**Forecast the 2015–2016 Influenza Season Collaborative Challenge**

**Objectives:**

To improve influenza forecasting, we will undertake a collaborative comparison of forecasts for the 2015-16 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 beginning week, the peak week, and the peak intensity of the 2015–2016 influenza season. The four-week ahead targets are the percent of outpatient visits experiencing 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 2, 2015, and will run until May 16, 2016. 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 Milestones:**

Participants can provide forecasts for three seasonal targets during the 2015-2016 influenza season: the beginning, the peak, and the intensity. Forecasts should provide probabilistic forecasts (i.e. 50% peak will occur on week 2; 30% chance on week 2) as well as the point prediction for each of the three milestones. The probabilities for each prediction for each milestone should be positive and sum to 1. If the sum is greater than 0.9 or 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. Participants can also provide forecasts for the ILINet percentage for the four weeks following the forecast submission. For example, ILINet data for week 41 will be posted on Friday, October 23 at 12:00PM Eastern Time. The four-week forecast submitted on Monday, October 26 should include predictions for ILINet values for weeks 42–45.

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 start of the season is defined as the MMWR surveillance week (<http://wwwn.cdc.gov/nndss/script/downloads.aspx>) when the percentage of visits reported through ILINet reaches or exceeds the baseline value for three consecutive weeks (updated 2015–16 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 11, 2015). Forecasted “start of the season” 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 2015–16 influenza season.
* The intensity will be defined as the highest numeric value that the weighted ILINet percentage reaches during the 2015-16 influenza season.

ILINet values will be rounded to one decimal point for determining the start week, the peak week, and peak ILINet percentage. 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 locations or row spacing) should not be modified in any way.** For submission, the filename should be modified to the following standard naming convention: a forecast submission using week 42 surveillance data submitted by John Doe University on October 26, 2015, should be named “EW42-JDU-2015-10-26.xls” where EW 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 2015-10-26 is the date of submission. The Pr(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 provided in 0.5 percentage point semi-open intervals (e.g. probability 0.6 that 3%<= ILINet peak <3.5%).

**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 alt is the set of probabilities for a given forecast, and alt is the probability assigned to the observed outcome, alt , the logarithmic score is:   
alt   
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 the correct bin is the first or last bin, the probabilities will be summed over the first three or last three bins, respectively. 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. Undefined natural logs (which occur when the probability assigned to the observed outcomes 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 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.

Unlike last year, 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 alt and an observation alt: alt   
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.

**Publication of forecasts:**

All participants provide consent that their forecasts can be published in real-time on a central 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). 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.