# Trends of Travel-Associated Disease in Florida and Evaluation of Reportable Disease Data Quality 2006-2015

Blake Scott, MPH

February 2017

# Research Outline

- Background
  - Trends in tourism
  - Growth of travel health surveillance

- Data Subset Development and Analysis
  - Merlin subset development
  - Trend analysis

- Data Quality Evaluation
  - Restrictions of analysis based on data quality
  - What information can't Merlin provide and how can it be improved

- Retrospective and Limitations
  - Lessons learned
  - What the analysis provides vs. what has to be inferred

- Conclusions
  - Trend analysis can assist in predicting travel-associated disease occurrence and can be made more reliable with improved data collection and storage
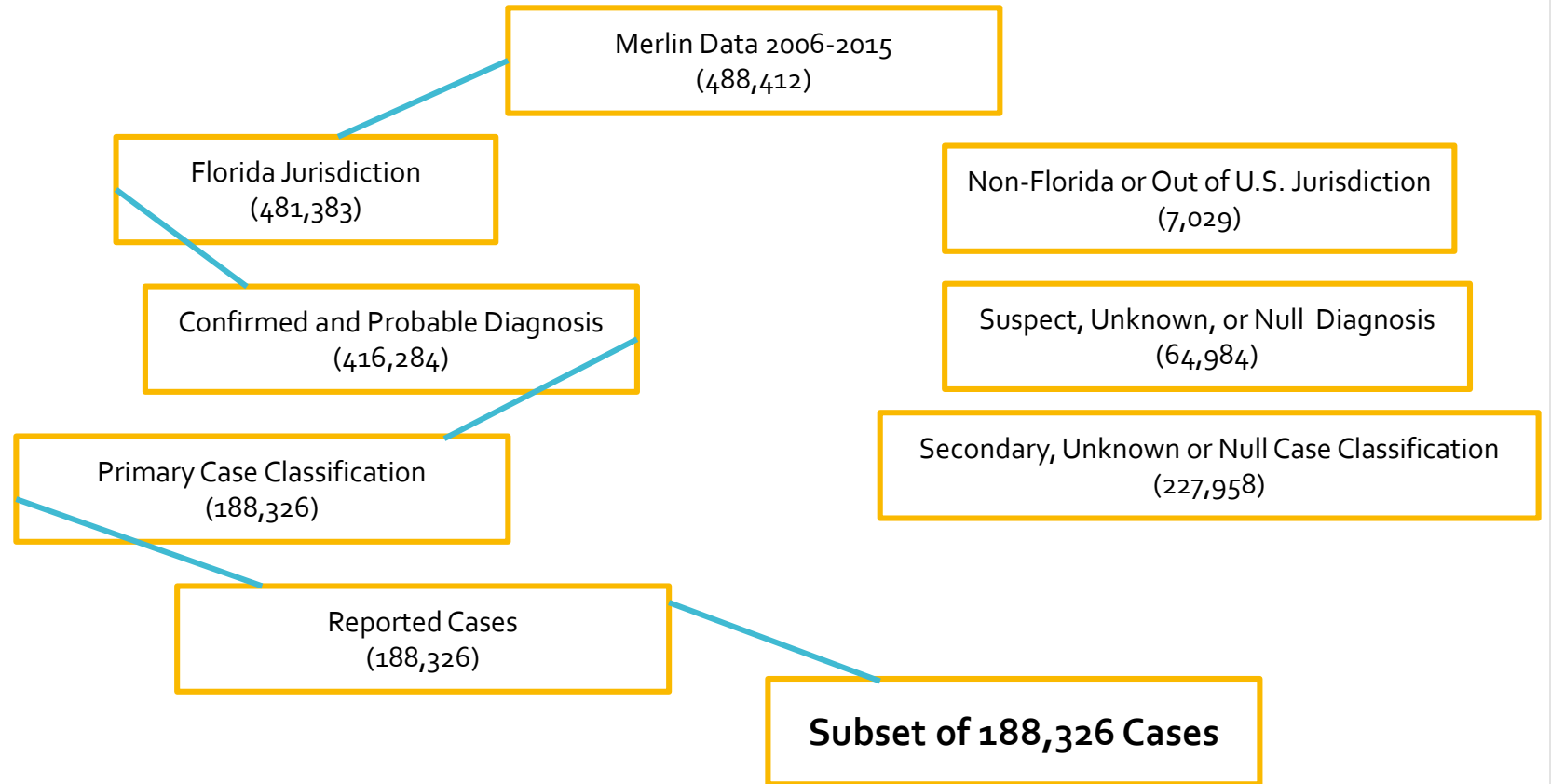
# Aims and Objectives

## Aims

- Develop a comprehensive trend analysis of travel-associated disease

- Improve data collection and storage of travel-associated cases in Merlin

## Objective

- Create a resource of predictable trends in travel-associated disease incidences in Florida residents for public health and travel health professionals

Data Analysis Methodology

• Creating the data subset

Merlin Data 2006-2015
(488,412)

Florida Jurisdiction
(481,383)

Non-Florida or Out of U.S. Jurisdiction
(7,029)

Confirmed and Probable Diagnosis
(416,284)

Suspect, Unknown, or Null Diagnosis
(64,984)

Primary Case Classification
(188,326)

Secondary, Unknown or Null Case Classification
(227,958)

Reported Cases
(188,326)

**Subset of 188,326 Cases**

• Additional Data Fields:

Imported Status
Date of Onset
ICD 9 Code
County
Gender
Age

# Data Table:

Imported Status by Year

| | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|---|---|---|---|---|---|---|---|---|---|---|
| Acquired in Florida | | 1 | 3111 | 12903 | 27195 | 21619 | 18704 | 19891 | 23009 | 22274 |
| Acquired Outside of the U.S. | | | 129 | 535 | 1764 | 1380 | 1099 | 1180 | 1483 | 1277 |
| Acquired in the U.S., not Florida | | | 57 | 225 | 344 | 600 | 508 | 634 | 759 | 640 |
| Unknown | 85 | 88 | 176 | 1313 | 6499 | 5525 | 3845 | 3745 | 2913 | 2816 |

# Preliminary Analysis

- Disease acquired in the U.S., but not in Florida
  - 55% Males and 45% Females
  - Most frequent age group: 60-69
  - Most frequent county of origin: Marion
  - Hepatitis C, Chronic
    - Highest observation frequency was from New York

- Diseases acquired outside of the U.S.
  - 47% Males and 53% Females
  - Most frequent age group:30-39
  - Most frequent county of origin: Miami-Dade
  - Acute Giardia
    - Highest observation frequency (by far) is Cuba

# Data Limitations

- 5 years of data versus 10
- Origin field is open text
- 'Unknown' Origin field is inconsistent
- Additional travel data in other tables within Merlin
- Difficult to compare analysis with tourism data
- Difficult to compare to GeoSentinel data
- No specific travel clinic data collection

# Data Quality Evaluation

- Improvements already made:
  - New imported field categories
  - Exposure during travel period questions
  - Drop-down list of states and countries
  - Requirement of travel related fields to be entered
  - Specific travel location fields added

- Suggested improvements:
  - More in-depth travel questions
  - Storing all travel data in one table
  - Data cleaning of Origin field
  - Auto-creating data of global region
  - Pre and Post travel clinic questions

# Future Analysis and Results

- Seasonality of travel-associated diseases

- Data quality of open text Origin field

- Data quality improvement of information collected from improvements already made to Merlin

- Discovering potential trends of specific travel-associated diseases
  - Season
  - Age Group
  - Origin
  - Statistical analysis of potential trends

# Next Steps

- IRB Requirements

- Data cleaning

- Trend and Statistical Analysis

- Discussion and implementation of data quality measures

- Continued research on data quality improvement

- Partnership with Florida travel clinics

# Discussion and Summary

- Several limitations to the Merlin data subset

- Trend analysis could still provide a fairly predictable report of travel-associated disease occurrence in Florida

- Trends can be made more reliable with improved data collection and storage in Merlin

- There will always be data limitations

# Future Studies

- Comparing a thorough analysis of travel-associated cases of a specific disease to the data collected from C.D.C and GeoSentinel data through collaboration

- Work in partnership with Florida travel clinic to survey habits of pre-travel consultation patients

# Resources

- Merlin Reportable Disease System of Florida

- Visit Florida tourism information

- Talbot, E.A., MD. ,et al (2010) Travel Medicine Research Priorities: Establishing an Evidence Base. *Journal of Travel Medicine ,Vol 17* (issue 6): 410-415

- Harvey, K., MPH et al (2013) Surveillance for Travel-Related Disease-GeoSentinel Surveillance System, United States, 1997-2011. *Morbidity and Mortality Weekly Report, Vol 62* (issue 3)

- Leder,K., et al (2013) Travel-associate Illness Trends and Clusters, 2000-2010. *Emerging Infectious Diseases, Vol 19* (issue 7)

# Questions?

Thank You!