Update on Analysis of Shared Syndromic Data
for Severe Respiratory Injury Surveillance (Internal Use Only)

Michael A. Coletta, MPH, Aaron Kite-Powell, MS

Outbreak of electronic-cigarette, or vaping, product use-associated lung injury (EVALI)
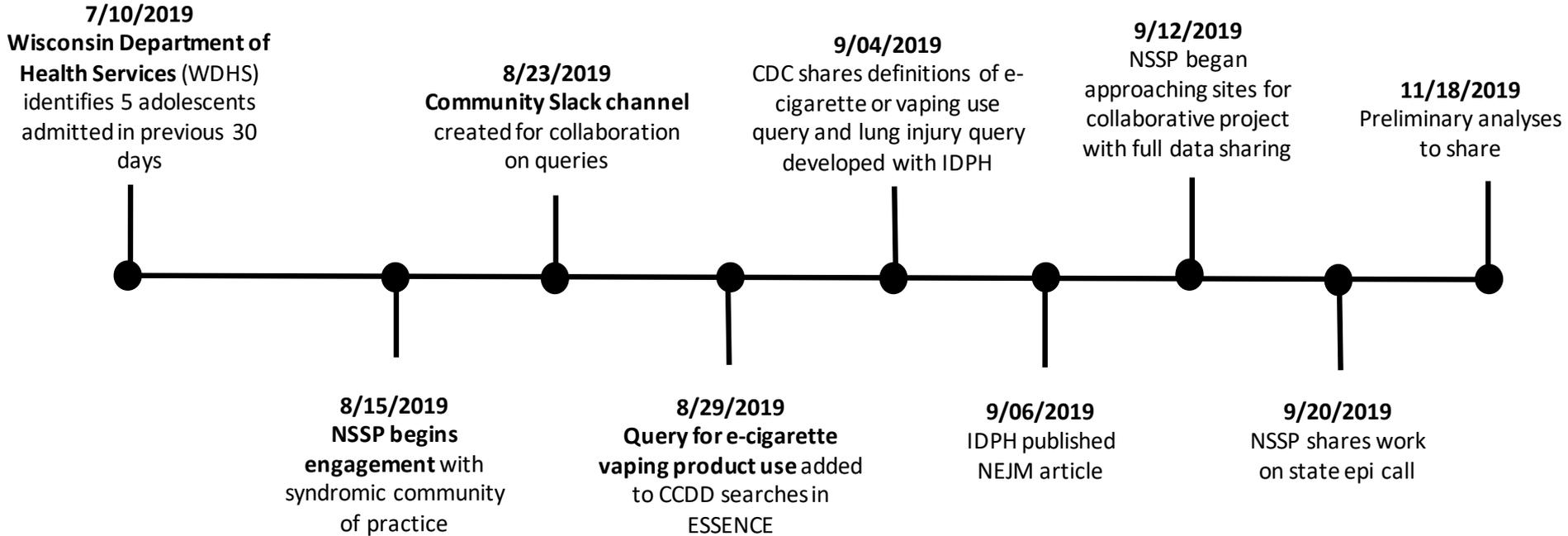
As of November 14, 2019:

- 49 states, Puerto Rico, District of Columbia (DC) and the U.S. Virgin Islands
- 2,172 confirmed and probable cases reported
- 42 deaths

As of October 15:

- In 3 months before symptoms, 11% reported exclusive nicotine use and 34% reported exclusive tetrahydrocannabinol (THC) use
- 86% reported any (THC)-containing products, 64% any nicotine

Timeline

7/10/2019
Wisconsin Department of Health Services (WDHS) identifies 5 adolescents admitted in previous 30 days

8/23/2019
Community Slack channel created for collaboration on queries

9/04/2019
CDC shares definitions of e-cigarette or vaping use query and lung injury query developed with IDPH

9/12/2019
NSSP began approaching sites for collaborative project with full data sharing

11/18/2019
Preliminary analyses to share

8/15/2019
NSSP begins engagement with syndromic community of practice

8/29/2019
Query for e-cigarette vaping product use added to CCDD searches in ESSENCE

9/06/2019
IDPH published NEJM article

9/20/2019
NSSP shares work on state epi call

Queries

- **E-cigarette or vaping**: Terms related to e-cigarette use or vaping in the chief complaint

- **Lung injury query, version 1 (sensitive)**: Includes acute respiratory illness, and excludes unrelated infectious disease, chronic disease, and injury; originally developed by IDPH for case-finding

- **Lung injury query, version 2 (specific)**: Includes only CDC-recommended codes for EVALI with same exclusions as version 1
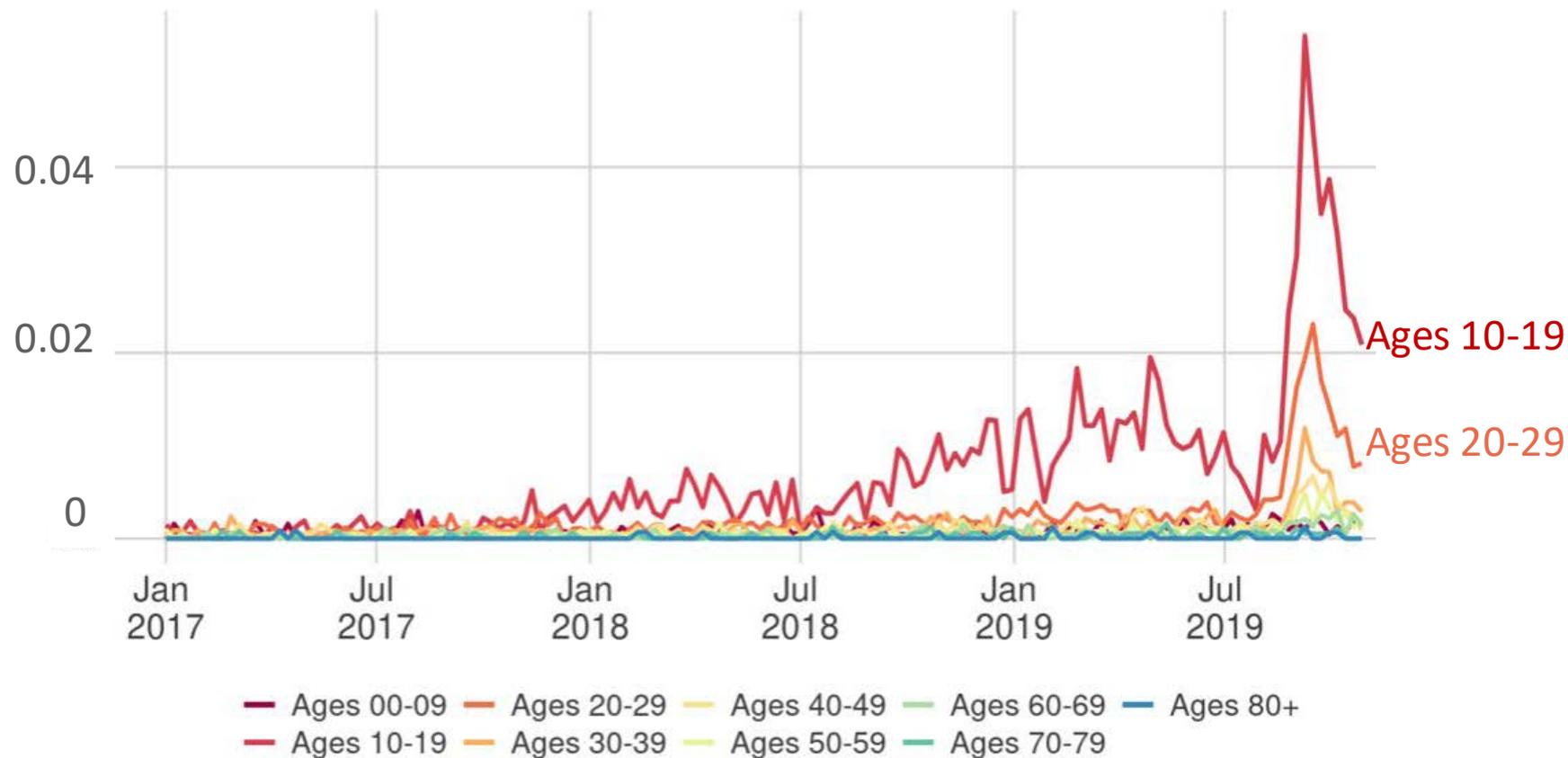
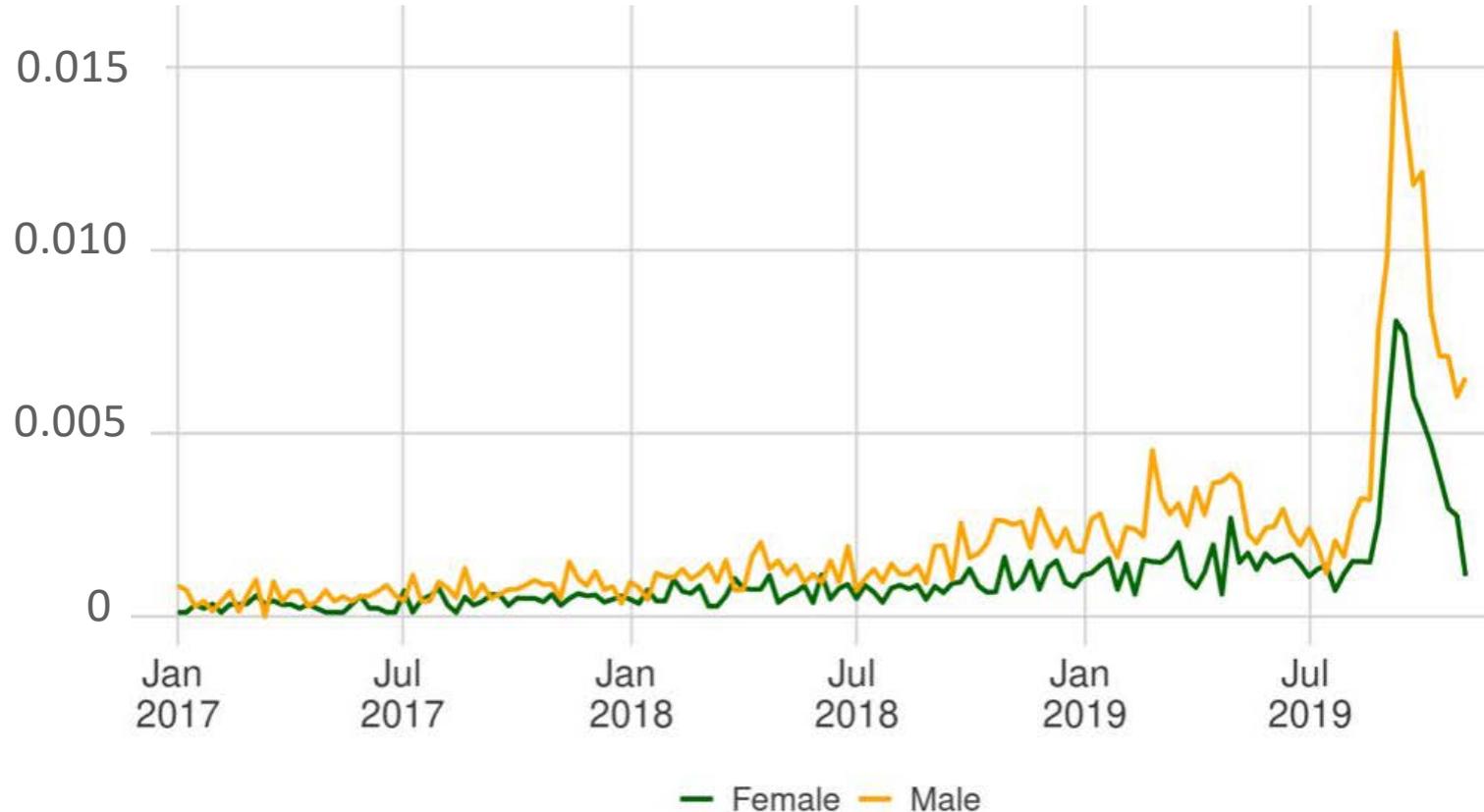# NSSP ESSENCE Use: Total number of queries for lung injuries or e-cigarette or vaping

Aug. 29: E-cigarette and vaping query added to ESSENCE
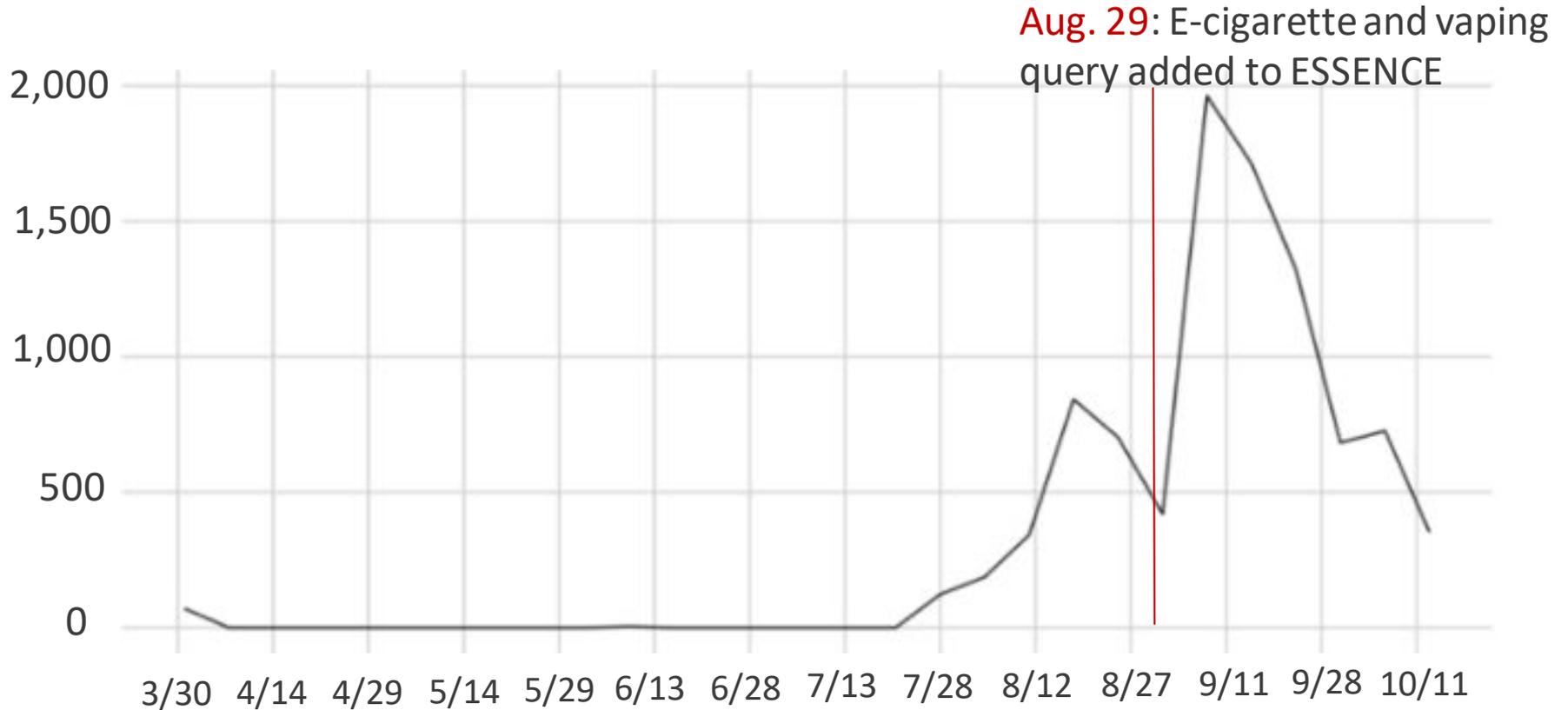
E-cigarette or vaping in chief complaint by age 2017–2019

Percent of all ED visits for each age group

E-cigarette or vaping in chief complaint by sex 2017–2019

Percent of all ED visits for each sex

Why share data?

- Without data sharing from sites, CDC cannot access details like patient age
- Response asked NSSP to help answer the question of when did this start?
  - More sophisticated methods to explore
  - Three main approaches pursued
    1. Difference of differences (ANCOVA)
    2. Text-based analyses
    3. Spatial and spatio-temporal analyses

1st Analytic Approach - ANCOVA

Trend analysis by sex

Observation

- At a national level we noticed a difference between men and women in the trend for severe respiratory illness.
- We saw this same trend in Illinois after they shared their data.

Methods

- Initially we used a time series approach to see if there was a any trend.
- We also used a difference in differences (ANCOVA) approach to determine the significance of the difference between men and women between two time periods.

Findings

- At a national level, large amount of seasonality within the data that is not controlled for with the denominator. However there is a statistically significant difference between men and women across both time periods.
- In Illinois, we see the data lacks a strong trend for men prior to April 2019, and has a steep upward trend after that point. There is also a statistically significant difference between men and women across both time periods.

Illinois lung injury syndrome (sensitive) by sex, ages 14–30 not discharged home
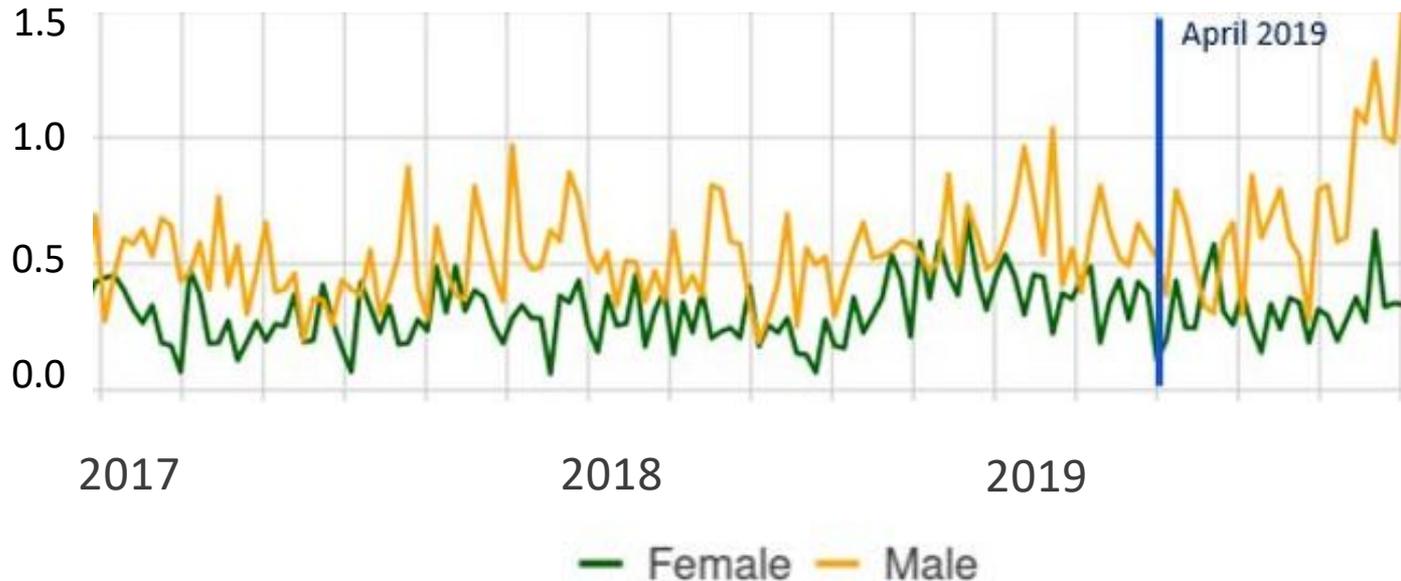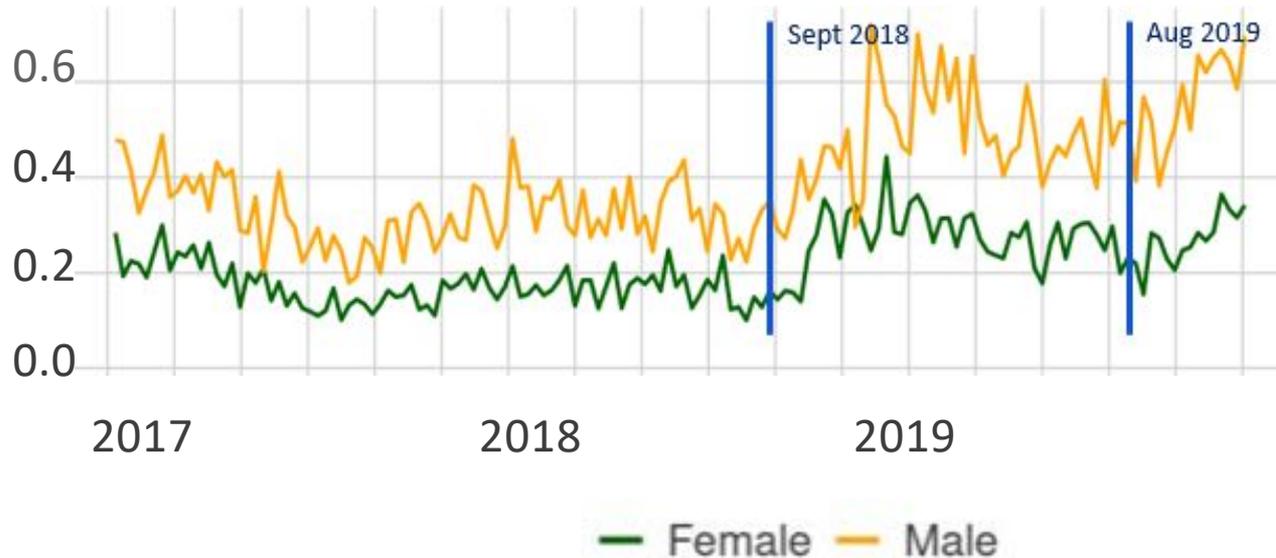
Percent of visits

Before April 2019: males ages 14–30 seem to have flat trend

Post April 2019: males ages 14–30 seem to have increasing trend

Shared site trends for lung injury syndrome (sensitive) by sex among ages 14–30, not discharged home

Percent of visits

**Post September 2018:** increase in SRI visits in ages 14–30

Post August 2019: increasing trend in both males and females ages 14–30

Lung injury syndrome (sensitive) – exploratory text analysis of shared data

- SRI Query (CCDD free text; Negations across multiple fields) limited to the sites that shared data

- Methods for Exploratory Text Analysis

  – Text and ICD Code N-gram Frequencies

  – Word network co-occurrence graphs

  – N-gram trends over time, by gender

  – Comparison of control and test time period

    • Purpose is to answer: What N-grams are more likely to come from the current outbreak period as compared with control periods?

Lung injury syndrome (sensitive) – weighted log odds ratio of top 10 discharge diagnosis unigrams
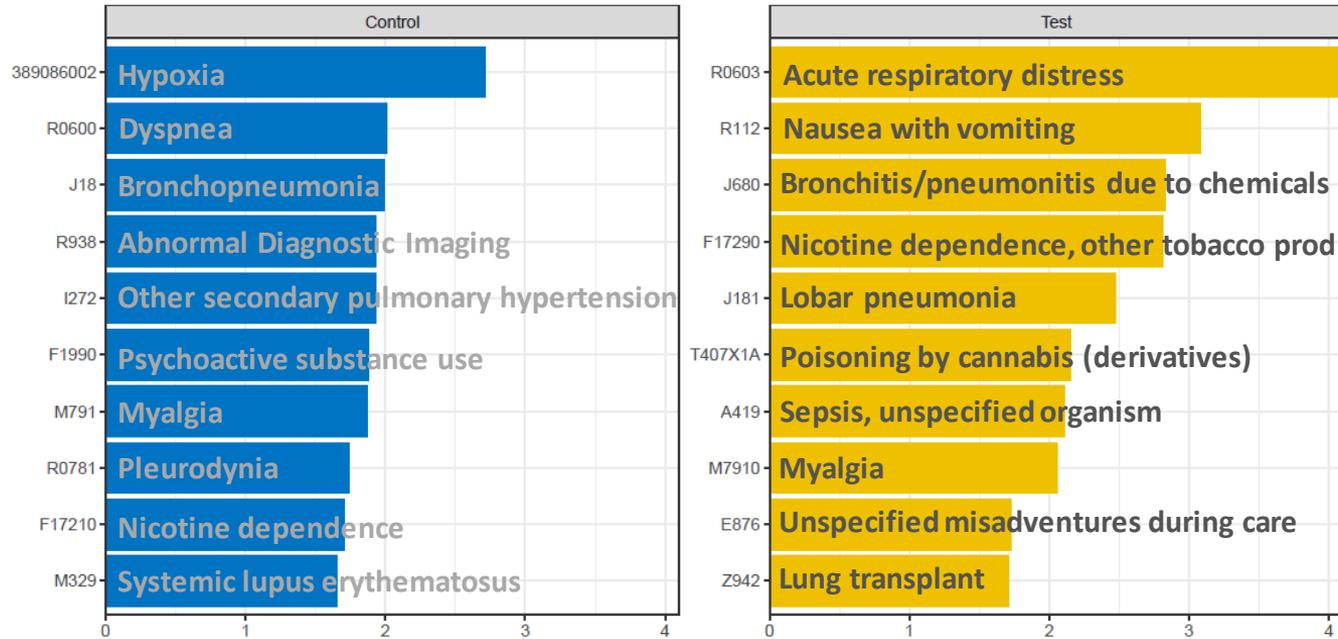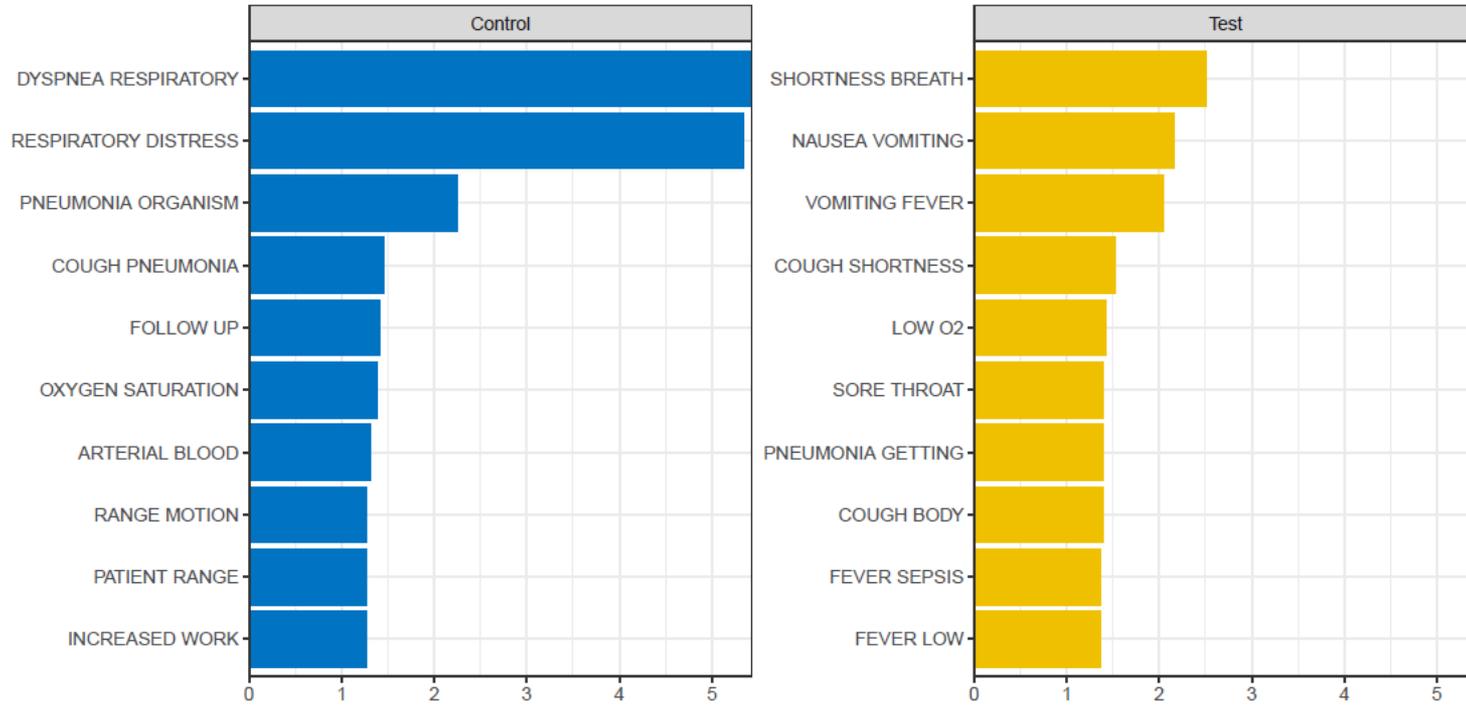

Tidylo R package:
https://github.com/juliasilge/tidylo

Lung injury syndrome (sensitive) – weighted log odds ratio of top 10 chief complaint bigrams
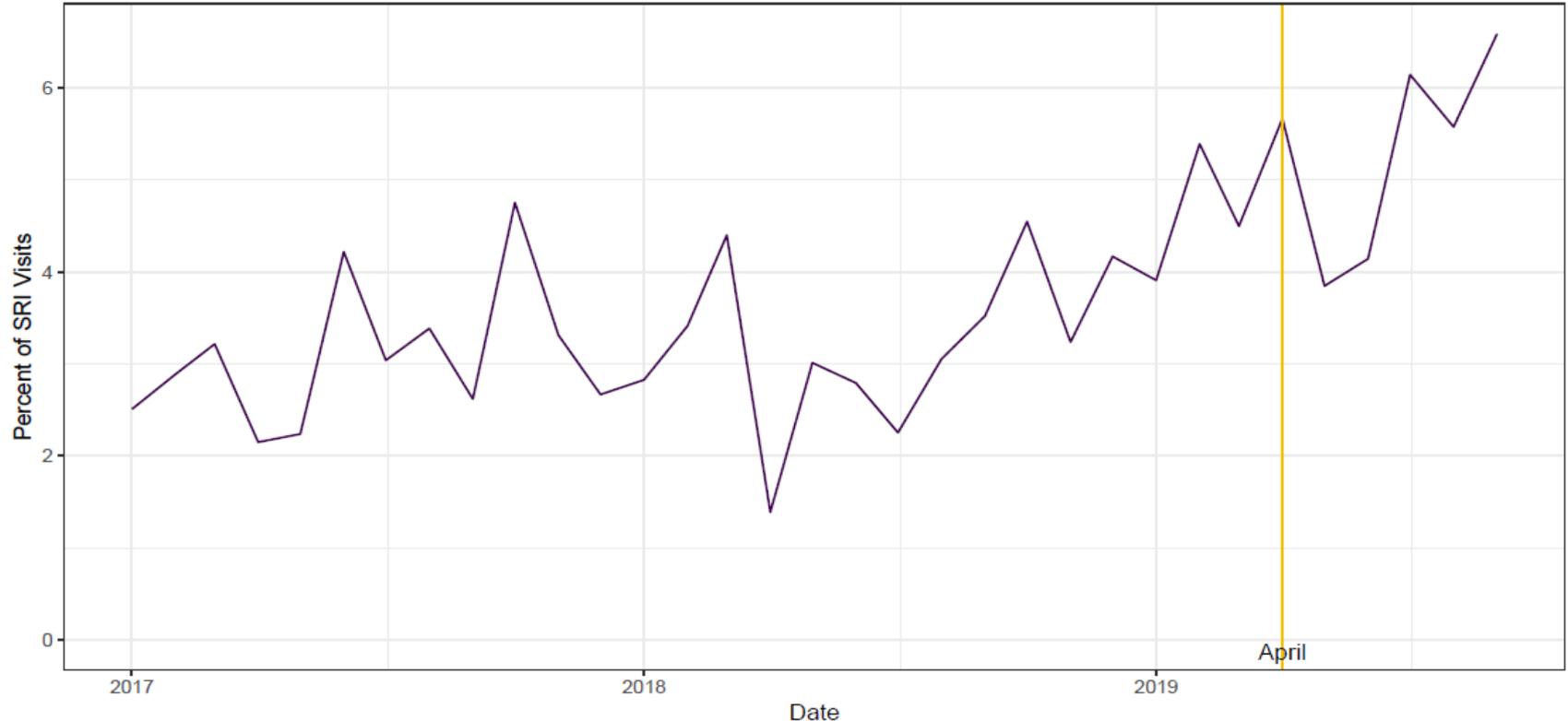

Tidylo R package:
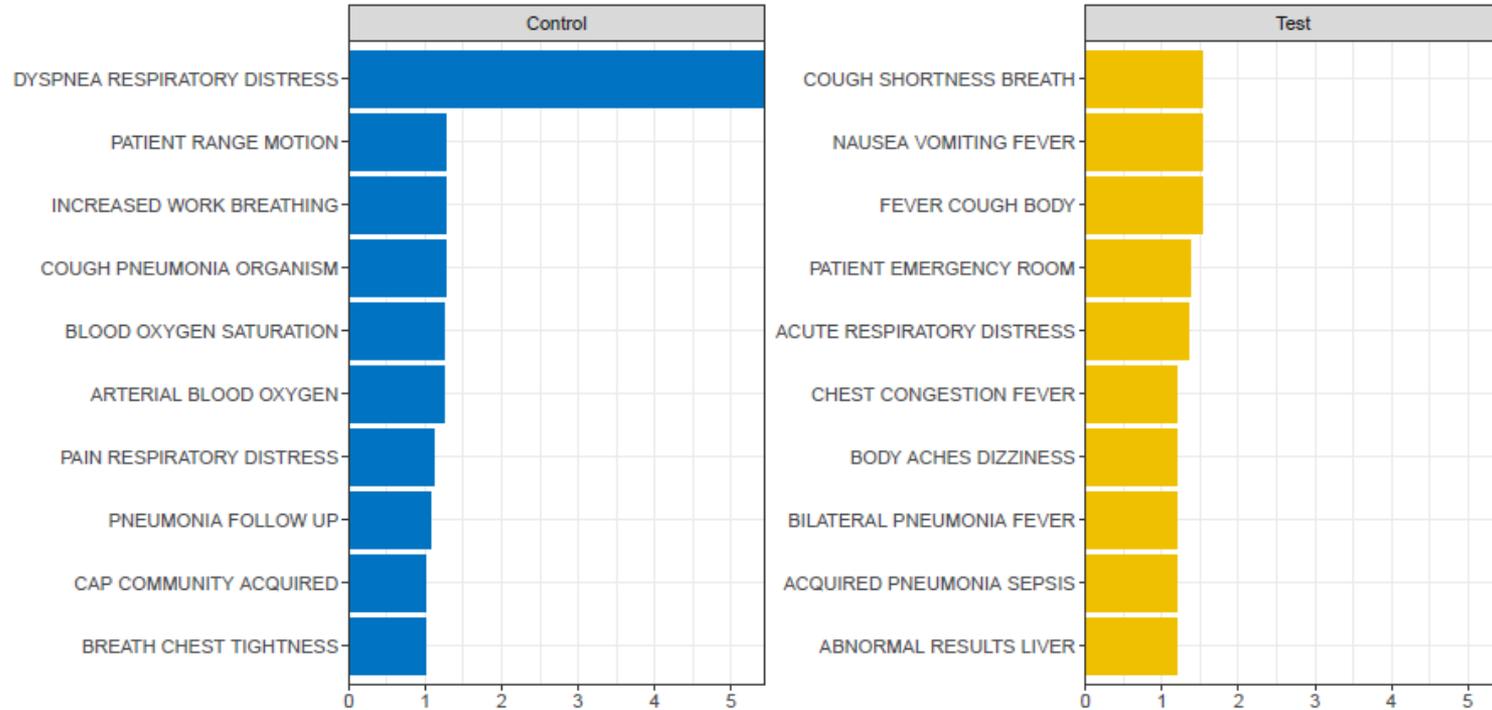https://github.com/juliasilge/tidylo

Trend for top 10 bigram chief complaint terms

■ Trend is as a percent of the lung injury (sensitive) query
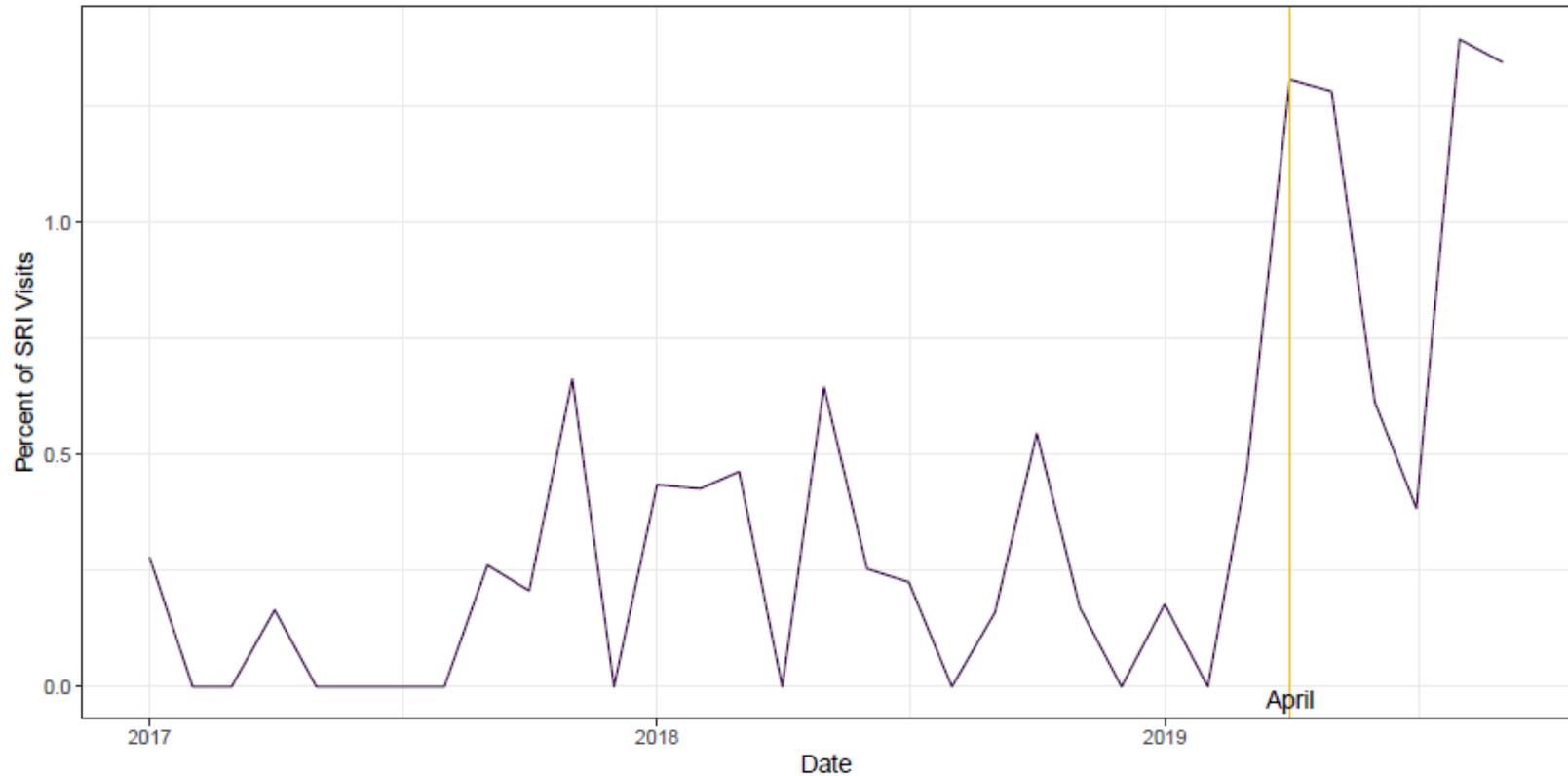
Lung injury syndrome (sensitive) –
weighted log odds ratio of top 10 chief complaint trigrams

# Trend for top 10 trigram chief complaint terms

Lung injury query – observations

- Chief complaints and ICD codes considered important in the test period included:

  - Chief Complaints: Shortness of breath, nausea/vomiting, cough shortness, low O2, bilateral pneumonia, fever sepsis

  - Discharge Diagnoses: acute resp distress, nausea and vomiting, bronchitis/pneumonitis due to chemicals, nicotine dependence (other products), lobar pneumonia, poisoning by cannabis and combinations

- Anecdotally, these are similar to diagnostic codes of actual cases

New lung injury (specific), ages 11-34 not discharged home

Percent of visits
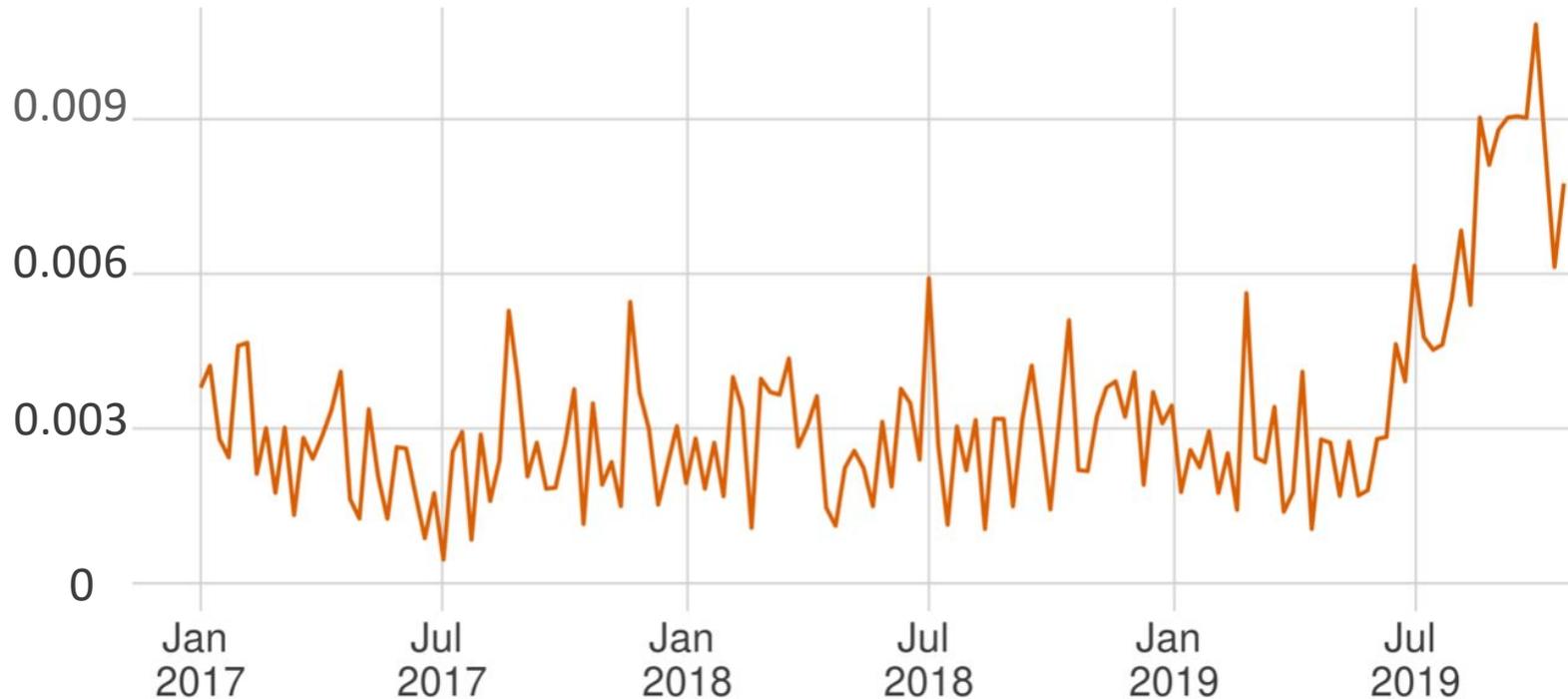
Lung injury query spatial analyses

- Methods and results are still in progress
  - Standard Morbidity Ratio (SMR) analysis.
    - Analyzing region counts / region population.
  - Pure Spatial analysis using SaTScan.
    - Analyzing geographical distribution of SRI visits.
  - Spatio-Temporal (Space-Time) analysis using SaTScan.
    - Analyzing geographical distribution within selected time parameters.

Conclusions

- ED visits with e-cigarette and vaping in the chief complaint began increasing among ages 10-19 before 2019; may be general health effects rather than signal of EVALI outbreak

- Text analysis of sensitive query for lung injury shows increase in symptoms such as nausea and vomiting associated with the EVALI outbreak in 2019

- Specific query for lung increase increased beginning in June 2019

Limitations

- Records returned by query for e-cigarette and vaping in the chief complaint not limited to lung injuries and likely strongly influenced by public awareness

- Sensitive SRI query includes unexplained pneumonia, which cannot be separated from influenza-like illness; specific version has small numbers

- Change in facilities sending data and quality of data sent may affect trend

Next Steps

- New England Journal of Medicine article about use of syndromic data

- Continuing to refine and share text mining methods

- Future spikes in trend or geographic clusters may be early signal of new risks, such as increasing exposure to dangerous e-cigarette or vaping products

For more information, contact CDC
1-800-CDC-INFO (232-4636)
TTY: 1-888-232-6348   www.cdc.gov

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.