# Supporting Online Materials for
# Evolution-informed forecasting of seasonal influenza A (H3N2)

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

* To whom correspondence should be addressed. Email: pascualmm@uchicago.edu

**This PDF file includes:**

Tables S1 to S5

Figs. S1 to S14

**Table S1**. H3N2 risk level forecasts based on leave-one-out cross validation using the cluster model for the period between 2003 and 2011. For a target season, incidence risk level is defined as high or low compared to a reference level chosen here as the median (50% quantile) of the seasonal total incidence over all seasons excluding the target season. We defined an observed season as high risk, when the observed total incidence surpass the reference level; and an observed season as low risk, otherwise. For the forecasts, the percentage of 1000 simulations that exhibit a high risk was obtained. When this percentage exceeded 40% (chosen based on Fig. S5), we forecasted a high risk season. Otherwise, a low-risk season was predicted.

| Seasons | Observed | % High (1000 simulations) | Forecasts (>40%: high) |
|---|---|---|---|
| 2003/2004 | High | 100.0 | High |
| 2004/2005 | High | 85.9 | High |
| 2005/2006 | High | 3.8 | Low |
| 2006/2007 | Low | 2.3 | Low |
| 2007/2008 | High | 97.5 | High |
| 2008/2009 | Low | 1.8 | Low |
| 2009/2010 | Low | 0.0 | Low |
| 2010/2011 | High | 98.7 | High |

**Table S2**. Observed and predicted H3N2 seasonal incidence rate for the US based on the cluster model for the out-of-fit period between 2011 and 2016 with the [2.5%, 97.5%] and [10%, 90%] intervals. Seasonal incidence rate is defined as the seasonal total incidence normalized by the average population size in that season.

| Seasons | Observed | Predicted | | |
|---|---|---|---|---|
| | Incidence rate | Incidence rate | [2.5%,97.5%] | [10%,90%] |
| 2011/2012 | 0.04 | 0.04 | [0.01,0.06] | [0.02,0.05] |
| 2012/2013 | 0.13 | 0.08 | [0.06,0.12] | [0.07,0.10] |
| 2013/2014 | 0.01 | 0.03 | [0.01,0.06] | [0.02,0.05] |
| 2014/2015 | 0.13 | 0.10 | [0.07,0.13] | [0.08,0.12] |
| 2015/2016 | 0.01 | 0.03 | [0.01,0.07] | [0.02,0.05] |
| 2016/2017 | 0.10* | 0.11 | [0.07,0.15] | [0.09,0.13] |

* Based on the updated data from the weekly US influenza surveillance report until week 14 ending on April 8, 2017

**Table S3.** H3N2 risk level forecasts based on the continuous model for the US. Seasonal risk level is defined as high or low for each influenza season between 2011 and 2017 , by comparison to a reference level defined here as the 50% quantile of the seasonal total incidence in the training dataset. 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 a H3N2 high risk was obtained. When this percentage exceeded 75% (chosen based on Fig. S13), we forecasted an H3N2 high risk season. Otherwise, an H3N2 low risk season was predicted.

| Seasons | Observed | % High (1000 simulations) | Forecasts (>70%: high) |
|---|---|---|---|
| 2011/2012 | Low | 50.4 | Low |
| 2012/2013 | High | 99.8 | High |
| 2013/2014 | Low | 58.2 | Low |
| 2014/2015 | High | 100.0 | High |
| 2015/2016 | Low | 67.9 | Low |
| 2016/2017 | High* | 100.0 | High |

* Based on the updated data from the weekly US influenza surveillance report until week 14 ending on April 8, 2017

**Table S4.** Model comparison based on data from the HHS region 3. See Table 1 for the different models and for how to compare them. Based on the likelihood ratio test, models with evolutionary change, including the cluster model, are significantly better than those without it (the refined model here, shaded). The continuous model is significantly better than other models with continuous evolutionary change (the transmission model and the immunity loss/transmission model), and the models with continuous evolutionary change are significantly better than the cluster model.

| Models | Epidemiology | | Evolution | | Number Parameters | AIC |
|---|---|---|---|---|---|---|
| | H1N1 | $\alpha \neq 1$ $\rho_{winter}$ & $\rho_{summer}$ | Loss of Immunity $\omega_{\varepsilon} \neq 0$ | Transmission $\omega_{\beta} \neq 0$ | | |
| Refined | √ | √ | × | × | 18 | 3729 |
| Immunity loss/Transmission | √ | √ | √ | √ | 21 | 3691 |
| Transmission | √ | √ | × | √ | 20 | 3706 |
| Immunity loss (continuous) | √ | √ | √ | × | 20 | 3684 |
| Immunity loss (cluster) | √ | √ | √ | × | 19 | 3712 |

**Table S5**. H3N2 risk level forecasts for the HHS region 3. Seasonal risk level is defined as high or low for each season of the out-of-fit period (2011-2017) compared to a reference level, defined here as the 50% quantile of the seasonal total incidence in the corresponding training dataset. 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 a H3N2 high risk was obtained. When this percentage exceeded 50% (chosen based on Fig. S14), we forecasted a H3N2 high risk season. Otherwise, a H3N2 low risk season was predicted.

| Seasons | Observed | Cluster Model | | Continuous Model | |
|---|---|---|---|---|---|
| | | % High (1000 simulations) | Forecasts (>50%: high) | % High (1000 simulations) | Forecasts (>50%: high) |
| 2011/2012 | Low | 0.1 | Low | 80.7 | High |
| 2012/2013 | High | 100.0 | High | 100.0 | High |
| 2013/2014 | Low | 0.0 | Low | 25.8 | Low |
| 2014/2015 | High | 91.0 | High | 100.0 | High |
| 2015/2016 | Low | 0.0 | Low | 36.0 | Low |
| 2016/2017 | High* | 100.0 | High | 100.0 | High |

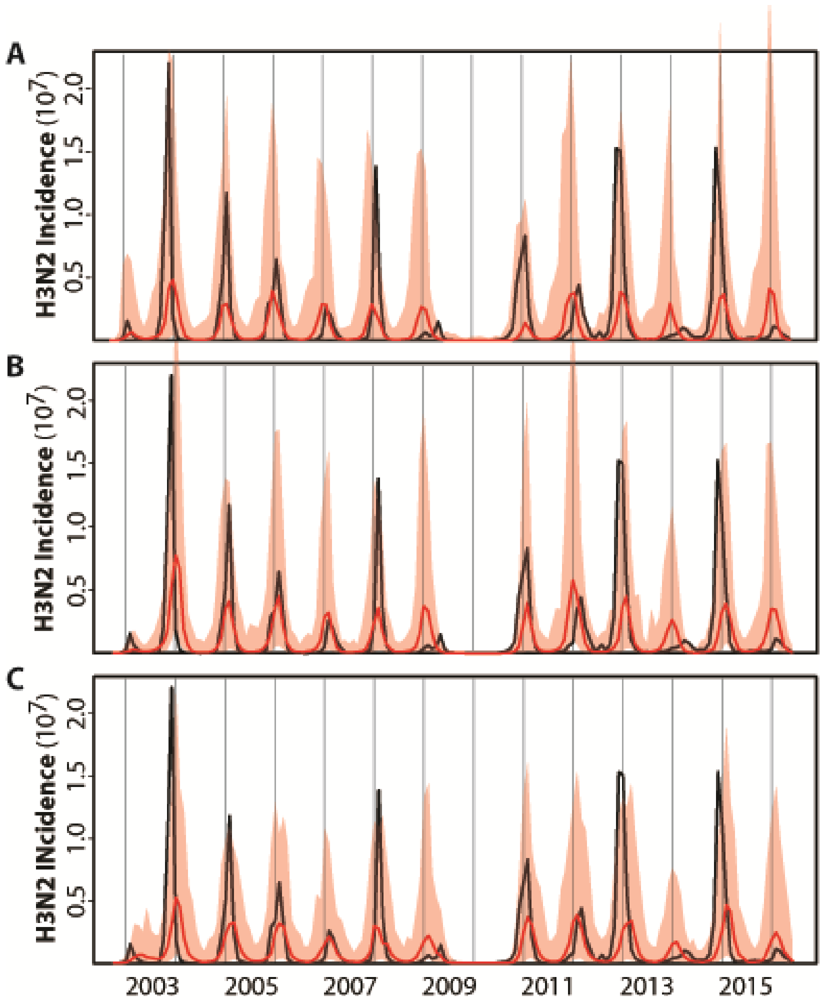


**Figure S1**. Illustration of the best model fits for **(A)** the basic H1, **(B)** the refined and **(C)** the humidity models. See Table 1 for the specification and statistical comparison of the different models. Here, simulations of the respective models with the MLE (Maximum Likelihood Estimates) parameters are shown for the median (in red) and 2.5-97.5% quantiles (shaded red) of 1000 simulations starting from estimated initial conditions in October 2002. For comparison, the data are shown in black. The basic-H1 model includes an additional class to rely on the observed incidence of H1N1 as a covariate. The refined model additionally considers a different reporting error for the summer and winter seasons, and allows for the sub-exponential growth of the epidemic curve. The humidity model uses a humidity-based function for the seasonal forcing. All three models only partially capture the interannual variation of H3N2 incidence.

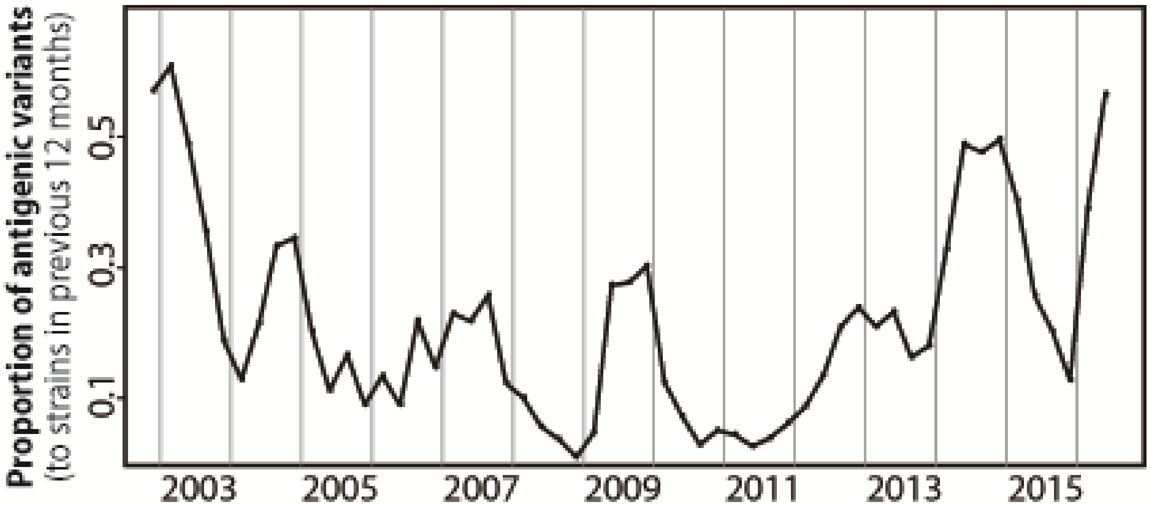


**Figure S2**. Proportion of antigenic variants (PAV). Antigenic variants were identified using the way described in the Materials and Methods. For strains in a given quarter, the number of antigenic variants was identified relative to the strains in the time window given by the previous 12 months. This number was then normalized by the number of all possible pairs. See the Materials and Methods for more details.

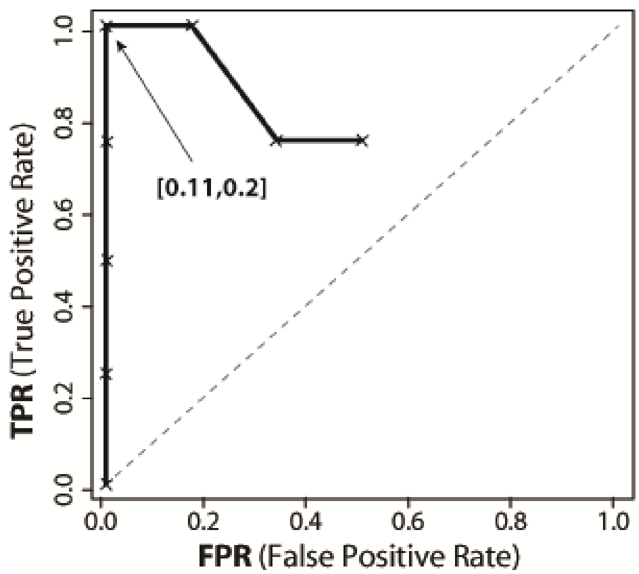


**Figure S3**. ROC curves for predicting antigenic cluster transitions based on the training dataset covering the period from 2002 to 2011 in US. Observed antigenic cluster transitions were identified based on CDC reports (see Materials and Methods for more details). Predictions were made based on the pattern of PAV change as explained in the Materials and Methods. When PAV for the third quarter crosses a cutoff, an antigenic cluster transition is anticipated for the next influenza season if PAV increases from the first quarter to the second quarter (or decreases but no antigenic cluster transition was assigned for the current influenza season; see Materials and Methods for more details). The range of the PAV cutoff with the best accuracy when predicting an antigenic cluster transition is also shown.

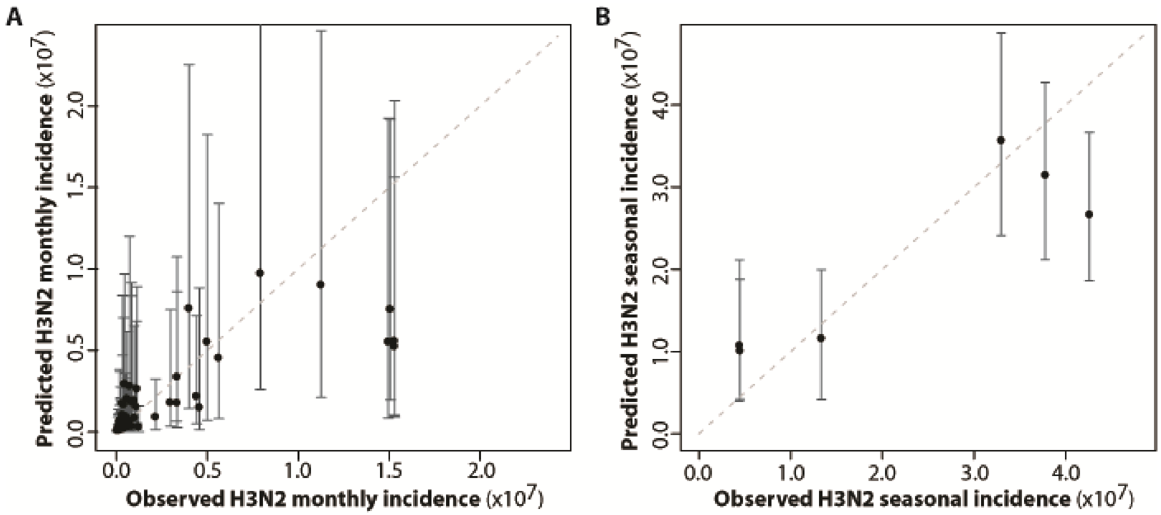


**Figure S4**. Comparison of observations and forecasts generated with the cluster model for the out-of-fit period covering the period from 2011 to 2017. **(A)** Scatter diagram for monthly observed and predicted medians of seasonal H3N2 incidence (with 2.5-97.5% quantiles) from 1000 random simulations with the best models ($r = 0.81$ and $r^2 = 0.67$). **(B)** Scatter diagram for seasonal observed and predicted medians of seasonal H3N2 incidence (with 2.5-97.5% quantiles) from 1000 random simulations with the best models ($r = 0.90$ and $r^2 = 0.81$). 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 available when this study was conducted, are based on data from the weekly US influenza surveillance report until week 14 ending on April 8, 2017. All correlations are statistically significant (p value < 0.05).

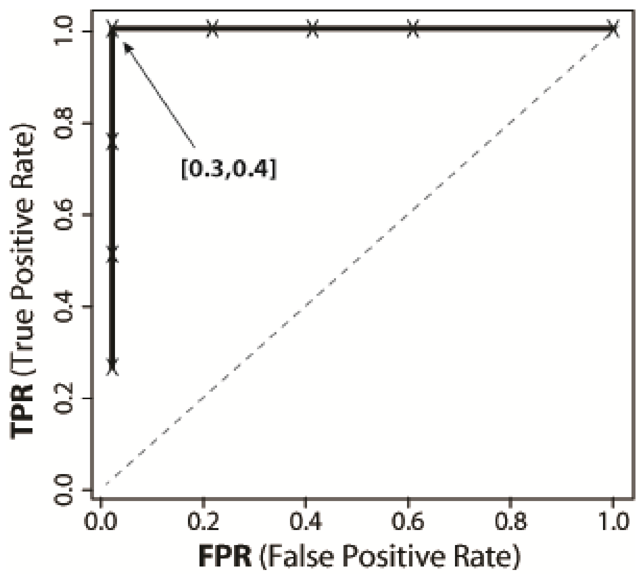


**Figure S5**. ROC curves for choosing the percentage cutoff applied to risk level prediction for the cluster model (Table 2 and Table S1). This choice is based only on the training dataset covering the period from 2002 to 2011 for the US. An observed high risk season is defined as a season whose total incidence is above a chosen threshold, here the 50% quantile of the seasonal totals from the whole training dataset. A high risk level is predicted for a season if the percentage of simulations that are classified as high risk among 1000 simulations is above a cutoff; otherwise, a low risk level is predicted for that season. The range of the percentage cutoff with the best accuracy when predicting seasonal risk level is also indicated.

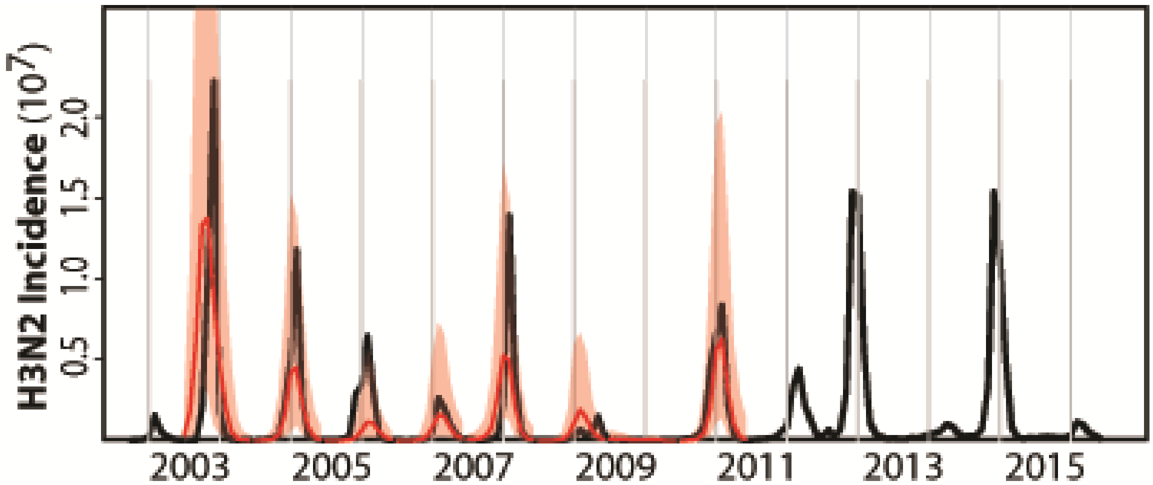


**Figure S6**. Leave-one-out cross validation for the cluster model based on the training dataset for the period between 2003 and 2011. These predictions are simulated on a seasonal basis from estimated initial conditions starting in June, and based on parameters estimated with all the data in the training dataset except the target season we seek to predict. We did so by not allowing the parameters whose values are being searched to vary in the target season during optimization, so that effectively the data in that season is not used in the parameter search and to evaluate the likelihood. Other methodological steps are those described in the Materials and Methods under Forecasts. The black curve is the monthly observed H3N2 incidence; the red curve is the predicted monthly medians with shaded 2.5-97.5% quantiles from 1000 random simulations with the best model. There is a significant correlation between observed incidence and median incidence from the forecasts for both monthly data ($r = 0.66$ and $r^2 = 0.44$) seasonal data ($r = 0.88$ and $r^2 = 0.77$). All correlations are statistically significant (p value < 0.05).

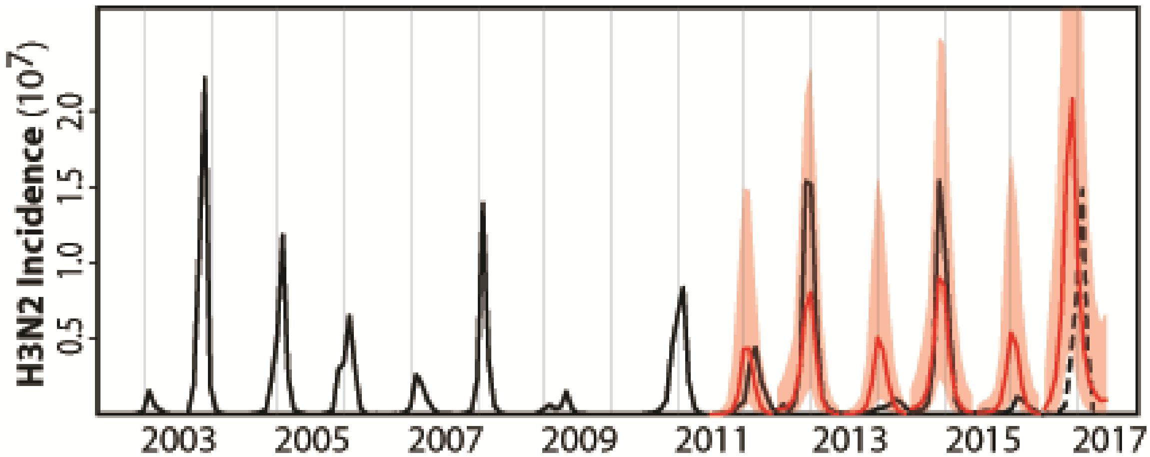


**Figure S7**. Forecasts based on the continuous model for the US. Both retrospective forecasts (for each influenza season from 2011/2012 to 2015/2016) and a real-time forecast for the 2016/2017 influenza season are presented. These forecasts are simulated on a seasonal basis from estimated initial conditions starting in June and based on parameters estimated with all the data up to that point in time. The black curve is the monthly observed H3N2 incidence; the red curve is the predicted monthly median incidence with shaded 2.5-97.5% quantiles from 1000 random simulations with the best model. The observed incidence data for the 2016/2017 influenza season, which were not available when this study was conducted, are shown with the dotted line, and are based on data from the weekly US influenza surveillance report until week 14 ending on April 8, 2017. Correlations between observed incidence and median incidence from the forecasts are $r = 0.51$ and $r^2 = 0.26$ for the monthly data, $r = 0.55$ and $r^2 = 0.30$ for the seasonal data. All correlations are statistically significant (with p value < 0.05).

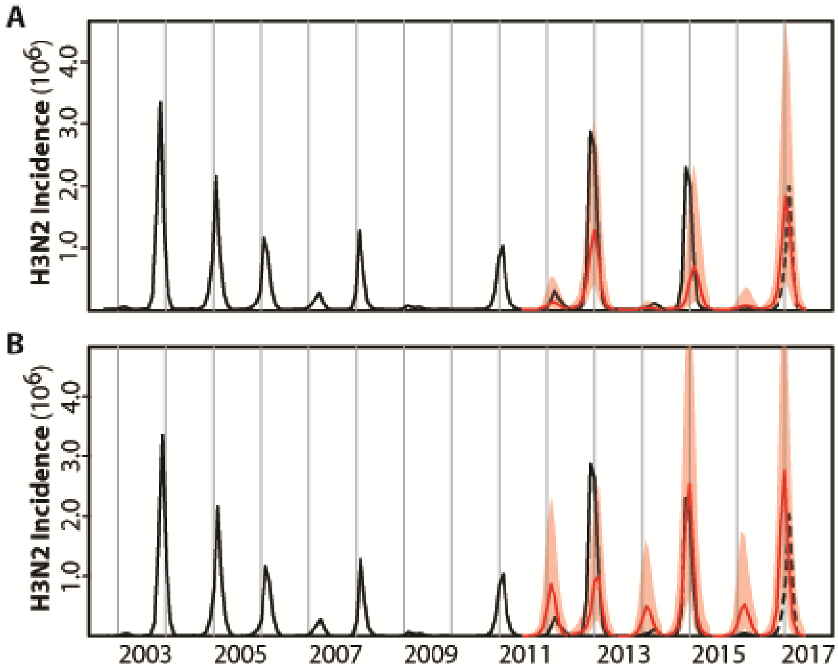


**Figure S8**. H3N2 incidence forecasts for the HHS region 3 based on **(A)** the cluster model and **(B)** the continuous model. Both retrospective forecasts (for each influenza season from 2011/2012 to 2015/2016) and a real-time forecast for the 2016/2017 influenza season are presented. These forecasts are simulated on a seasonal basis from estimated initial conditions starting in June, and are based on parameters estimated with all the data up to that point in time. The black curve is the monthly observed H3N2 incidence; the red curve is the predicted monthly median incidence with shaded 2.5-97.5% quantiles from 1000 random simulations with the best model. The observed incidence data for the 2016/2017 influenza season, which were not available when this study was conducted, are shown with the dotted line, and are based on data from the weekly US influenza surveillance report until week 14 ending on April 8, 2017. For the cluster model, correlations between observed incidence and median incidence from forecasts are $r = 0.68$ and $r^2 = 0.47$ for the monthly data, $r = 0.76$ and $r^2 = 0.57$ for the seasonal data (all correlations are significant, with p value < 0.05). For the continuous model, correlation between observed incidence and median incidence from forecasts are $r = 0.62$ and $r^2 = 0.38$ for the monthly data, $r = 0.60$ and $r^2 = 0.36$ for the seasonal data. All correlations are statistically significant, with p value < 0.05).

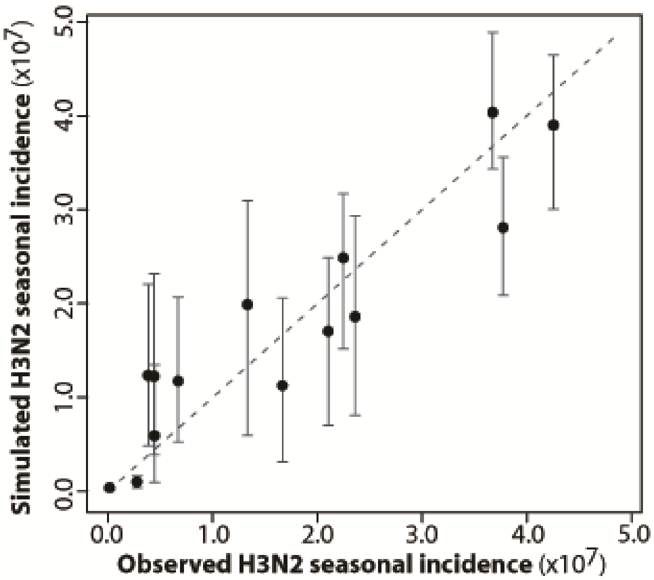

**Figure S9**. Scatter diagram for seasonal simulations and observations for the US dataset covering the period from 2002 to 2016 based on the cluster model. The correlation between seasonal observed H3N2 incidence and median incidence from the simulations is $r = 0.83$ and $r^2 = 0.69$, which is statistically significant (p value < 0.05). Confidence intervals are also shown based on 2.5-97.5% quantiles from 1000 random simulations with the best model. The seasonal incidence is the sum of the monthly incidence for a specific influenza season.

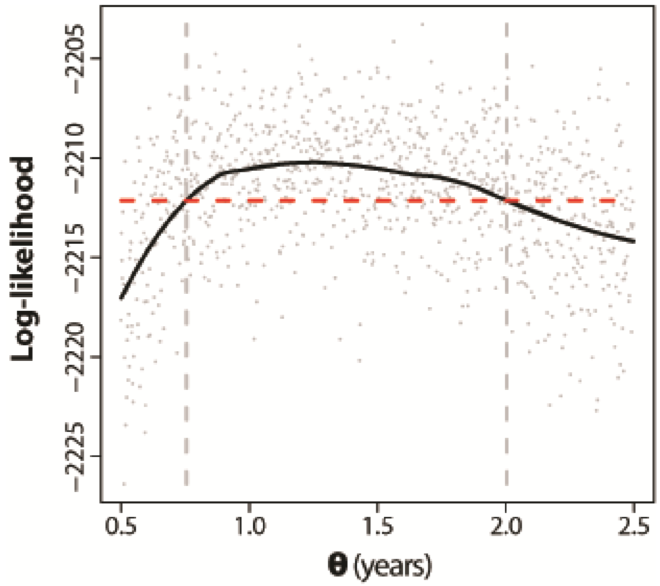


**Figure S10**. Log-likelihood profiling for average effective time $\theta$. The best likelihood among 10 random repeats for each $\theta$ fixed at a given value was obtained by allowing other parameters to be freely optimized (grey dots). The black curve is the fitting based on local polynomial regression. The intersection between the red dashed line and the vertical dashed lines indicates the 95% confidence interval.

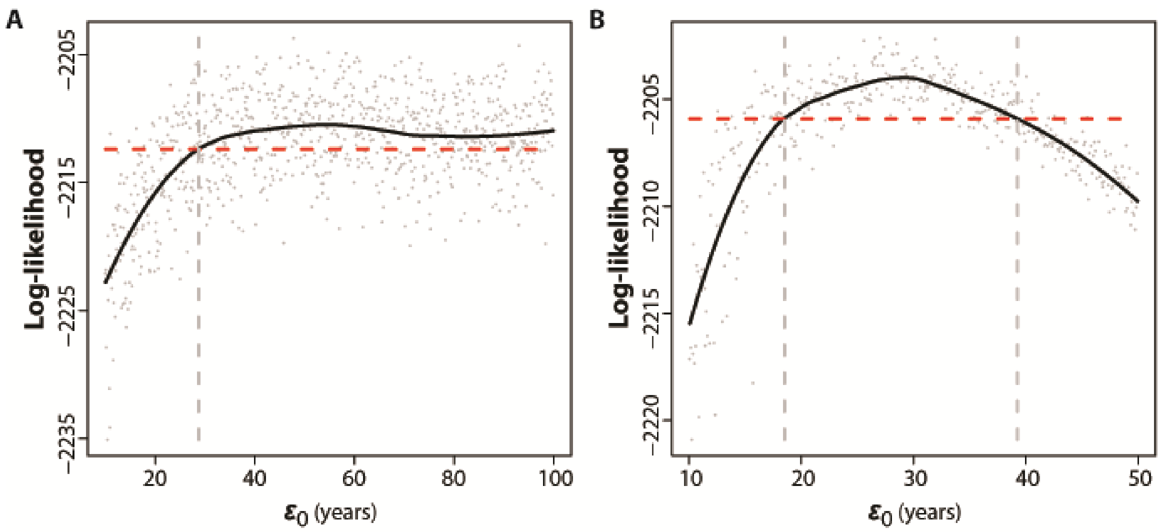


**Figure S11**. Log-likelihood profiling of the basic average latent time $\varepsilon_0$ for both **(A)** the best fitness model and **(B)** the cluster model. The intersection between the red dashed line and the vertical dashed lines indicates the 95% confidence interval.

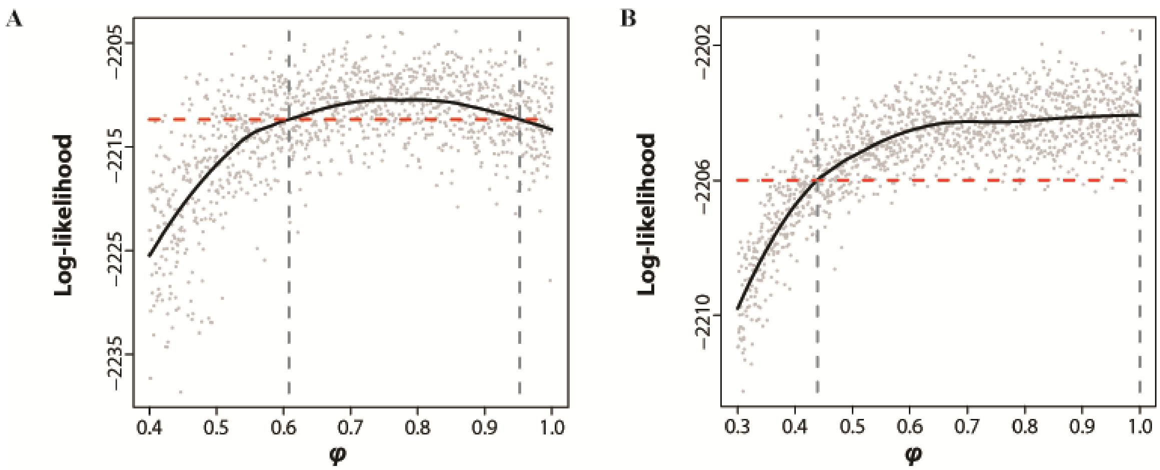


**Figure S12**. Log-likelihood profiling of the reporting rate $\varphi$ for both **(A)** the continuous model and **(B)** the cluster model. The intersection between the red dashed line and the vertical dashed lines indicates the 95% confidence interval.

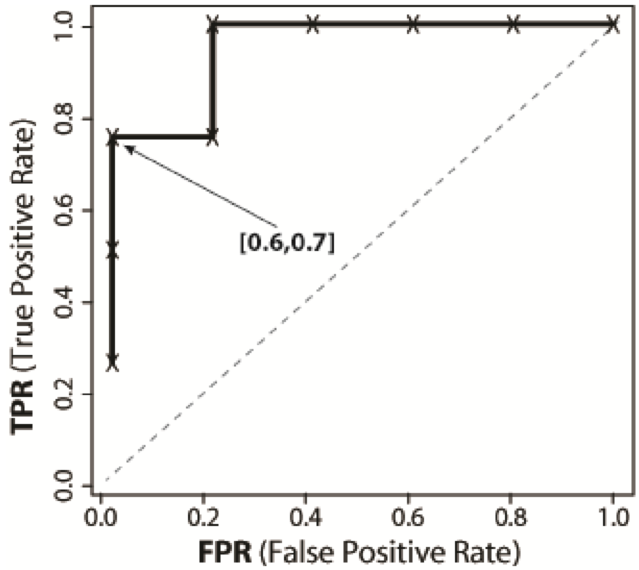


**Figure S13**. ROC curves for choosing the percentage cutoff applied to risk level prediction for the continuous model (Table S3). This choice is based on the training dataset covering the period from 2002 to 2011 in US. An observed high risk season is defined as a season with total incidence above a threshold, defined here as the 50% quantile of the seasonal totals from the whole training dataset. A high risk level is predicted for a given season when the percentage of simulations classified as high risk among 1000 simulations is larger than a cutoff; otherwise a low risk level is predicted for that season. The range of the percentage cutoffs with the best accuracy when predicting seasonal risk level is also indicated.

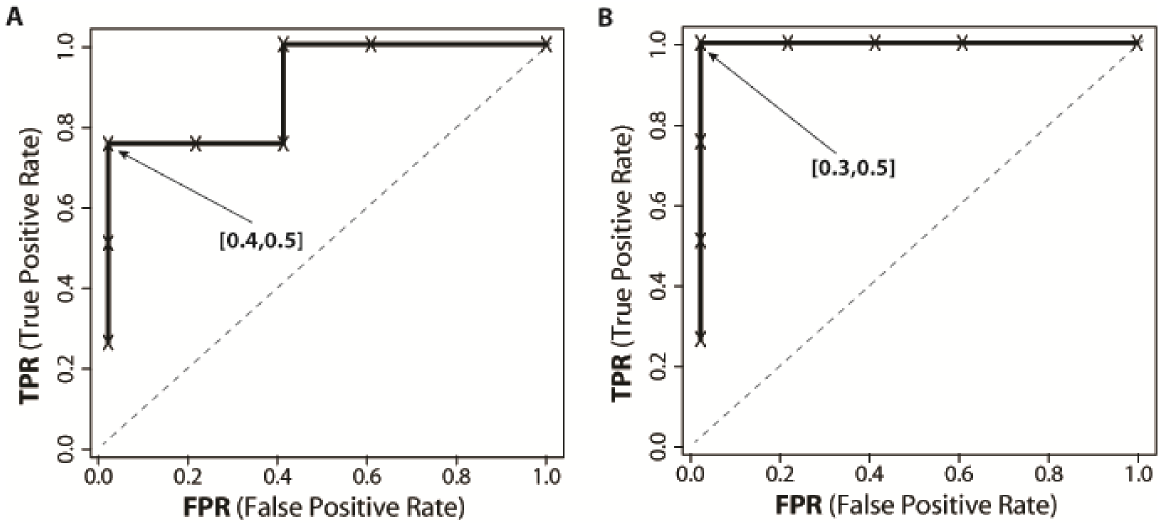


**Figure S14**. ROC curves for choosing the percentage cutoff applied to risk level prediction for **(A)** the cluster model and **(B)** the continuous model for the HHS region 3 (Table S5). This choice is based on the training dataset that covers the period from 2002 to 2011. An observed high risk season is a season whose total incidence surpasses a chosen threshold, here the 50% quantile of the seasonal totals from the whole training dataset. A high risk level is predicted for a season if the percentage of simulations classified as high risk among 1000 simulations is above a given cutoff; otherwise, a low risk level is predicted for that season. The ranges of the percentage cutoffs with the best accuracy when predicting seasonal risk level are also indicated.