Dengue forecasting
Model and challenges

Michael A. Johansson
Epidemic Analytics Unit, Dengue Branch, Division of Vector-Borne Diseases
San Juan, Puerto Rico
Dengue – San Juan, Puerto Rico

Number of dengue cases

0 100 200 300 400

Evaluate forecasts on out-of-sample data (over multiple years for dengue).
Dengue forecast error - Mexico

![Graph showing mean absolute error over months ahead for Mexico. The graph indicates a relatively flat line, suggesting consistent forecast error throughout the 6 months.]
Dengue forecast error - Mexico

![Graph showing mean absolute error over months ahead]

- **Mean**
- **Monthly mean**

Johansson et al. Scientific Reports 2016
Dengue forecast error - Mexico

Mean Monthly mean Temperature

Mean Monthly mean

mean absolute error

months ahead

Johansson et al. Scientific Reports 2016
Dengue forecast error - Mexico

![Graph showing the mean, autocorrelation, and temperature over months ahead.](image)
Dengue forecast error - Mexico

Compare to a baseline model.

Graph showing mean absolute error over months ahead with lines for Mean, Autocorrelation, Temperature, Monthly mean, and Autocorrelation + seasonality.
Forecasts - Mexico

Assess the uncertainty.
BS Checklist for Forecasts

☐ Evaluate forecasts on out-of-sample data.
☐ Compare to a baseline model.
☐ Assess the uncertainty.
Dengue forecasting research

“[poor prediction was] the result of the unusual behavior that occurred between 2009 and 2011”
The state of dengue forecasting

- Many models
- Mostly retrospective
- Varying targets & evaluation metrics
- Little sense of appropriateness of models for decision-making
- No quantitative models being routinely used for decision-making

As of 2010: 60+ published mechanistic dengue models

Dengue Forecasting Project

- Pandemic Prediction & Forecasting Science & Technology Working Group
- June–September, 2015
- Targets: Peak week, peak incidence, and total incidence over 8 seasons in Iquitos, Peru and San Juan, Puerto Rico
- 16 teams; 10,000 forecasts
- dengueforecasting.noaa.gov, predict.cdc.gov
Correlation of point forecasts is not enough

We need to assess both accuracy and confidence (i.e. certainty/uncertainty).
Error metrics are simple and straightforward

<table>
<thead>
<tr>
<th>Team</th>
<th>Forecast Peak Week</th>
<th>Observed Peak Week</th>
<th>Error (weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team A</td>
<td>23</td>
<td>32</td>
<td>9</td>
</tr>
<tr>
<td>Team B</td>
<td>23</td>
<td>32</td>
<td>9</td>
</tr>
<tr>
<td>Team C</td>
<td>22</td>
<td>32</td>
<td>10</td>
</tr>
</tbody>
</table>
Probabilistic forecasts have more information

Team A
Point prediction

Team B
Point prediction

Team C
Point prediction
Assessing probabilistic forecasts

Team A

Point prediction

Observed peak

p = 0.02

Team B

Point prediction

Observed peak

p = 0.04

Team C

Point prediction

Observed peak

p = 0
Forecast calibration

Well-calibrated  Over-confident  Under-confident  No confidence  No resolution
Week 12 forecast for San Juan 2012/2013
When are forecasts best?

~12,000 forecasts

- 2 locations
- 8 seasons
- 19 models
Nowcast/situational awareness

SEASONAL DENGUE
Peak week forecasts

SHORT-TERM INFLUENZA
1- to 4-week ahead forecasts

San Juan
logarithmic score

Week of forecast

Skill

Weeks-ahead
Promising approaches

- Simpler models
  - No climate data (dengue)
  - No vector model (dengue)

- Ensembles
  - Simple ensembles (across targets, seasons, & diseases)
  - Prospectively defined
  - Current standard for influenza (since 2017/18)
Key questions

- What are the key surveillance data?
- How much do vectors matter?
- What is the contribution of weather?
- What is the role of immunity and enhancement?
- What is the role of mobility and spatial heterogeneity?
Conclusions
“Dengue is a disease of the tropical and subtropical regions, and within these zones it has a marked preference for the hot season - for summer.”

- Hermann Nothnagel, 1905
“It is difficult to make predictions, especially about the future.”
How can infectious disease forecasting improve? (How has weather forecasting forecasting improved?)

- Data
- Analytical tools
- Computational power
- Evaluation
- Standardization & interoperability
CDC Epidemic Prediction Initiative

- Connect researchers to data
  - Dengue, influenza (github.com/cmu-delphi/delphi-epidata), Zika (github.com/cdcepi/zika)

- Develop an analytical pipeline
  - predict.cdc.gov
  - Current: Influenza, *Aedes*

- Build a community
  - Centers for Disease Control and Prevention, Researchers, Multiple US Departments & Agencies, Council of State and Territorial Epidemiologists
Conclusions

- Surveillance and forecasting go hand in hand.
- Current forecasting methods improve upon expert knowledge and can be helpful for situational awareness.
- Improved analytics can improve our ability to predict and respond effectively to arboviral disease epidemics.
Key considerations

- Connect forecasts to decision making needs.
- Evaluate forecasts on out-of-sample data.
- Compare to a baseline model.
- Assess the uncertainty (including calibration).
- Use more than one model.
- Use forecasts as one input for decision making.
Acknowledgements

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

The Epidemic Prediction Initiative community

CDC Epidemic Prediction Initiative

Matt Biggerstaff
Craig McGowan
Juan Sanchez Montalvo
Luis Mier-y-Teran Romero
Dania Rodriguez Vargas
F. Scott Dahlgren
Chelsea Lutz

Office of Public Health Preparedness and Response
Division of Vector-Borne Diseases
Influenza Division
Council of State and Territorial Epidemiologists

Nicholas Reich (University of Massachusetts Amherst)
Aditi Hota (Harvard University)
Mauricio Santillana (Harvard University)
John Brownstein (Boston Children’s Hospital)

For more information, contact CDC
1-800-CDC-INFO (232-4636)

mjohansson@cdc.gov