkin, J. Dushoff, S. A. Levin, Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6263–6268 (2002).

42. T. Bedford, M. A. Suchard, P. Lemey, G. Dudas, V. Gregory, A. J. Hay, J. W. McCauley, C. A. Russell, D. J. Smith, A. Rambaut, Integrating influenza antigenic dynamics with molecular evolution. *eLife* **3**, e01914 (2014).

43. J. B. Axelsen, R. Yaari, B. T. Grenfell, L. Stone, Multiannual forecasting of seasonal influenza dynamics reveals climatic and evolutionary drivers. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9538–9542 (2014).

44. E. A. Belongia, M. E. Sundaram, D. L. McClure, J. K. Meece, J. Ferdinands, J. J. VanWormer, Waning vaccine protection against influenza A (H3N2) illness in children and older adults during a single season. *Vaccine* **33**, 246–251 (2015).

45. S. V. Scarpino, A. Allard, L. Hébert-Dufresne, The effect of a prudent adaptive behaviour on disease transmission. *Nat. Phys.* **12**, 1042–1046 (2016).

46. R. A. Neher, T. Bedford, R. S. Daniels, C. A. Russell, B. I. Shraiman, Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E1701–1709 (2016).

47. C. Li, M. Hatta, D. F. Burke, J. Ping, Y. Zhang, M. Ozawa, A. S. Taft, S. C. Das, A. P. Hanson, J. Song, M. Imai, P. R. Wilker, T. Watanabe, S. Watanabe, M. Ito, K. Iwatsuki-Horimoto, C. A. Russell, S. L. James, E. Skepner, E. A. Maher, G. Neumann, A. I. Klimov, A. Kelso, J. McCauley, D. Wang, Y. Shu, T. Odagiri, M. Tashiro, X. Xu, D. E. Wentworth, J. M. Katz, N. J. Cox, D. J. Smith, Y. Kawaoka, Selection of antigenically advanced variants of seasonal influenza viruses. *Nat. Microbiol.* **1**, 16058 (2016).

48. N. M. Bouvier, P. Palese, The biology of influenza viruses. *Vaccine* **26** (suppl. 4), D49–D53 (2008).

49. K. Koelle, D. A. Rasmussen, The effects of a deleterious mutation load on patterns of influenza A/H3N2's antigenic evolution in humans. *eLife* **4**, e07361 (2015).

50. S. E. Hensley, S. R. Das, A. L. Bailey, L. M. Schmidt, H. D. Hickman, A. Jayaraman, K. Viswanathan, R. Raman, R. Sasisekharan, J. R. Bennink, J. W. Yewdell, Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. *Science* **326**, 734–736 (2009).

51. E. Goldstein, S. Cobey, S. Takahashi, J. C. Miller, M. Lipsitch, Predicting the epidemic sizes of influenza A/H1N1, A/H3N2, and B: A statistical method. *PLOS Med.* **8**, e1001051 (2011).

52. C. A. Russell, T. C. Jones, I. G. Barr, N. J. Cox, R. J. Garten, V. Gregory, I. D. Gust, A. W. Hampson, A. J. Hay, A. C. Hurt, J. C. de Jong, A. Kelso, A. I. Klimov, T. Kageyama, N. Komadina, A. S. Lapedes, Y. P. Lin, A. Mosterin, M. Obuchi, T. Odagiri, A. D. M. E. Osterhaus, G. F. Rimmelzwaan, M. W. Shaw, E. Skepner, K. Stohr, M. Tashiro, R. A. M. Fouchier, D. J. Smith, The global circulation of seasonal influenza A (H3N2) viruses. *Science* **320**, 340–346 (2008).

53. V. Charu, S. Zeger, J. Gog, O. N. Bjørnstad, S. Kissler, L. Simonsen, B. T. Grenfell, C. Viboud, Human mobility and the spatial transmission of influenza in the United States. *PLOS Comput. Biol.* **13**, e1005382 (2017).

54. D. Zinder, T. Bedford, E. B. Baskerville, R. J. Woods, M. Roy, M. Pascual, Seasonality in the migration and establishment of H3N2 Influenza lineages with epidemic growth and decline. *BMC Evol. Biol.* **14**, 272 (2014).

55. P. Bogner, I. Capua, N. J. Cox, D. J. Lipman, A global initiative on sharing avian flu data. *Nature* **442**, 981 (2006).

56. R. C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

57. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLOS ONE* **5**, e9490 (2010).

58. Centers for Disease Control and Prevention, FluView (2017); http://gis.Cdc.Gov/grasp/fluview/fluportaldashboard.html.

59. United States Census Bureau, Population and housing unit estimates (2016); www.census.gov/programs-surveys/popest.html.

60. National Aeronautics and Space Administration, Land Data Assimilation System (2017); http://ldas.gsfc.nasa.gov/.

61. A. Bhadra, E. L. Ionides, K. Laneri, M. Pascual, M. Bouma, R. C. Dhiman, Malaria in northwest India: Data analysis via partially observed stochastic differential equation models driven by Lévy noise. *J. Am. Stat. Assoc.* **106**, 440–451 (2011).

62. Centers for Disease Control and Prevention, Past Weekly Surveillance Reports (2017); www.cdc.gov/flu/weekly/pastreports.htm.

63. S. L. Davlin, L. Blanton, K. Kniss, D. Mustaquim, S. Smith, N. Kramer, J. Cohen, C. N. Cummings, S. Garg, B. Flannery, A. M. Fry, L. A. Grohskopf, J. Bresee, T. Wallis, W. Sessions, R. Garten, X. Xu, A. I. A. Elal, L. Gubareva, J. Barnes, D. E. Wentworth, E. Burns, J. Katz, D. Jernigan, L. Brammer, Influenza activity—United States, 2015–16 season and composition of the 2016–17 influenza vaccine. *Morb. Mortal. Wkly. Rep.* **65**, 567–575 (2016).

64. G. D. Appiah, L. Blanton, T. D'Mello, K. Kniss, S. Smith, D. Mustaquim, C. Steffens, R. Dhara, J. Cohen, S. S. Chaves, J. Bresee, T. Wallis, X. Xu, A. I. A. Elal, L. Gubareva, D. E. Wentworth, J. Katz, D. Jernigan, L. Brammer; Centers for Disease Control and Prevention (CDC), Influenza activity—United States, 2014–15 season and composition of the 2015–16 influenza vaccine. *Morb. Mortal. Wkly. Rep.* **64**, 583–590 (2015).

65. S. Epperson, L. Blanton, K. Kniss, D. Mustaquim, C. Steffens, T. Wallis, R. Dhara, M. Leon, A. Perez, S. S. Chaves, A. A. Elal, L. Gubareva, X. Xu, J. Villanueva, J. Bresee, N. Cox, L. Finelli, L. Brammer; Influenza Division, National Center for Immunization and Respiratory Diseases, CDC, Influenza activity—United States, 2013–14 season and composition of the 2014–15 influenza vaccines. *Morb. Mortal. Wkly. Rep.* **63**, 483–490 (2014).

66. Centers for Disease Control and Prevention (CDC), Influenza activity—United States, 2012–13 season and composition of the 2013–14 influenza vaccine. *Morb. Mortal. Wkly. Rep.* **62**, 473–479 (2013).

67. A. A. King, D. Nguyen, E. L. Ionides, Statistical inference for partially observed Markov processes via the R Package pomp. *J. Stat. Softw.* **69**, 1–43 (2016).

68. E. L. Ionides, C. Bretó, A. A. King, Inference for nonlinear dynamical systems. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18438–18443 (2006).

69. R. L. Iman, J. C. Helton, J. E. Campbell, An approach to sensitivity analysis of computer models. Part I. Introduction, input variable selection and preliminary variable assessment. *J. Qual. Technol.* **13**, 174–183 (1981).

70. H. Bozdogan, Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika* **52**, 345–370 (1987).

71. J. Neyman, E. S. Pearson, On the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. R. Soc. Lond.* **231**, 289–337 (1933).

Citation: X. Du, A. A. King, R. J. Woods, M. Pascual, Evolution-informed forecasting of seasonal influenza A (H3N2). *Sci. Transl. Med.* **9**, eaan5325 (2017).

Science Translational Medicine

Evolution-informed forecasting of seasonal influenza A (H3N2)

Xiangjun Du, Aaron A. King, Robert J. Woods and Mercedes Pascual

Foreseeing future flu

Although two strains of influenza A have been circulating globally for decades, these strains mutate quickly to try to escape host immunity. Many models exist to predict influenza seasons, but Du *et al.* have now incorporated evolutionary information to try to forecast upcoming H3N2 seasons in advance. After developing the model, they used it to predict that H3N2 in the pending 2016/2017 season will be at a higher incidence than has been seen recently and will involve substantial antigenic change in the hemagglutinin. Being able to predict this type of information could affect influenza countermeasures worldwide.

ARTICLE TOOLS	http://stm.sciencemag.org/content/9/413/eaan5325
SUPPLEMENTARY MATERIALS	http://stm.sciencemag.org/content/suppl/2017/10/23/9.413.eaan5325.DC1
RELATED CONTENT	http://stm.sciencemag.org/content/scitransmed/9/388/eaaj1928.full http://stm.sciencemag.org/content/scitransmed/9/412/eaam5752.full http://stm.sciencemag.org/content/scitransmed/9/382/eaaf9194.full http://science.sciencemag.org/content/sci/358/6362/450.full http://science.sciencemag.org/content/sci/358/6362/496.full http://science.sciencemag.org/content/sci/359/6373/255.full http://science.sciencemag.org/content/sci/359/6373/290.full http://science.sciencemag.org/content/sci/359/6373/277.full http://stm.sciencemag.org/content/scitransmed/10/428/eaan8405.full http://stm.sciencemag.org/content/scitransmed/10/441/eaaj2345.full http://science.sciencemag.org/content/sci/361/6404/755.full http://science.sciencemag.org/content/sci/361/6404/800.full http://science.sciencemag.org/content/sci/361/6408/1207.full http://science.sciencemag.org/content/sci/362/6410/75.full http://science.sciencemag.org/content/sci/362/6410/29.full http://science.sciencemag.org/content/sci/362/6414/511.full http://science.sciencemag.org/content/sci/362/6414/598.full http://science.sciencemag.org/content/sci/362/6419/1094.full http://science.sciencemag.org/content/sci/363/6428/676.full http://science.sciencemag.org/content/sci/363/6431/eaar6221.full http://stm.sciencemag.org/content/scitransmed/11/502/eaau5485.full http://stm.sciencemag.org/content/scitransmed/11/515/eaax5866.full http://science.sciencemag.org/content/sci/367/6480/852.full http://science.sciencemag.org/content/sci/367/6480/eaau0810.full http://science.sciencemag.org/content/sci/368/6497/1335.full http://stm.sciencemag.org/content/scitransmed/12/554/eabc1126.full http://science.sciencemag.org/content/sci/early/2020/08/12/science.aaz8432.full

Use of this article is subject to the Terms of Service