**Forecast Influenza-related Hospitalizations during the 2017–2018 Influenza Season Collaborative Challenge**

**Objectives:**

Influenza-related hospitalizations are a major contributor to the overall burden of influenza in the United States. Accurate prediction of hospitalization rates could help ensure an appropriate public health response during an influenza season. To work towards this goal, CDC will coordinate a collaborative forecasting challenge for influenza-related hospitalizations during the 2017-2018 influenza season. For each week during the season, participants will be asked to provide national probabilistic forecasts for seasonal and short-term targets. The seasonal targets include the peak week and peak weekly rate of hospitalizations, while the short-term targets include weekly hospitalization rates one, two, three, and four weeks ahead from the date of forecast. Forecasts will be compared with values from the Influenza Hospitalization Surveillance Network (FluSurv-NET: <https://www.cdc.gov/flu/weekly/#S6>).

**Eligibility:**

All are welcome to participate in this collaborative challenge, including individuals or teams that have not participated in previous CDC forecasting challenges.

**Dates:**

The Challenge Submission Period will begin December 11, 2017, and will run until May 7, 2018. Weekly forecasts must be submitted by 11:59PM Eastern each Monday. Missed or late submissions will not preclude participation in this challenge but will adversely affect submission scores.

**Forecasting Targets:**

Forecasts should provide probabilistic forecasts for each of the two seasonal targets and four week-ahead targets. The probabilities for each prediction for each target should be positive and sum to 1. If the sum is greater than 0.9 and less than 1.1, the probabilities will be normalized to 1.0. If any probability is negative or the sum is outside of that range, the forecast will be discarded. Forecasts for the hospitalization rate for the four weeks following the forecast submission should be relative to the most recent week of FluSurv-NET data released. For example, FluSurv-NET data for week 48 will be posted on Friday, December 8 at 11:00AM Eastern Time. The four-week forecast submitted on Monday, December 11 should include predictions for FluSurv-NET values for weeks 49-52.

Forecasts should be provided at the national level for the overall hospitalization rate and each FluSurv-NET age group specific hospitalization rate. Initial submissions should include a brief narrative describing the methodology and data used in the prediction model. Model methodology and source data can be changed during the course of the challenge, but an updated narrative explanation of the model should be provided if models are changed.

*Target definitions*

* The peak week will be defined as the MMWR surveillance week that the overall FluSurv-NET hospitalization rate is the highest for the 2017-2018 influenza season.
* The peak weekly rate will be defined as the highest numeric value that the overall FluSurv-NET hospitalization rate reaches during the 2017-2018 influenza season.
* One- to four-week ahead forecasts will be defined as the weekly overall FluSurv-NET hospitalization rate for the target week.

FluSurv-NET rates will be rounded to one decimal point for determining the peak week, peak weekly rate, and weekly forecast targets. In the case of multiple peak weeks (i.e. there is an identical peak FluSurv-NET value in two or more weeks), both weeks will be considered the peak week.

**Forecast Submission:**

Forecasts should be submitted via online submission to the Epidemic Prediction Initiative website at <https://predict.phiresearchlab.org>. Forecasts should be structured using the provided submission template (named “FluSurv\_Submission\_Template.csv”). **The structure of the spreadsheet (e.g. the column or row locations) should not be modified in any way to ensure accurate forecast upload.** Forecasts for peak weekly rate and for 4-weeks-ahead should be given in the provided 0.1 case per 100,000 intervals labeled as “bin\_start\_incl” on the submission sheet (e.g. the bin for 3.1 represents probability that the rounded FluSurv-NET rate = 3.1). The probability assigned to the final bin (labeled 13 for all age groups except 65+, where it is labeled 60) includes the probability of all FluSurv-NET rates greater than or equal to that value.

To conform to the current structure of the Epidemic Prediction Initiative website’s submission system, some labels on the submission template are not 100% aligned with the information being presented. First, the submission template column labeled “location” actually refers to the age group being forecast. Second, while FluSurv-NET data are measured in rates per 100,000 population, weekly targets and peak intensity targets are labeled as “percent.” Finally, the peak weekly rate target is labeled as “Season peak percentage.” These slight inaccuracies are necessary to allow the submitted forecasts to be visualized correctly on the EPI website. Please note the relevant areas on the template and ensure submitted forecasts are properly labeled.

In the event that online submission is unsuccessful for a particular week due to technical issues, submissions should be emailed to flucontest@cdc.gov. In this case, the filename should be modified to the following standard naming convention: a forecast submission using week 45 surveillance data submitted by John Doe University on December 11, 2017, should be named “EW48-JDU-Hosp-2017-12-11.csv” where EW48 is the latest week of FluSurv-NET data used in the forecast, JDU is the name of the team making the submission (e.g. John Doe University), and 2017-12-11 is the date of submission.

**Evaluation Criteria:**

*Probabilistic forecasts*

All forecasts will be evaluated using the hospitalization rates pulled from the FluSurv-NET system during MMWR week 28 of 2018. The logarithmic scoring rule will be used to measure the accuracy of the probability distribution of a forecast. If is the set of probabilities for a given forecast, and   is the probability assigned to the observed outcome , the logarithmic score is:
For peak week, the probability assigned to that correct week plus the probability assigned to the preceding and following weeks will be summed to determine the probability assigned to the 3-week window centered on the observed outcome. In the case of multiple peak weeks, the probability assigned to each peak week and the corresponding preceding and following bins will be summed.

For peak rate and 1-4 week-ahead forecasts, the probability assigned to the correct 0.1 bin plus the probability assigned to the age group-specific number of preceding and following bins will be summed to determine the probability assigned to the observed outcome, with a minimum of 1 preceding and following bin. For each age group, and for overall rates, bins including up to plus or minus 10% of the observed value will be applied to each side of the observed value, rounded to the nearest 0.1. For example, if the observed overall peak hospitalization rate is 3.3 per 100,000, the bins encompassing 10% of this value (0.3) above and below the observed value of 3.3 will be included. Therefore, the probabilities assigned to all bins ranging from 3.0 to 3.6 would be summed to determine the probability assigned to the observed outcome.

In the case of very small observed values, a minimum of one bin proceeding and following the observed bin will be included with the observed bin. For example, if the observed weekly hospitalization rate is 0.2 per 100,000, the probabilities assigned to bins representing 0.1 and 0.3 will also be included. For all targets, if the correct bin is near the first or last bin, the number of bins will be truncated at the respective boundary.

 For example, if the observed overall hospitalization rate for a given week is 0.0 per 100,000, probabilities assigned to bins ranging from 0.0 to 0.1 will be summed. If the observed rate is in the highest bin (i.e. an overall rate ≥ 13.0), only those bins with values contained in the calculated window based on the value of the highest bin will be summed. For example, if the correct maximum overall hospitalization rate is 13.2, the window will be plus or minus 1.3 and probabilities assigned to bins ranging from 11.7 upwards will be summed.

Undefined natural logs (which occur when the probability assigned to the observed outcome was 0) will be assigned a value of -10. Forecasts which are not submitted (e.g. if a week is missed) or that are incomplete (e.g. sum of probabilities less than 0.9 or greater than 1.1) will also be assigned a value of -10. Logarithmic scores will be averaged across different submission time periods, the seasonal targets, and the four-week ahead targets to provide both specific and generalized measures of model accuracy.

**Example:** At the conclusion of the season, FluSurv-NET showed that the 2016/2017 overall weekly hospitalization rates peaked at 5.4 per 100,000. As a result, the window of plus or minus 0.54 rounds to 0.5, which spans the probability bins from 4.9 to 5.9. If a forecast predicts there is a probability of 0.1 (i.e. a 10% chance) that hospitalization rates peak at 5.4 per 100,000, with an additional 0.3 probability that they peak between 4.9 and 5.3 and a 0.2 probability that they peak between 5.5 and 5.9, then the forecast would receive a score of log(0.6) = -0.51.

Forecast accuracy will be measured by log score only. Nonetheless, forecasters are requested to continue to submit point predictions to aid in communication and dissemination of forecasts, which should aim to minimize the Absolute Error (AE). Absolute error (AE) is the absolute difference between a prediction   and an observation : . If no point prediction is provided, CDC will calculate the point prediction as the median value from the submitted distribution.

**Data**

The historical national surveillance data that can be used to enable training and model development are available at <https://gis.cdc.gov/GRASP/Fluview/FluHospRates.html>; these data are updated every Friday at noon Eastern Time. Teams will be responsible for downloading weekly updates to FluSurv-NET surveillance data during the challenge. The ‘cdcfluview’ R package can also be used to access and download FluSurv-NET surveillance data. Historical FluSurv-NET information dating to the 2010-2011 season with weekly updates from the 2013-2014 season onwards are available to participating teams on the CDC Epidemic Prediction Initiative website (https://predict.phiresearchlab.org). Teams are welcome to utilize additional data beyond FluSurv-NET - additional potential data sources include but are not limited to:

Carnegie Mellon University’s [Delphi group](http://delphi.midas.cs.cmu.edu/)’s [Epidata API](https://github.com/undefx/delphi-epidata)

Health Tweets: <http://www.healthtweets.org/>

**Publication of forecasts:**

All participants provide consent that their forecasts can be published in real-time on the CDC’s Epidemic Prediction Initiative website and, after the season ends, in a scientific journal describing the results of the challenge. The forecasts can be attributed to a team name (e.g. John Doe University) or anonymous (e.g. Team A) based on the individual team’s preference. Team names should be limited to 25 characters for display online. Additionally, teams are requested to inform CDC if they prefer that their probabilistic forecast data be published with their team name attached or published anonymously. No participating team can publish the results of another team’s model in any form without the team’s consent. The manuscript describing the accuracy of forecasts across teams will be coordinated by a representative from CDC. If discussing the forecasting challenge on social media, teams are encouraged to use the hashtag #CDCflusight to promote visibility of the challenge.

**Ensemble Model and Historical Model:**

 Participant forecasts will be combined into an unweighted average ensemble forecast to be published in real-time along with the participant forecasts. In addition, forecasts will be displayed alongside the output of a null model for comparison, based solely on the historical distribution of the value of interest (i.e. peak week, peak weekly rate, or hospitalization rate in a given MMWR week), excluding the 2009/2010 H1N1 pandemic season.