DSI Project 4: West Nile Virus in Chicago

Team:
GAS the Vectors

Introduction - Who are We?

G wen
A lfred
S cott

GAS the Vectors

A vector is mathematical construction that conveys direction and magnitude.

A vector is also the term for an organism that spreads a disease such as insects, rodents, etc.

West Nile Transmission Path

West Nile virus lives in birds ➔ Mosquitoes feed on birds
Infected mosquitoes ➔ Bites humans

Key take away: Newly hatched mosquitoes do not carry the virus. They obtain the virus from other animals.

First human case hits city in 2002

- Discovered in horses and birds in 2001

- 225 confirmed cases in first year
  - 42 deaths reported in Cook County
- Chicago leads nation in WNV deaths in 2003

West Nile Virus in Chicago

First human case hits city in 2002

- Discovered in horses and birds in 2001

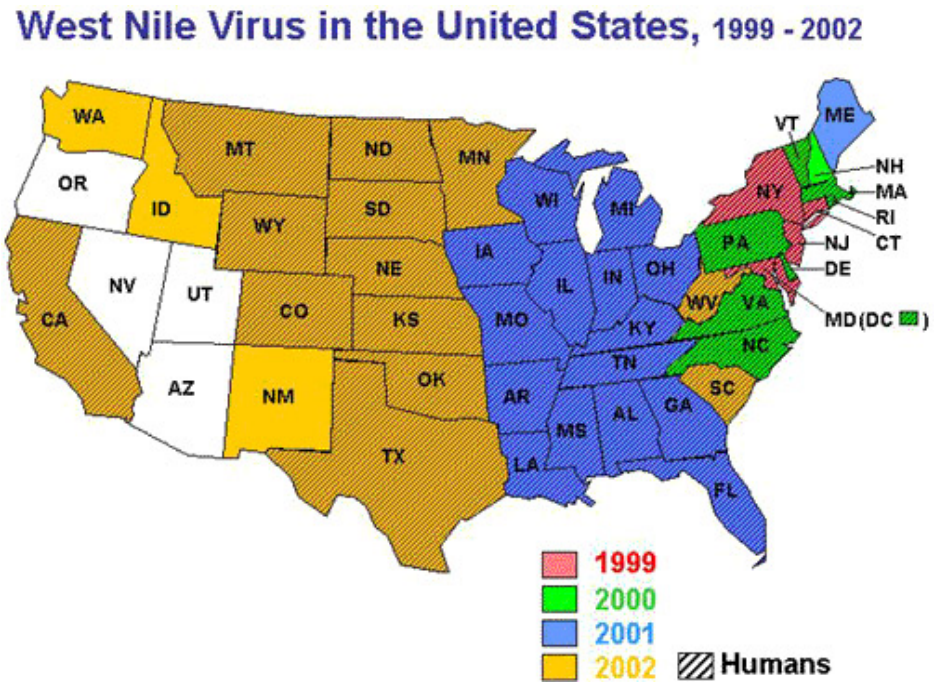
225 confirmed cases in first year

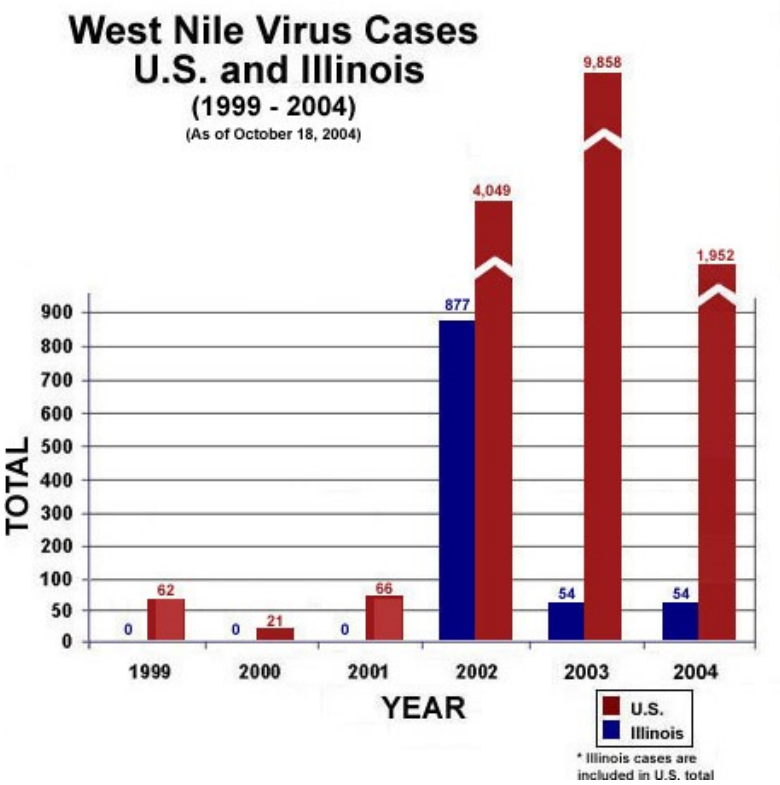
- 42 deaths reported in Cook County

Chicago leads nation in WNV deaths in 2003

- 64 deaths
- About half of deaths are in people over 60
  - Only 10% of city population is over 60

History of West Nile Virus in Chicago





Public Health Impact of West Nile Virus

8 out of 10 are asymptomatic

1 in 5 will have some febrile illness

- Head and body ache, joint pain, vomiting, diarrhea, rash

1 of 150 result in serious symptoms (neuroinvasive)

- Encephalitis (inflammation of the brain)
- Meningitis (inflammation of membranes surrounding brain and spinal cord
  - 1 in 10 result in death

Description of Data

Spray

- data from 8 dates of spraying
- 14,294 observations after data munging

Weather

- from 2 stations, for all years
- May – Oct; 2007 – 2014

Trap data

- 2007, 2009, 2011, 2013
- 10,506 records, ~116,000 predictions requested

Data Cleaning / Interpretation – Key Issues

Spray data

- issue with date/time & locations, could infer actual amounts

Traps

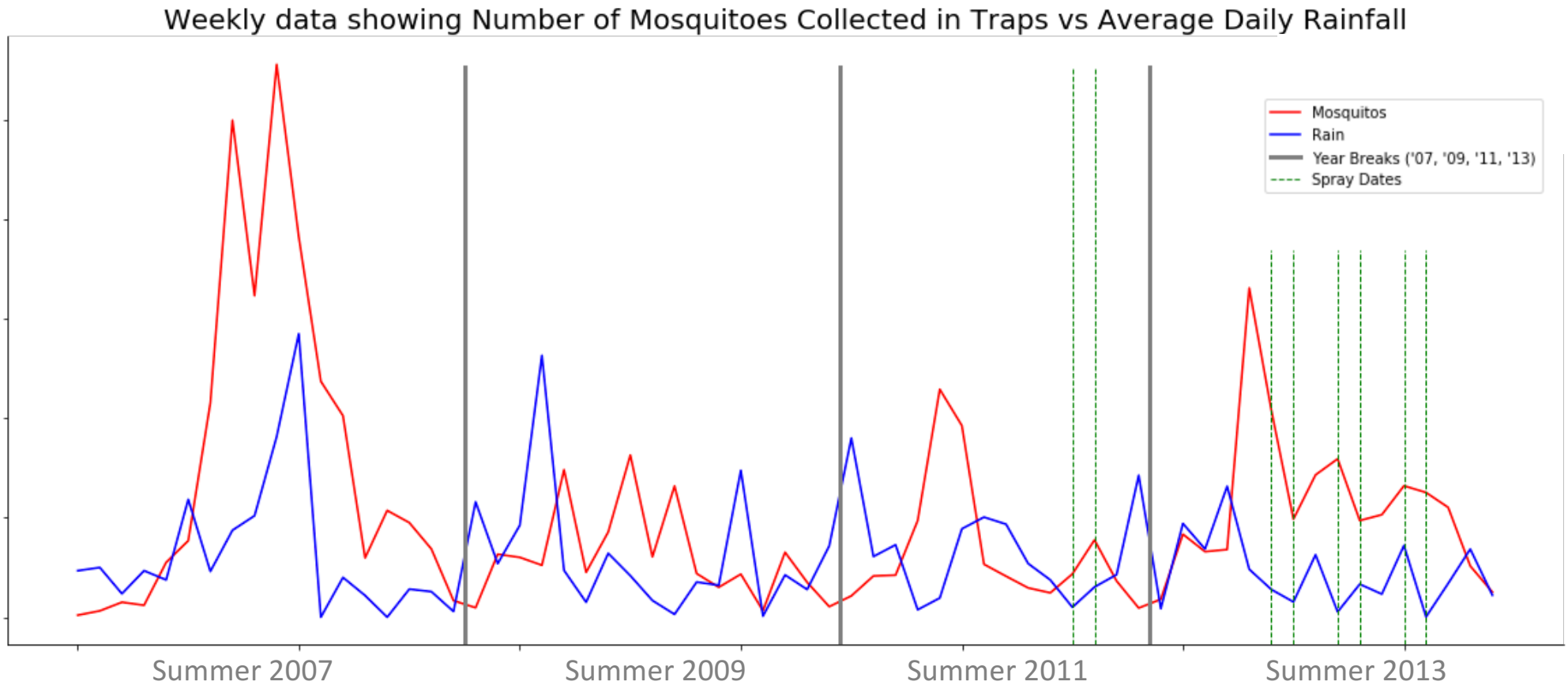
- one trap had two (long, lat)

Weather

- some “M” and “T”, Sunset times of 1860.

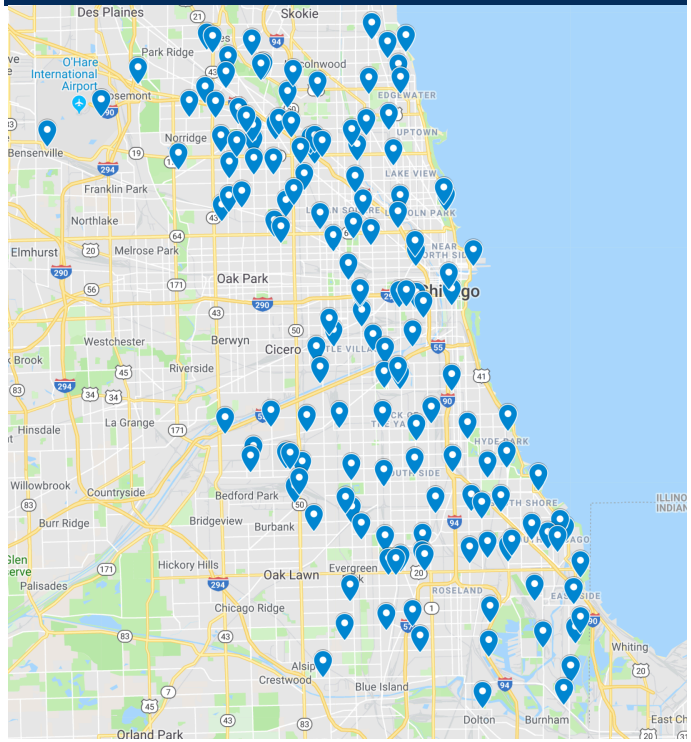
50 maximum mosquitoes in a trap per observation

Key Data Elements Over Time – Rain, Mosquitos, and Spray Data

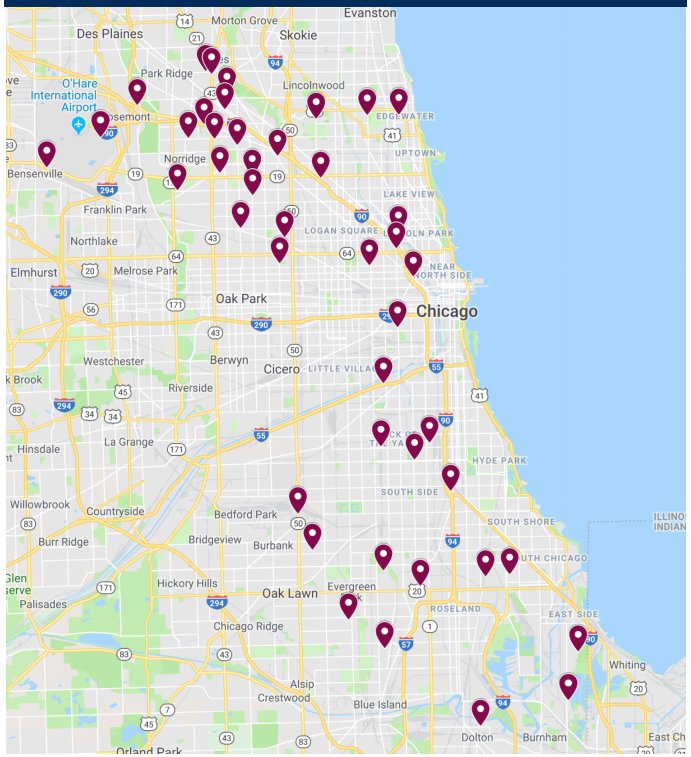


Mapping Traps & Hotspots

Map of All Traps



Map of Hotspots



Traps with higher empirical probabilities of collecting samples that contained West Nile. Shows a hotspot near O'Hare.

Modeling

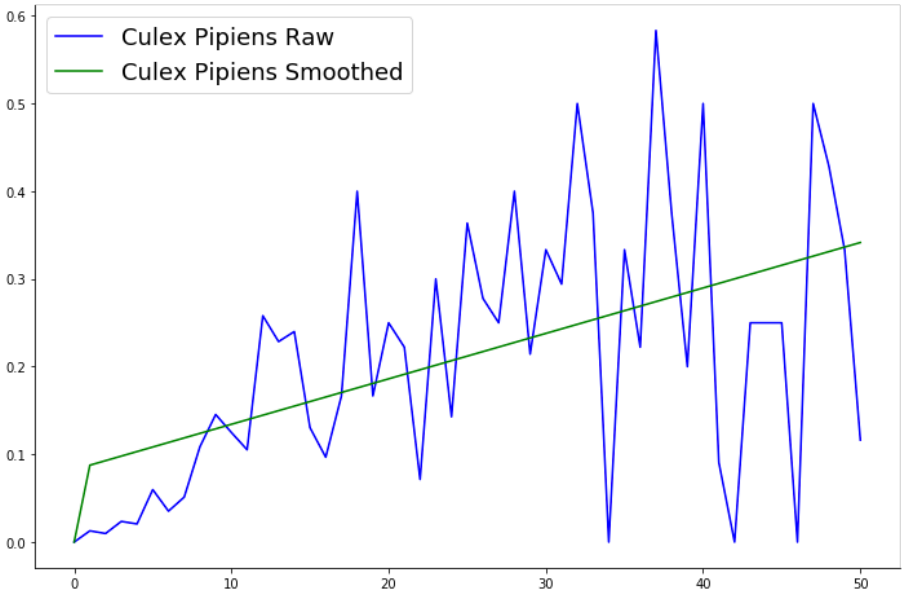
Overall goal of the modeling is to predict probability of WNV at various locations in Chicago during summer months. This then informs a cost/benefit analysis on prevention.

We first predicted the mosquitos levels expected (as measured in the traps) and then used that information to then predict the probability of the presence of WNV.

Modeling – Key Data Elements Useful in our Models

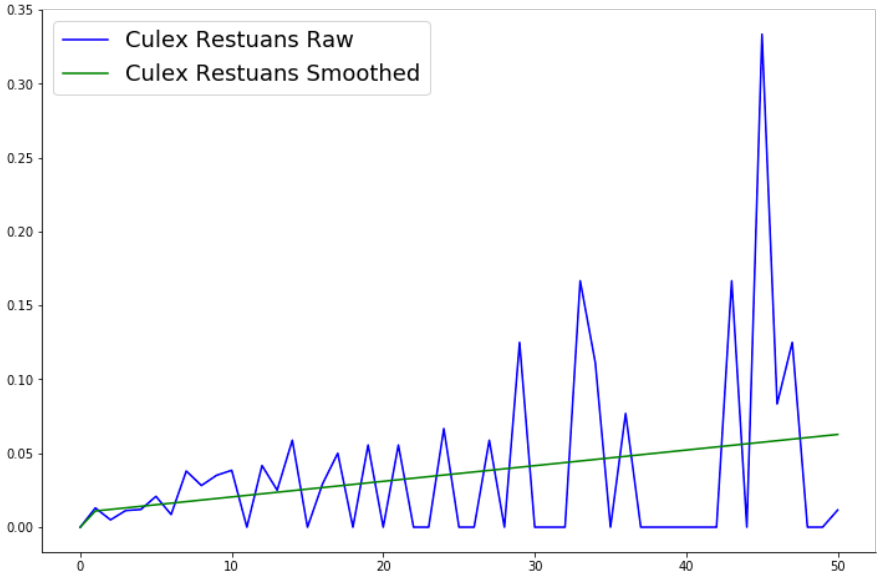
- Longitude and Latitude
- The week of the year when observations were made
- The type of mosquitoes present in traps
- The type of mosquitoes that transmit WNV
- Tried clustering because more traps were listed in the test versus tests in the training data. (not useful)
- Date since emptied for traps (not useful)
- Week of the year was useful
- Days since spraying if there was spraying (not robust)
- Aggregation of duplicate rows (lost information)

Modeling – Use EDA to estimate P(WNV | (Spec & # in Trap))



Could have probably raised the ‘smoothed’ probabilities

Step 1, predict # of Mosquitos
Step 2, look up probabilities from EDA



Modeling – Predicting Mosquitoes per Trap

- 8 Regression models: LR was poor, Ridge, Elastinet, Lasso, Random Forrest, Gradient Boosting, AdaBoost
- Grid search used to tune hyper-parameters
- GradientBoosting was the best model

- Forced predictions to fit our range of possible values (negative values to zero, if test had multiple rows, we set value to 50).

Modeling – Predicting Probability of WNV

Used the Mosquito prediction to augement test data.

Then modeled probability of WNV present by:

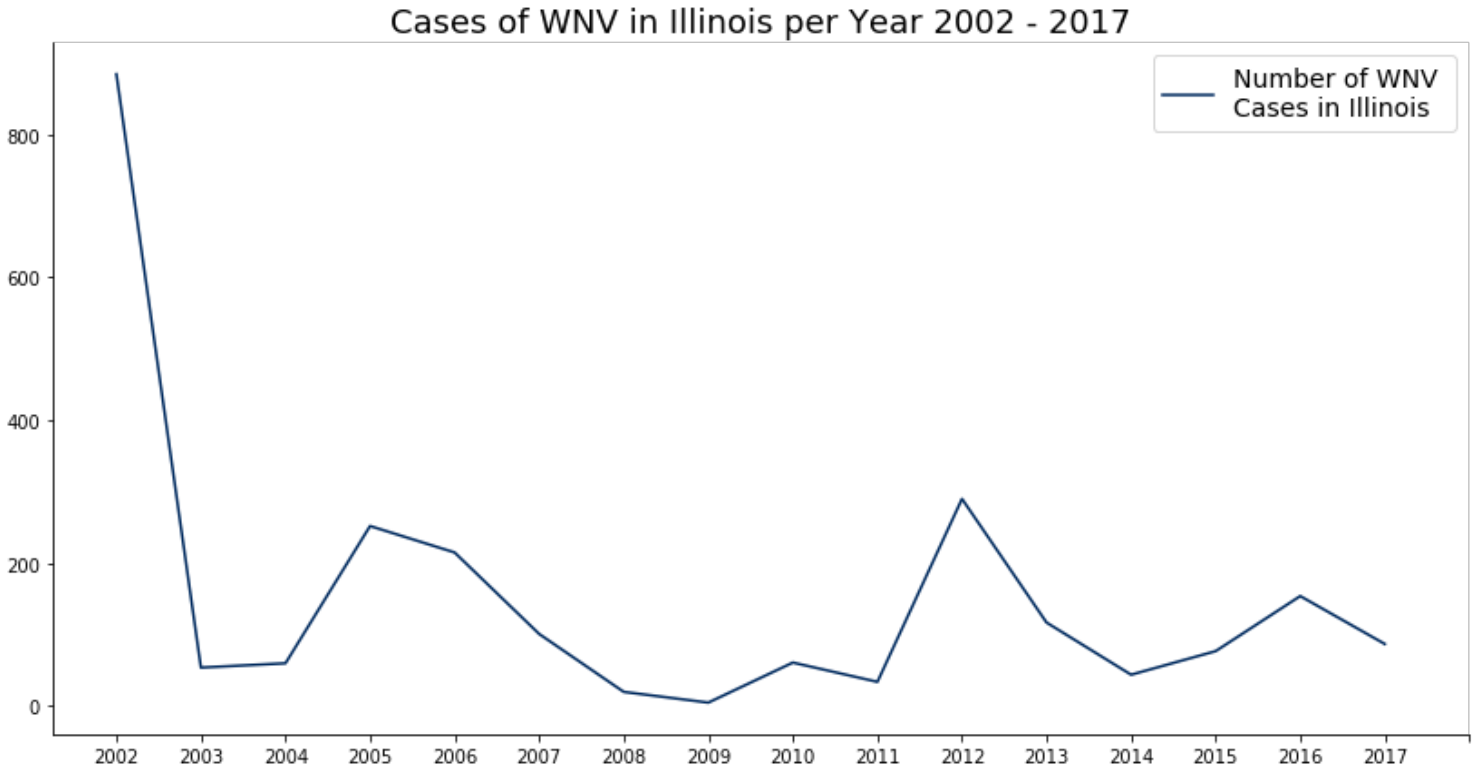
- Classification modeling with: LogisticRegression, KNN, and various decision trees.
- All gave relatively good scores
- LogisticRegression used as final model
- Train/Test split AUC ROC was 78% (not Kaggle score
- Only 3 species with non-zero probabilities

Modeling Next Steps

Things we would like to pursue to improve the model.

1. Expand dataset to include observations where there were no mosquitos found.
2. Weather data included in prediction models
3. Length of day using sunrise, sunset information

Cost Benefit Analysis – Neuro and Non-Neuro Invasive Cases



Cases have run as high as 880 per year.

Chicago has 84% of Illinois population.

Source: CDC

Cost Benefit Analysis – Current State / Costs

City of Chicago currently uses Zenvix as of August 2018

- Adulticide formulated to also kill larva
  - Approved in all states except New York
- Approved for use on crops and in Urban areas

- 30 Gallon barrel -- $11,595
  - At recommended application rate of 0.007 lb/acre
  - $1,171.20 per sq mile

- City of Chicago allocations $5.3 million for mosquito control
  - On a steady 3% increase year to year

Cost Benefit Analysis – Recommended Areas for Spraying

With the use of our data, we can narrow the optimal areas for increased spraying:

Targeting 45 square miles of Chicago with

- Highest occurrences
  - East of O'Hare to Lake Michigan (16 sq miles)
- Heavy traffic
  - O'Hare Airport (9 sq miles), Lincoln Park area (4 sq miles)
- Areas with vulnerable persons
  - Hospitals and senior centers

Cost Benefit Analysis – Public Health Benefits

Benefit of preventing a case of WNV in humans: $27,000 - $133,000 with a mean of $33,000.

Ref: "Cost effectiveness of a targeted age-based West Nile virus vaccination program" by Shankar, et al. Elsevier public health journal, May 2017

Average number of cases of West Nile in Chicago ~100.

Maximum cost benefit = $3.3m if WNV eliminated

Recommendation

We are recommending that the Chicago Public Health department implement the following measures:

1. Selective Spraying
   - 45 square miles x 25 weeks x 1,170 = $1.3m
   - Would need approximately 1/3 reduction in cases to be C/B neutral
2. Use of Ovillanta trap
3. Biological Control

Trapping

Ovillanta

- Developed at Laurentian University in Ontario
- Cost $3.08
- Traps and kills
- Self-sustaining

Biological Control

Methods for long term mosquito deterrence

- Work with parks department on tree/flower selections
  - Permanent incorporation of plants that repel mosquitos
  - Marigold, citronella, lavender, Eucaluptus
- Work with Wildlife & Watershed Management on species control, incorporate
  - Dragonflies, bats
  - Koi and Carp
- Mosquito Sterilization – Tap into Federal Efforts

Thank you for your time

Questions?