**Forecast the 2016–2017 Influenza Season Collaborative Challenge**

**Objectives:**

To improve influenza forecasting, we will undertake a collaborative comparison of forecasts for the 2016-2017 influenza season. For each week during the season, participants will be asked to provide national and regional probabilistic forecasts for the entire influenza season (seasonal targets) and for the next four weeks (four-week ahead targets). The seasonal targets are the onset week, the peak week, and the peak intensity of the 2016-2017 influenza season. The four-week ahead targets are the percent of outpatient visits experiencing influenza-like illness (ILI) one week, two weeks, three weeks, and four weeks ahead from date of the forecast. All forecasts will be compared to the weighted values from the U.S. Outpatient Influenza-like Illness Surveillance Network (the ILINet system: <http://www.cdc.gov/flu/weekly/overview.htm>). Participants can submit forecasts for the seasonal targets, the four-week ahead targets, or both.

**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 November 7, 2016, and will run until May 15, 2017. 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 (i.e. 50% peak will occur on week 2; 30% chance on week 3) as well as the point prediction for each of the three 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 ILINet percentage for the four weeks following the forecast submission should be relative to the most recent week of ILINet data released. For example, ILINet data for week 43 will be posted on Friday, November 4 at 12:00PM Eastern Time. The four-week forecast submitted on Monday, November 7 should include predictions for ILINet values for weeks 44-47.

Forecasts must be provided at the national level. Forecasts at the HHS region level are also encouraged. 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 onset of the season is defined as the MMWR surveillance week (<http://wwwn.cdc.gov/nndss/script/downloads.aspx>) when the percentage of visits for influenza-like illness (ILI) reported through ILINet reaches or exceeds the baseline value for three consecutive weeks (updated 2016-2017 ILINet baseline values for the US and each HHS region will be available at <http://www.cdc.gov/flu/weekly/overview.htm> the week of October 10, 2016). Forecasted “onset” week values should be for the first week of that three week period.
* The peak week will be defined as the MMWR surveillance week that the weighted ILINet percentage is the highest for the 2016-2017 influenza season.
* The intensity will be defined as the highest numeric value that the weighted ILINet percentage reaches during the 2016-2017 influenza season.
* One- to four-week ahead forecasts will be defined as the weighted ILINet percentage for the target week.

ILINet values will be rounded to one decimal point for determining the onset week, the peak week, peak ILINet percentage, and weekly forecast targets. In the case of multiple peak weeks (i.e. there is an identical peak ILINet value in two or more weeks within a geographic region), both weeks will be considered the peak week.

**Forecast Submission:**

Forecasts should be submitted to [flucontest@cdc.gov](mailto:flucontest@cdc.gov) using the provided .csv spreadsheet (named “Weekly\_Submisson\_Spreadsheet”). **The structure of the spreadsheet (e.g. the column or row locations) should not be modified in any way.** For onset, the “none” field in the spreadsheet is to indicate if no influenza season is forecasted (e.g. the ILINet value never reaches or exceeds the baseline for at least three consecutive weeks during the season). This value should be used to indicate no season for the season onset. Forecasts for peak percent and for 4-weeks-ahead should be given in the provided 0.1 percentage intervals labeled as “bin\_start\_incl” on the submission sheet (e.g. the bin for 3.1% represents probability that 3.05%<= ILINet peak <3.15%). The probability assigned to the final bin labeled 13% includes the probability of ILINet values greater than or equal to 13%.

For submission, the filename should be modified to the following standard naming convention: a forecast submission using week 43 surveillance data submitted by John Doe University on November 7, 2016, should be named “EW43-JDU-2016-11-07.xls” where EW43 is the latest week of ILINet data used in the forecast, JDU is the name of the team making the submission (e.g. John Doe University), and 2016-11-07 is the date of submission.

At some point during the season, teams may be able to submit their forecasts directly to the CDC’s Epidemic Prediction Initiative website. More guidance will be provided at that time for how to submit forecasts in that manner.

**Evaluation Criteria:**

*Probabilistic forecasts*

All forecasts will be evaluated using the weighted observations pulled from the ILINet system during week 28, and 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 onset week and peak week, the probability assigned to that correct bin (based on the weighted ILINet value) plus the probability assigned to the preceding and proceeding bins will be summed to determine the probability assigned to the observed outcome. If onset is never reached during the season, only the probability assigned to the bin for “no onset” will be scored. In the case of multiple peak weeks, the probability assigned to the bins containing the peak weeks and the preceding and proceeding bins will be summed. For peak percentage and 4-weeks-ahead forecasts, the probability assigned to the correct bin plus the probability assigned to the five preceding and five proceeding bins will be summed to determine the probability assigned to the observed outcome. For example, if the correct peak ILINet value is 6.5%, the probabilities assigned to all bins ranging from 6.0% to 7.0% will be summed to determine the probability assigned to the observed outcome.

For all targets, if the correct bin is near the first or last bin, the number of bins summed will be reduced accordingly. No bin farther than one bin (onset and peak week) or five bins away (percentage forecasts) from the correct bin will contribute to the score. For example, if the correct ILINet percentage for a given week is 0.3%, probabilities assigned to bins ranging from 0% to 0.8% 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 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, the four-week ahead targets, and locations to provide both specific and generalized measures of model accuracy.

**Example:** A forecast predicts there is a probability of 0.2 (i.e. a 20% chance) that the flu season starts on week 44, a 0.3 probability that it starts on week 45, and a 0.1 probability that it starts on week 46 with the other 0.4 (40%) distributed across other weeks according to the forecast. Once the flu season has started, the prediction can be evaluated, and the ILINet data show that the flu season started on week 45. The probabilities for week 44, 45, and 46 would be summed, and the forecast would receive a score of log(0.6) = -0.51. If the season started on another week, the score would be calculated on the probability assigned to that week plus the values assigned to the preceding and proceeding week.

Forecast accuracy will be measured by log score only. Nonetheless, forecasters are requested to continue to submit point predictions, which should aim to minimize the Absolute Error (AE). Absolute error (AE) is the absolute difference between a prediction   and an observation :     
For example,a forecast predicts that the flu season will start on week 45; flu season actually begins on week 46. The absolute error of the prediction is |45-46| = 1 week. For season onset, if the point prediction is for no onset, please report a point prediction of “NA”.

**Data**

The historical national surveillance data that could be used to enable training and model development are available at <http://gis.cdc.gov/grasp/fluview/fluportaldashboard.html>; these data are updated every Friday at noon Eastern Time. The “cdcfluview” package for R can be used to retrieve these data automatically. Teams are welcome to utilize additional data beyond ILINet - 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 their probabilistic forecast data can be published with their team name attached, published anonymously, or if they prefer not to share their forecast data. 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 Null Models:**

New this year, participant forecasts will be combined into an ensemble forecast to be published in real-time along with the participant forecasts. All teams are welcome to contribute to the development of this ensemble model and interested teams should contact CDC. In addition, forecasts will be displayed alongside the output of two null models for comparison. One null model will be based solely on the historical distribution of the value of interest (i.e. onset week, peak week, peak percentage, or wILI percentage in a given MMWR week), excluding the 2009/2010 H1N1 pandemic season, while the second null model will be a simple SARIMA model fit to prior years’ ILINet activity, excluding the 2008/2009 and 2009/2010 years to eliminate activity due to the H1N1 pandemic.

**Hospitalization Rates Working Group**

Based on feedback at the annual forecasting meeting, CDC is exploring adding weekly rates of laboratory-confirmed influenza hospitalizations as a target to be modeled in future years. Currently, CDC’s FluSurv-NET system covers hospitals in 13 states and estimates age-specific rates of laboratory-confirmed influenza hospitalization. Initial targets would be peak national hospitalization rate, peak week of hospitalization, and 4-week ahead forecasts, though teams are invited to provide feedback on target definitions. Forecasts would be probabilistic in nature, similar to those for ILINet percentages. Interested teams should contact CDC to form a working group to explore the feasibility of these targets. Historical surveillance data of influenza hospitalization rates from FluSurv-Net are available at http://gis.cdc.gov/GRASP/Fluview/FluHospRates.html

**State-based ILINet Working Group**

Teams interested in being part of a working group to discuss opportunities and challenges regarding potentially including state-based ILINet forecasts in future competitions should contact CDC.