**Forecast 2017–2018 State Level Influenza Season Collaborative Challenge**

**Objectives:**

To improve influenza forecasting, we will undertake a collaborative comparison of forecasts for the 2017-2018 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 (four-week ahead targets). The seasonal targets are the peak week and the peak intensity of the 2017-2018 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 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 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 6, 2017, and will run until May 14, 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 (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 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 3 at 12:00PM Eastern Time. The four-week forecast submitted on Monday, November 6 should include predictions for ILINet values for weeks 44-47.

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 ILINet percentage is the highest for the 2017-2018 influenza season.
* The intensity will be defined as the highest numeric value that the ILINet percentage reaches during the 2017-2018 influenza season.
* 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 geographic region), both weeks will be considered the peak week.

**Forecast Submission:**

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.** 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 rounded ILINet = 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.phiresearchlab.org/>) using the instructions included in Appendix A. In the event forecasts cannot be submitted online, they may be emailed to flucontest@cdc.gov using the provided .csv spreadsheet. For an email 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 6, 2017, should be named “EW43-JDU-StateILI-2017-11-06.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 2017-11-06 is the date of submission.

At some point during the season, teams may be able to submit their forecasts using an API. More guidance will be provided at that time for how to submit forecasts in that manner.

**Evaluation Criteria:**

All forecasts will be evaluated using the weighted observations pulled from the ILINet system during week 28 of 2018, 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 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 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 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 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 : . If a point prediction is not provided, then the point prediction will be estimated by using the median of the submitted distribution.

**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.

**Method to determine overall team rankings**

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. 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. If teams opt to not provide forecasts for all states and all seasonal and short-term targets, then they will be ineligible to be named the overall top performing team but they will still be ranked for the states and targets they provided.

**Data**

Historical state-specific ILI data will be shared with interested teams prior to the beginning of the challenge. These data are not publically available, and teams are required to use them only for the purposes of model training and development. During the season, weekly updates for each state’s level of ILI will be distributed through the FluSight website each Friday at noon ET. Data will only be accessible after signing into the website.

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 (<https://predict.phiresearchlab.org/>) and 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 the individual team’s 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 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:**

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. 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.

**Appendix A: FluSight website forecast submission instructions**

1. If you have not yet created an account and activated the project, please see the instructions at the end of this document to do so.
2. Log into the website at <https://predict.phiresearchlab.org/>
3. Click on “Go to project” under your team name in the FluSight box



1. In the left menu bar, click on “Submit” to access the submission page



1. In the row for the week you are submitting (Nov 21 for the first submission), click “Upload CSV” and select your submission file.
2. You should see “Sending…” appear next to the Upload button while your data are transmitted and the system checks to make sure your submission is valid.
3. If the submission is successful, you will see an “Open” link next to the upload button. Clicking this will open the JSON file that your submission has been converted to that will eventually be used for visualizations on the webpage.
4. If your submission is not successful, an informative error message will popup explaining what rows/columns are out of agreement with a valid submission. Fix those errors and resubmit as above.
5. Please also email your CSV to flucontest@cdc.gov until otherwise directed.

**Account creation instructions**

Create an account:

1. Go to <https://predict.phiresearchlab.org/> and click “Create Account” in the top right corner.
2. Fill in the fields and click “Register”
3. Send an email to episupport@cdc.gov as indicated to get your account approved.

Once your account is active:

1. Log into your account at <https://predict.phiresearchlab.org/>
2. You should see a field to create a team for the FluSight project.
3. Enter your model name here. While the interactive display is still under development, this is the name that will eventually be publically linked to your submission. If your team has preferred to remain anonymous, I will email you a team name to use so CDC knows who each submission is.
4. Send an email to mrx6@cdc.gov so I can approve your model to enter the project.