# Estimate RSV onset and peak timing

## Tutorial 2 for transition

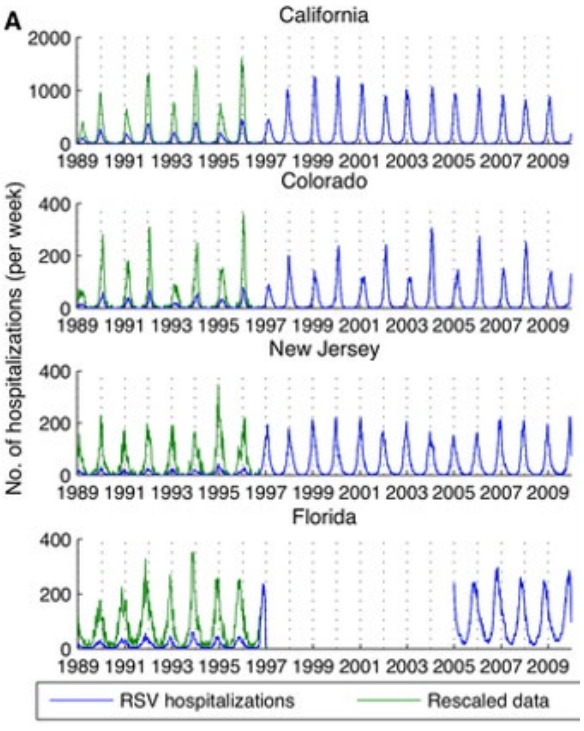
Gigi (Zhe Zheng)[1] Dan (Daniel Weinberger) Ginny (Virginia Pitzer)

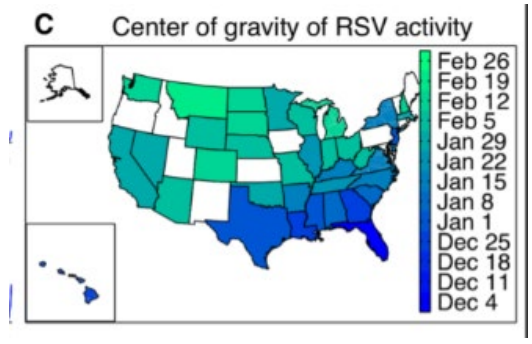
2023-01-04

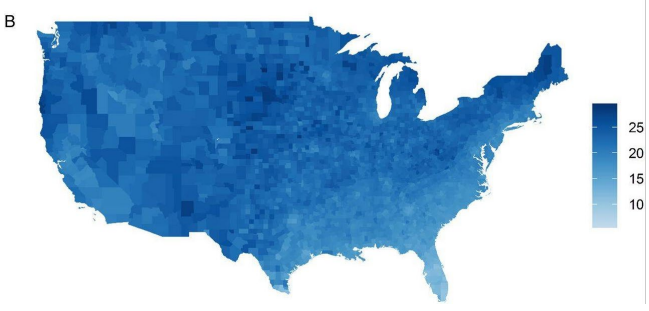
[1]zhe.zheng@yale.edu; zhe.zheng@aya.yale.edu; gigi.zhe.zheng@gmail.com

# Background

## Before the COVID-19 pandemic, the timing of RSV seasonal epidemics exhibited notable spatial patterns




Virginia E. Pitzer, 2015, PLoS Pathog
Daniel Weinberger, 2015, CID

# Outline

- In section 1, we will first introduce how to find the peak timing of RSV epidemics using harmonic regression (given regular annual/biennial seasonality). We will then talk about identifying the onset of RSV epidemics using second derivative method, regardless of the seasonality of RSV.
- In section 2, we will learn how to use R to identify peak timing and onset of RSV epidemics.
- In section 3, we will incorporate the spatial component of RSV epidemics. We will first introduce the concept of spatial autocorrelation and then lean to use R to account for spatial autocorrelation.

# References

Relevant readings:

- RSV onset timing at county level

- RSV peak timing on state level

- Comparing RSV onset timing before and during the COVID-19 pandemic

- RSV peak timing at ZIP code level and the drivers of RSV spread

# Harmonic regression to estimate the peak timing of RSV epidemics

**Note: Most of the following materials came from Dan and Ginny's Lecture notes abd Harmonic Regression by NCSS[2].**

Please check out:

- Dan's class: Public Health Surveillance
- Ginny's class: Quantitative Methods in Infectious Diseases

$$X_t = \mu + R\cos(ft + d) + e_t$$

Where
$X_t$ is the time-series contains a periodic (cyclic) component.
$\mu$ is mean of the series.
$R$ is the amplitude of seasonality.
$f = \frac{1}{period}$ is the frequency of the periodic.
$d$ is the phase or horizontal offset.
$e_t$ is the random error (noise) of the series.
$t$ is the time step

[2] https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Harmonic_Regression.pdf

# Pseudo-RSV data: Simulate time series with a 12 month period

Imagine this is RSV case data from 2 states, and we want to investigate the epidemic characteristics in these states and the lag between states.

```
set.seed(123)
  n=120 # 10 years
  t <- seq(1,n)
 amp1=2.5 # high amplitude
 freq=1/12 # frequency = 1/period
 amp2=0.5 # low amplitude

 xt1a=amp1*cos(2*3.14159*t*freq)

#other series shifted by 3 months
 xt2a=amp1*cos(2*3.14159*t*freq+1)

#Simulate some poisson count data
 xt1=rpois(n,exp((1+xt1a)))
 xt2= rpois(n,exp(2+xt2a))
```

# The observed pseudo-RSV cases over time in two states

```
plot(t, xt1,type="l", lwd=3, xaxt="n",bty="l",
     ylab="Cases of RSV", xlab="Month(n)",
     ylim=range(c(xt1,xt2)))
lines(t,xt2,col="red")
```