**Forecast 2018–2019 State Level Influenza Season: Elective Collaborative Challenge Year 2**

**Objectives:**

To improve influenza forecasting, we will undertake a collaborative challenge of state-level forecasts for the 2018-2019 influenza season. For each week during the season, participants will be asked to provide state and territorial level probabilistic forecasts for the entire influenza season (seasonal targets) and for the next four weeks ahead (short-term targets). The seasonal targets are the peak week and the peak intensity of the 2018-2019 influenza season. The short-term (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 state-specific 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 short-term 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 Period will begin October 29, 2018 and will run until May 13, 2019. Participants must submit weekly forecasts by 11:59PM Eastern Standard Time each Monday. Missed or late submissions will not preclude participation in this challenge but will adversely affect submission scores.

**Forecasting Targets:**

* Seasonal Targets
  + The peak week is defined as the MMWR surveillance week that the ILINet percentage is the highest for the 2018-2019 influenza season for each state/territory.
  + The peak intensity is defined as the highest numeric value that the ILINet percentage reaches during the 2018-2019 influenza season for each state/territory.
* Short-term Targets
  + One- to four-week ahead forecasts will be defined as the ILINet percentage for the target week.

ILINet values will be rounded to one decimal point for the purposes of determining all 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 state or territory), both weeks will be considered the peak week.

**Forecast Submission:**

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 two seasonal targets and four-week ahead targets. The probabilities for each target prediction should be non-negative 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 the 0.9-1.1 range, the forecast will be discarded. Short-term forecast submissions 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 2 at 12:00PM Eastern Standard Time. Each short-term forecast (1- , 2- , 3- , and 4-week ahead) submitted on Monday, November 5 should include predictions for ILINet values for weeks 44-47.

A description of methodology should be submitted to CDC by November 16 using the methodology form provided. This form captures key model factors, such as data source(s) and model type(s) in a standardized way. Model methodology and source data may be changed during the course of the challenge, but teams should submit a new methodology form as soon as possible after the change. Please submit the completed form and forward any questions to [flucontest@cdc.gov](mailto:flucontest@cdc.gov).

**Submission Structure:**

All forecasts should be structured to match the attached spreadsheet (named “StateILI\_Submission\_Template.csv”). **The structure of the spreadsheet (e.g., the column or row locations) should not be modified in any way.** Peak intensity and week-ahead forecasts should be given in the provided 0.1 percentage intervals labeled as “bin\_start\_incl” on the submission sheet. For example, the bin for 3.1% represents the probability that rounded ILINet equals 3.1%. The probability assigned to the final bin labeled 13% includes the probability of ILINet values greater than or equal to 13%.

Forecasts should be submitted online through the FluSight website (<https://predict.cdc.gov/>). Instructions will be provided once testing on the new website is complete. In the event forecasts cannot be submitted online, they may be emailed to [flucontest@cdc.gov](mailto:flucontest@cdc.gov) using the provided .csv spreadsheet. For an email submission, the file name should be modified to the following standard naming convention: a forecast submission using week 43 surveillance data submitted by John Doe University on November 6, 2018, should be named “EW43-JDU-StateILI-2018-11-05.csv” 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 2018-11-05 is the date of submission.

At some point during the season, teams may be able to submit their forecasts using an application programming interface (API). Additional guidance will be provided at that time.

**Evaluation Criteria:**

*Log Score*

All forecasts will be evaluated using the weighted observations pulled from the ILINet system during week 28 of 2019, 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 peak week, the probability assigned to that correct bin (based on the weighted ILINet value) plus the probability assigned to the immediately preceding and proceeding bins will be summed to determine the probability assigned to the observed outcome. In the case of multiple peak weeks, the probability assigned to the bins containing each peak week and the respective preceding and proceeding bins will be summed. For peak intensity and short-term 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 (peak week) or five bins (peak intensity and short-term, i.e. percentage forecasts) away 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 is 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.

**Example:** A forecast predicts, for a given state, there is a probability of 0.2 (i.e., a 20% chance) that peak week is on week 5, a 0.3 probability that peak is on week 6, and a 0.1 probability that peak is on week 7, 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 peak week for this given state was on week 6. The probabilities for week 5, 6, and 7 would be summed, and the forecast would receive a score of log(0.6) = -0.51. If the peak occured on another week, the score would be calculated on the probability assigned to that week plus the values assigned to the immediately preceding and proceeding weeks.

*Absolute Error*

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 such that: . If a point prediction is not provided, CDC will estimate the point prediction using the median of the submitted distribution.

**Example:** A forecast predicts that the peak week for a given state will occur on week 5; flu actually peaks in this state on week 6. The AE of the prediction is |5-6| = 1 [week].

**Method to Determine Overall Team Rankings**

Logarithmic scores for seasonal and short-term forecasts will be averaged across different submission time periods and locations to provide both specific and generalized measures of model accuracy. The overall team rankings at the end of the season will be determined by averaging scores across all of the state targets over the entire forecasting period. Teams that do not provide all seasonal and short-term targets for all states will be ineligible to be named the overall top performing team; however, they will still be ranked for the states and targets they provided. Although teams may choose to participate in more than one challenge (e.g. FluSight and the state challenge described here), rankings for one challenge will not influence rankings for another, and an overall top-score will not be determined.

**Data Sources**

Historical jurisdiction-specific ILI data are currently available for all jurisdictions (except Florida) directly through the FluView website or via the get\_flu\_data function in the ‘cdcfluview’ package in R. During the season, publicly available data for each jurisdiction will be updated each Friday on the FluView website. Teams are welcome to utilize additional data beyond ILINet - additional potential data sources include but are not limited to: Carnegie Mellon University’s Epidata API ([Delphi group](http://delphi.midas.cs.cmu.edu/) <<http://delphi.midas.cs.cmu.edu/>> and <https://github.com/undefx/delphi-epidata>) and 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 (<https://predict.cdc.gov/>), GitHub page (<https://github.com/cdcepi>), 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 individual team preference. Team names should be limited to 25 characters for display online. The team name registered with the EPI website will be displayed alongside a team’s forecasts – any team that wishes to remain anonymous should contact CDC to obtain an anonymous team name to use. No participating team may 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:**

Participant forecasts will be combined into one or more ensemble forecasts to be published in real-time along with the participant forecasts. All teams are welcome to contribute to the development of the ensemble or have their forecasts contribute to the ensemble, and interested teams should contact CDC at [flucontest@cdc.gov](mailto:cdccontest@cdc.gov). In addition, forecasts will be displayed alongside the output of one null model for comparison, which is based solely on the historical distribution of the value of interest (i.e., peak week, peak percentage, or ILI percentage in a given MMWR week), excluding the 2009/2010 H1N1 pandemic season.

**FluSight Challenge**

Teams interested in participating in the sixth year of the FluSight Forecast the Influenza Season Collaborative Challenge should contact CDC at [flucontest@cdc.gov](mailto:flucontest@cdc.gov). Historical national surveillance data from ILINet are available at <http://gis.cdc.gov/grasp/fluview/fluportaldashboard.html> and archived national and regional ILI baseline values are available at <https://github.com/cdcepi/FluSight-forecasts/blob/master/wILI_Baseline.csv>.

**Hospitalization Rates**

Teams interested in participating in the second year of the FluSurv-Net Hospitalization Rate Forecasting Challenge should contact CDC at [flucontest@cdc.gov](mailto:cdccontest@cdc.gov). Historical surveillance data of influenza hospitalization rates from FluSurv-Net are available at <http://gis.cdc.gov/GRASP/Fluview/FluHospRates.html>.