# Forecasting influenza activity using machine-learned mobility map

<https://doi.org/10.1038/s41467-021-21018-5>)

Background

* Forecasting is key to decision making
* 291,000–645,000 seasonal influenza-related deaths each year
* Active area of research in epidemiology
* CDC forecasting challenge
* Common predictors: search trends, social media, medical claims, weather data
* Statistical forecasting – capturing patterns in the time series
* Mechanistic representation – captures mechanism by which people are exposed

Study

* Disease transmits through person-to-person contact, in turn influenced by how much people move around
* So use Mobility data to predict – call data records, location history (machine-learned anonymized mobility map)
* Metapopulation Model (SEIR – susceptible, exposed, infected, recovered)
* Used MAPE to compare models
  + AMM
  + Commute
  + Gravity
  + Radiation
  + No Mobility
* 1-4 week ahead predictions
* Better predictions with higher quality, more granular mobility data (e.g. location history vs. commuter surveys)
* Code: https://github.com/NSSAC/PatchSim, https://github.com/NSSAC/

# A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States

<https://pubmed.ncbi.nlm.nih.gov/30647115/>

Background

* Highest incidence typically occurring in colder and drier months
* Forecasting can inform decisions about hospital staff, where to target public heath interventions, timing of public health campaigns
* FluSight – Forecast the Influenza Season Collaborative Challenge – CDC – starting in 2013/2014 <https://www.cdc.gov/flu/weekly/flusight/how-flu-forecasting.htm>

Study

* Comparison of 22 different models over seven flu seasons (2010/2011 – 2016-2017)
* Possible predictors:
  + Prexisting population immunity
  + Temp and humidity
  + Vaccine effectiveness
  + Timing of school vacations
* CDC evaluates forecasts using log score (evaluates precision and accuracy)
  + The log score for a model m is defined as log fm(z\*jx), where fm(zjx) is the predicted density function from model m for target Z conditional on some data x, z\* is the observed value of the target Z, and log is the natural logarithm.
* Baseline model - ReichLab-KDE achieved between 0.12 and 0.37
* Best model: CUEKF SIRS (compartmental model) achieved a region-specific average forecast score for week-ahead targets between 0.32 and 0.55
* 15/22 outperformed baseline in forecasting week ahead, only 7 of 22 outperformed baseline in 4-week ahead targets
* <http://flusightnetwork.io/>

# A Systematic Review of Studies on Forecasting Dynamics of Influenza Outbreaks

Background

* Why forecast? Allows for allocation of public heath resources like vaccines, antivirals and doctors/nurses

Study

* Existing Approaches
  + Measures predicted:
    - Magnitude
    - Peak Timing, Intensity
    - Duration
  + Evaluation metrics:
    - Correlation coefficient between predicted and observed values (allows for comparison of trends but doesn’t measure closeness between observations and predictions
    - Percent Error
    - RMSE
    - Confidence Intervals
    - Comparing predicted peak time with actual peak time
  + Common Models
    - ARIMA
    - Method of Analogs
    - Compartmental Models
    - Agent-based models
    - Metapopulation models
  + Features
    - Temperature and humidity

# A Comparative Evaluation of Time Series Models for predicting influenza outbreaks from sentinel sites of healthcare centers in Iran

Background

* Forecasting accurately allows for public health messaging and raising awareness
* 3-5 million severe illnesses annually
* 250,000 – 500,000 deaths/year
* One of most important causes of mortality worldwide

Study

* SVM, random forest, ANN
* Best model = random forest:
  + RMSE = 22.78
  + MAE = 14.99
  + ICC = 0.88
* Data retrieved from WHO’s FluNet
* Predictors
  + 52 previous weekly case numbers
  + Year
  + Season
  + Week