1-1-2016

Respiratory Syncytial Virus In Connecticut: Predictors Of Seasonal Epidemic Timing

Douglas Brian Noveroske
Yale University, dougnov@gmail.com

Respiratory Syncytial Virus in Connecticut: Predictors of Seasonal Epidemic Timing

Douglas Noveroske

Advisors: Daniel M. Weinberger, Virginia E. Pitzer, Joshua L. Warren

Yale School of Public Health

2016

ABSTRACT

Introduction: Respiratory syncytial virus (RSV) is a primary cause of hospitalizations in children worldwide. Prophylaxis (Palivizumab) can be given to infants who are at high-risk of a severe RSV infection, but the timing of seasonal RSV epidemics need to be known in order to administer prophylaxis at the appropriate time.

Methods: This study used data from the Connecticut State Inpatient Database to identify RSV hospitalizations based on ICD-9 diagnosis codes. A harmonic regression analysis was used to evaluate RSV epidemic timing at the county level; subsequently, a hierarchical model was fit to assess RSV epidemic timing at the ZIP code level. Finally, a linear regression model was used to investigate demographic characteristics that were predictive of RSV epidemic timing.

Results: 9,740 hospitalizations coded as RSV occurred among children less than 2 years old in Connecticut between July 1, 1997 and June 30, 2013. The seasonal RSV epidemic in the earliest county (Fairfield County) peaked 2.55 weeks earlier than the latest county (Tolland County). The earliest ZIP code had a seasonal RSV epidemic that peaked 4.64 weeks earlier than the latest ZIP code. Demographic characteristics that were significantly associated with ZIP code level RSV peak-timing included population density of children <5 and percent of the population that is black.

Conclusions: Seasonal RSV epidemics in Connecticut occurred earlier in areas that were more urban, had higher population density, and larger black populations. These findings could be used to better time the administration of prophylaxis to high-risk infants.

Acknowledgements: DMW is supported by grants R56AI110449 from NIH/NIAID, P30AG021342 NIH/NIA, UL1 TR000142 (NIH/CTSA), and OPP1114733 (The Bill and Melinda Gates Foundation).

TABLE OF CONTENTS

List of tables……v

List of figures……..vi

Introduction…… ….1

Methods……2

Results……..4

Discussion……8

Supplemental Information…….14

References……..15

LIST OF TABLES

Table 1: Study Population Demographics…….11

Table 2: Timing and Incidence of RSV in Connecticut Counties…….11

LIST OF FIGURES

Figure 1: Weekly RSV Hospitalizations from July 1, 1996 to June 30, 2013……...12

Figure 2: Timing of Seasonal RSV Epidemics by County……12

Figure 3: Zip Code Peak-Timing, Incidence, Population Density, & Percent Black……13

Figure 4: Correlation between Peak Timing/Incidence Rate Ratio and Population Density <5/Proportion Black……...14

INTRODUCTION

Respiratory syncytial virus (RSV) is a primary cause of hospitalizations in children worldwide [1]. Most children will have an RSV infection before 2 years of age [2]; however, certain high-risk groups are at an increased risk of a severe RSV infection [3]. Prophylaxis (Palivizumab) can be given to high-risk groups including: infants less than 12 months old at the beginning of RSV season who were born premature (earlier than 29 weeks gestation), infants with some congenital pulmonary and heart diseases, and immunocompromised infants [4]. Palivizumab has been available since 1998, and has been shown to be moderately effective in lowering the risk for RSV hospitalizations [5, 6]. Up to five doses of Palivizumab are given to provide protection for high-risk infants, and the timing of these doses needs to coincide with RSV season. Understanding the timing of RSV epidemics allows for prophylaxis to be administered at the appropriate time to provide adequate protection. Starting prophylaxis too soon is an inefficient use of resources, as the risk of infection is not high before the seasonal RSV epidemic begins; starting prophylaxis too late can put susceptible infants at risk for a severe RSV infection as they would not be protected when local RSV transmission begins. To optimize the use of resources and maximize the protection afforded to high-risk infants, it is critical to understand variations in RSV epidemic timing between locations so that prophylaxis can be appropriately timed [7].

Seasonal RSV epidemics can begin anywhere from late November to late January, and can also differ in duration [8, 9]. The timing of the epidemics varies by as much as 10 weeks across the United States, with earlier peaks in the southeastern part of the country and later peaks in the north and west; on a more regional scale, earlier epidemics have been noted in urban counties compared with rural counties [10]. It is not clear, however, how these broad national

and regional patterns of epidemic timing translate to the local scale. This study aimed to evaluate the timing and magnitude of seasonal RSV epidemics at the ZIP-code level in the state of Connecticut and to investigate the geographical and demographic characteristics that were associated with these patterns.

METHODS

Data Sources

Data consisting of all hospitalizations among children less than 2 years old in Connecticut from 1997 to 2013 were obtained from the Connecticut State Inpatient Database through the Connecticut Department of Public Health (CT-DPH). Variables included in the data were ZIP code of residence, patient age, primary insurance used for hospital admission, ICD-9 defined diagnoses (up to 10 diagnoses per hospital admission), and week, month, and year of hospital admission.

ZIP code-level data on total population, population under 5 years old, and population that is black were obtained through the United States Census Bureau [11] (by ZIP code tabulation area). These data were used to calculate population density for each ZIP code as well as population density for children less than 5 years old and percent of population that is black.

In total, there were 28,252 hospitalizations among children less than 2 years old in Connecticut between July 1, 1997 and June 30, 2013. An RSV case was defined as a hospitalization with at least one of the following documented diagnoses: 079.6 (respiratory syncytial virus), 466.11 (acute bronchiolitis due to respiratory syncytial virus), or 480.1

(pneumonia due to respiratory syncytial virus). A total of 9,740 RSV hospitalizations between July 1, 1996 and September 1, 2013 occurred that were included in the analysis. The study was approved by the Human Investigation Committees at Yale and the CT-DPH. Certain data used in this publication were obtained from the CT-DPH. The authors assume full responsibility for analyses and interpretation of these data.

County-Level estimates of timing and peak incidence

A harmonic regression analysis was first conducted to evaluate how timing of RSV epidemics differed between the eight counties in Connecticut [12]. A generalized linear model with a Poisson link was fit to the data; the data did not show evidence of overdispersion. Using the 12-month sine and cosine parameter estimates, the average peak-time of the annual epidemics and 95% confidence interval was estimated for each county [12]. Annual incidence was calculated using the mean annual number of RSV hospitalizations and the county population under 5 [(Mean annual cases / (0.4*population <5))*10,000].

ZIP code-level estimates of timing and peak incidence

A hierarchical model was used to estimate the RSV peak time and incidence rate ratio for each ZIP code. For ZIP code *i* at time *t*:

$$\text{RSV cases/population size} = \exp(\beta_0 + \beta_1 * \sin 12_t + \beta_2 * \cos 12_t + \delta_i + \eta_i * \sin 12_t + \tau_i * \cos 12_t)$$

The average peak-time of the annual RSV epidemic in each ZIP code was estimated using the 12-month sine and cosine fixed and random effects coefficients[12]. Incidence rate ratios for each ZIP code were estimated by exponentiating the random intercept (δ_i).

Correlates of peak timing at ZIP code level

We explored associations between ZIP code-level variations in peak-timing or incidence rate ratios and demographic characteristics. The demographic characteristics that were evaluated were population density of children <5 years old and proportion of the population that is black. We calculated the Pearson correlation to evaluate univariate associations and used linear regression to fit a multivariate model.

Harmonic regression and linear regression models were fit using PROC GENMOD, the hierarchical model was fit using PROC GLIMMIX, and correlation was measured using PROC CORR. The analyses were carried out using the SAS V9.4, Cary, NC.

RESULTS

Demographic and disease characteristics

Among the children less than 2 years old hospitalized for RSV, annual RSV incidence per 10,000 among children less than 6 months old (176.26 , 95% C.I. 161.14, 191.39) was more than three times the annual incidence among children 6-12 months old and more than nine times the annual incidence among children 1-2 years old (**Table 1**). The annual incidence per 10,000 of RSV was highest in ZIP codes where the population density was in the highest quintile (103.27,

95% C.I. 95.05, 111.49). Annual incidence in ZIP codes where population density was in the three lowest quintiles was less than half of the incidence observed in ZIP codes in the highest density quintile. In ZIP codes with black populations <5%, annual incidence was 43.38 cases per 10,000 (95% C.I. 39.30, 47.47). The annual incidence in ZIP codes with black populations of 20% or higher was more than twice that of incidence in ZIP codes with black populations <5% (113.83 cases per 10,000, 95% C.I. 103.84, 123.81). More than half of the RSV hospitalizations were paid for by Medicaid (53.03%), a proxy for low income.

Annual RSV incidence per 10,000 in New Haven County (116.30, 95% C.I. 105.84-126.76) was almost double the incidence in any other county. After New Haven County, annual incidence was next highest in Fairfield County (68.82 per 10,000, 95% C.I. 61.94-75.69), which along with New Haven County comprise two of the three most populous counties in Connecticut (out of eight). Middlesex County (50.07, 95% C.I. 44.04-56.09) and Tolland County (28.13, 95% C.I. 21.27-34.99) had the lowest annual RSV incidence per 10,000 people.

Statewide, 9,740 hospitalizations coded as RSV occurred among children less than 2 years old between July 1, 1997 and June 30, 2013. The number of weekly RSV hospitalizations varied from 0 to 73 (**Figure 1**). Seasonal RSV epidemics peaked in late December/early January, and the amplitude of the yearly peak varied by year. The number of RSV cases in a July – June season ranged from 461 hospitalizations in the July 2008 – June 2009 RSV season to 764 hospitalizations in the July 2006 - June 2007 RSV season. The average number of RSV hospitalizations in an epidemiological year was 608.75 (s.d. = 95.10 hospitalizations).

Variation in epidemic timing by county

The seasonal RSV epidemic peaked earliest in Fairfield county (28.71 weeks after July 1, 95% C.I. 28.46 – 28.95 weeks after July 1), 2.55 weeks earlier than the seasonal RSV peak in Tolland County (31.26 weeks after July 1, 95% C.I. 30.33 – 32.18 weeks after July 1) (**Table 2, Figure 2**). The three counties with the largest populations (Fairfield, Hartford, and New Haven Counties) had seasonal RSV epidemics that all peaked earlier than the other five, smaller, counties (Litchfield, New London, Windham, Tolland, and Middlesex Counties). Two of the three counties with the earliest peak-timing also had the highest RSV incidence (New Haven and Fairfield Counties).

Variation in epidemic timing and incidence by ZIP code

Across all ZIP codes, the average seasonal RSV peak happened 29.68 weeks after July 1 (**Figure 3**). The seasonal peaks among ZIP codes ranged from 27.07 weeks after July 1 to 31.71 weeks after July 1. Large differences in RSV incidence between zip codes were also present, as measured by an incidence rate ratio comparing incidence in each ZIP code to the average incidence for the state of Connecticut. The incidence rate ratio varied from 0.18 to 5.18, reflecting rates that were 5-fold above or below average (**Figure 3 C, D**). Seasonal peak timing and incidence rate ratio were also correlated (r = -0.35, 95% C.I. -0.46, -0.24).

Correlates of ZIP code-level timing

Seasonal RSV epidemics that peaked earlier were associated with higher population density <5 (r = -0.45, 95% C.I. -0.54, -0.34) and higher proportion of the population that is black (r = -0.46, 95% C.I. -0.55, -0.36) (**Figure 4**).

A similar relationship was seen with incidence rate ratios. ZIP codes with higher RSV incidence were associated with higher population density <5 5 (r = 0.46, 95% C.I. 0.46, 0.55) and a larger black populations (r = 0.58, 95% C.I. 0.49, 0.65).

Finally, a multivariate model was used to evaluate these relationships (**Table 3**). Population density <5 (estimate = -0.11, 95% C.I. -0.18, -0.04), and proportion black (estimate = -0.13, 95% C.I. -0.20, -0.06) were independently associated with peak timing. A 1-log increase in population density < 5 was associated with a peak timing of 0.12 weeks earlier; a 1-log increase in proportion black was associated with the seasonal RSV epidemic peaking 0.13 weeks earlier.

Significant predictors of ZIP code incidence were also density of children <5 (estimate = 0.07, 95% C.I. 0.01, 0.12) and proportion black (estimate = 0.18, 95% C.I. 0.12, 0.23). A 1-log increase in density of children <5 was associated with an increase in the log RSV incidence rate ratio of 0.07. Independently, a 1-log increase in proportion black was associated with an increase in log incidence rate ratio of 0.18.

DISCUSSION

Using a comprehensive statewide hospitalization database from Connecticut, we detected significant local variations in the timing of seasonal RSV epidemics, with more than 4.5 weeks between epidemic peaks in the earliest and latest ZIP codes. The ZIP codes where RSV epidemics peaked earlier tended to be larger, more urban locations with higher population density and larger black populations. These variations should be taken into account to ensure that prophylaxis for children at-risk of a severe RSV infection is appropriately timed with the local epidemiology of RSV.

Other studies have examined RSV epidemics and factors that are related with higher RSV risk; risk factors associated with RSV have included sharing bedrooms and having more than 4 children in a house[13]. These measures of crowding could have a similar influence on RSV incidence as population density and population density of children <5. The R_0 of RSV has also been shown to be significantly associated with population density [14], which could help explain the finding of seasonal RSV epidemics happening earlier in urban areas. Urban environments may have higher incidence and earlier peak-timing due to the fact that individuals who live in urban places have contact with more people. This could account for increased transmission which happens quicker, and seasonal RSV epidemics are able to establish earlier in urban areas than in less-densely populated rural areas where individuals may contact others less often.

In order to appropriately provide RSV prophylaxis, it is important to understand when local RSV season begins as well as how long the season lasts. Our findings on the ZIP code-level in Connecticut build on previous findings on the county-level which showed an earlier optimal time to begin RSV prophylaxis that was associated with urbanization, percent black, and higher

population density [10]. Drastic differences in demographic characteristics can exist within counties, so understanding epidemic timing at the ZIP code-level can better guide the optimal time to begin RSV prophylaxis.

In addition to ZIP code-level differences in RSV timing, high variability in the duration of RSV seasons at the ZIP code level has also been observed [9]. Longer RSV seasons have been associated with urban areas, crowding, and percentage of children <5. Analogous mechanisms associated with urban environments and high population density (increased contact with higher number of individuals, situations in which people live closer to each other) could help explain the similarities observed in RSV epidemic timing and epidemic duration. Insight into the relationship between RSV epidemic timing and duration could provide even better guidance into the optimal use of RSV prophylaxis.

This analysis has both strengths and weaknesses. Through the use of data from the Connecticut State Inpatient Database, this study captured 100% of hospitalizations in Connecticut. This allowed for the analysis of all hospitalizations of children <2 years old during this time period. Additionally, a large amount of data was available to conduct the analysis. Nearly 10,000 RSV hospitalizations throughout 16 seasonal epidemics provided a large data pool to analyze the timing of, and factors associated with, RSV epidemics throughout the state of Connecticut on the county and ZIP code level.

There are also weaknesses with the data used and analysis conducted. Since hospitalization data was used, we were reliant on ICD-9 coding. This may lead to some misclassification of RSV cases, as data on viral tests were not available for analysis. Without laboratory testing, other respiratory illnesses such as influenza could have been incorrectly coded

as RSV. Furthermore, RSV cases may not have been correctly coded as such and were therefor not included in the analysis. Other studies, however, have found good concordance between ICD-coded RSV and laboratory [14]. Additionally, this analysis only evaluated incidence and peak-timing of seasonal RSV epidemics. Since children who are at high-risk of a severe RSV infection need to be protected throughout the duration of seasonal RSV epidemics, understanding the duration of seasonal epidemics is also important to effectively protect high-risk children.

Understanding when seasonal RSV epidemics happen is critical in providing adequate prophylaxis to children who are at high-risk of a severe RSV infection. This analysis can be used to guide policy decisions for when to provide prophylaxis to children who live in different parts of Connecticut. Further studies that investigate how local RSV timing differs in ZIP codes throughout the United States should be used to appropriately provide prophylaxis based on local RSV epidemiology. By better understanding the factors that predict seasonal RSV epidemic patterns, better, more cost-effective, protection can be provided to children who need it.

Table 1: Study Population Demographics*

	Annual Incidence per 10,000 (95% C.I.) or Number (percent)
Age	
<6 months	176.26 (161.14, 191.39)
6-12 months	51.71 (46.09, 57.32)
13-24 months	19.18 (16.99, 21.37)
Percent Black	
<5%	43.38 (39.30, 47.47)
5-10%	67.00 (59.57, 74.43)
10-20%	97.78 (87.32, 108.25)
20% or more	113.83 (103.84, 123.81)
Population Density	
Quintile 1 (lowest)	39.83 (31.46, 48.20)
Quintile 2	34.70 (26.79, 42.61)
Quintile 3	39.00 (33.88, 44.12)
Quintile 4	45.61 (41.25, 49.98)
Quintile 5 (highest)	103.27 (95.05, 111.49)
Insurance	
Commercial	4,304 (44.19)
Medicaid	5,165 (53.03)
Medicare	2 (0.02)
Other Public	95 (0.98)
Uninsured	174 (1.79)

*Numbers may not sum to total due to missing data, and percentages may not sum to 100% due to rounding

Table 2: Timing and Incidence of RSV in Connecticut Counties

	Peak timing (weeks since July, 1), (95% CI)	Incidence (Annual RSV Hospitalizations per 10,000), (95% CI)
Peak Timing		
Fairfield	28.71 (28.46, 28.95)	68.82 (61.94, 75.69)
New Haven	29.09 (28.90, 29.27)	116.30 (105.84, 126.76)
Hartford	29.09 (28.83, 29.35)	60.73 (54.59, 66.87)
Litchfield	30.12 (29.60, 30.64)	61.10 (48.53, 73.68)
New London	30.70 (30.30, 31.10)	65.26 (56.64, 73.87)
Middlesex	31.13(30.47, 31.78)	50.07 (44.04, 56.09)
Windham	31.18 (30.52, 31.84)	64.57 (54.03, 75.10)
Tolland	31.26 (30.33, 32.18)	28.13 (21.27, 34.99)

Figure 1: Weekly RSV Hospitalizations from July 1, 1996 to June 30, 2013

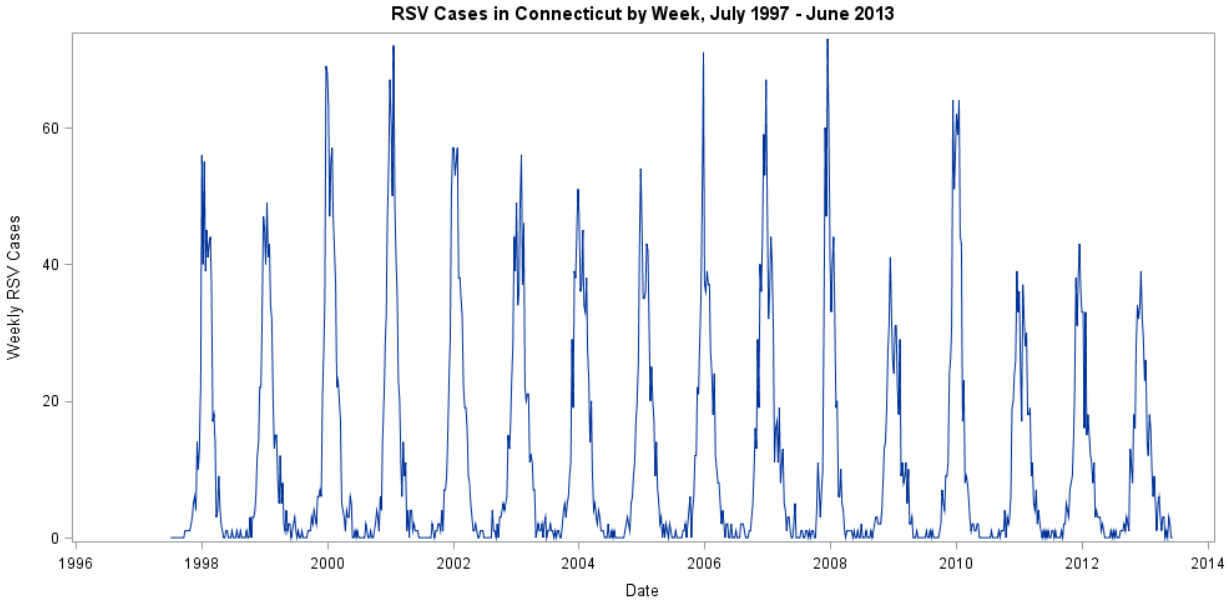


Figure 2: Timing of Seasonal RSV Epidemics by County

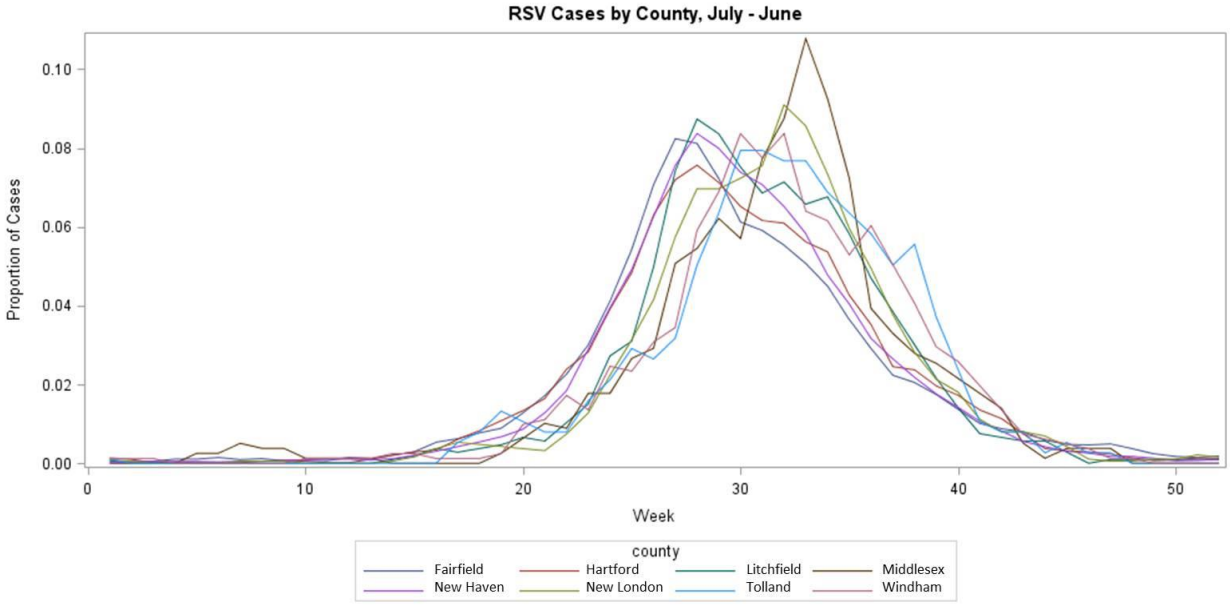


Figure 3: Zip Code Peak-Timing, Incidence, Population Density, & Percent Black

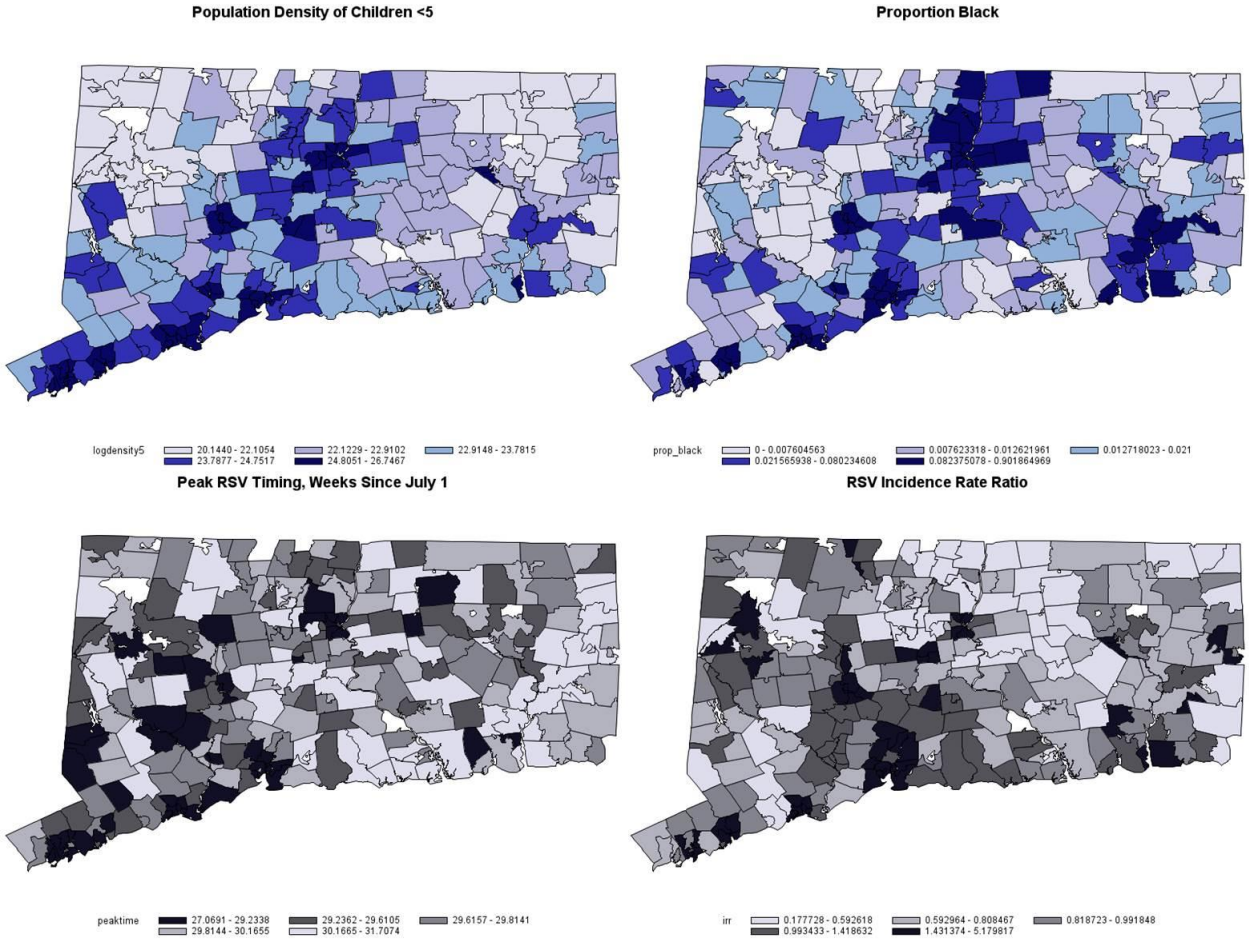
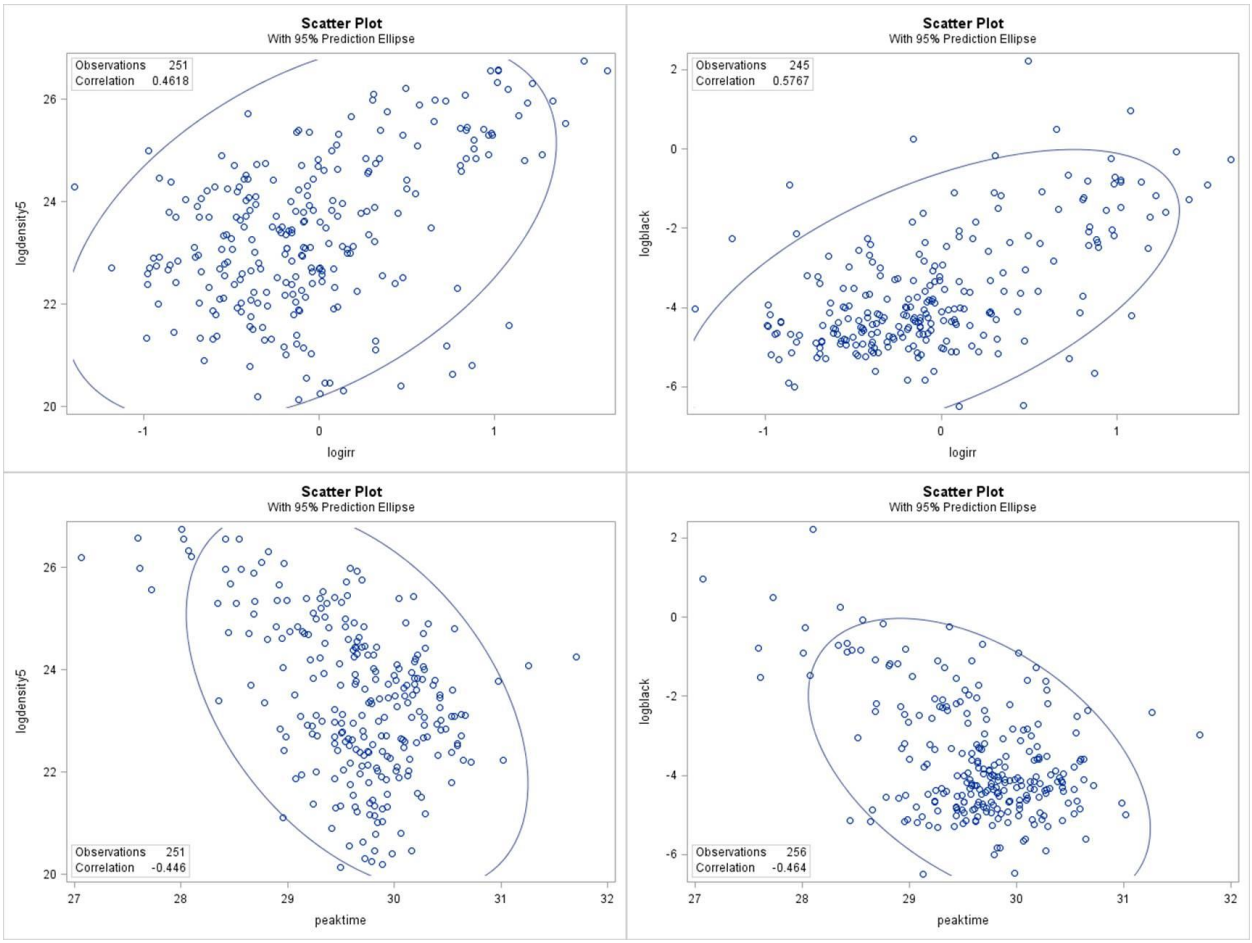


Figure 4: Correlation Between Peak Timing/Incidence Rate Ratio and Population Density <5/Proportion Black



SUPPLEMENTAL INFORMATION: SAS code for the hierarchical model

```
proc glimmix data = ds1 method = laplace;
class patzip;
model weekcase =  sin12 cos12 / dist = poisson link = log offset = log5 s;
random intercept sin12 cos12 / subject = patzip s;
output out = modelzippop pred = predzippop;
ods output parameterestimates = fixedeffectspop solutionr = randomeffectspop;
run; quit;
```

REFERENCES

1. Nair, H., Nokes DJ, Gessner BD, et al., *Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis.* The Lancet, 2010. **375**: p. 1545-1555.
2. Kneyber, M., Steyerberg EW, de Groot R, Moll HA, *Long-term effects of respiratory syncytial virus (RSV) bronchiolitis in infants and young children: a quantitative review.* Acta Paediatrica, 2000. **89**: p. 654-660.
3. RC, W., *Review of Epidemiology and Clinical Risk Factors for Severe Respiratory Syncytial Virus (RSV) Infections.* The Journal of Pediatrics, 2003. **143**: p. S112-S117.
4. *Updated Guidance for Palivizumab Prophylaxis Among Infants and Young Children at Increased Risk of Hospitalization for Respiratory Syncytial Virus Infection.* Pediatrics, 2014. **134**(2): p. 415-420.
5. Pedraz, C., Carbonell-Estrany X, Figueras-Aloy J, Quero J, The IRIS Study Group, *Effect of palivizumab prophylaxis in decreasing respiratory syncytial virus hospitalizations in premature infants.* The Pediatric Infectious Disease Journal, 2003. **22**(9): p. 823-827.
6. Feltes, T., Cabalka A, Meissner C, Piazza F, Carlin D, Top F, Connor E, Sondheimer H, *Palivizumab Prophylaxis Reduces Hospitalization Due to Respiratory Syncytial Virus in Young Children with Hemodynamically Significant Congenital Heart Disease.* The Journal of Pediatrics, 2003. **143**: p. 532-540.
7. Gutfraind, A., Galvani AP, Meyers LA, *Efficacy and Optimization of Palivizumab Injection Regimens Against Respiratory Syncytial Virus Infection.* JAMA Pediatrics, 2015. **169**(4): p. 341-348.
8. Mullins, J., Lamonte A, Bresee J, Anderson L, *Substantial variability in community respiratory syncytial virus season timing.* The Pediatric Infectious Disease Journal, 2003. **22**(10): p. 857-862.
9. Zachariah, P., Shah S, Gao D, Simoes EAF, *Predictors of the Duration of the Respiratory Syncytial Virus Season.* The Pediatric Infectious Disease Journal, 2009. **28**(9): p. 772-776.
10. Weinberger, D., Warren J, Steiner C, Charu V, Viboud C, Pitzer V, *Reduced-Dose Schedule of Prophylaxis Based on Local Data Provides Near-Optimal Protection Against Respiratory Syncytial Virus.* Clinical Infectious Diseases, 2015. **61**(4): p. 506-514.
11. *Age groups and sex: 2010*. United States Census Bureau: American FactFinder.
12. Lofgren, E., Wenger JB, Fefferman NH, Bina D, Gradus S, Bhattacharyya S, Naumov YN, Gorski J, Naumova EN, *Disproportional effects in populations of concern for pandemic influenza: insights from seasonal epidemics in Wisconsin, 1967–2004.* Influenza and Other Respiratory Viruses, 2010. **4**: p. 205-212.
13. Simoes, E., *Environmental and Demographic Risk Factors for Respiratory Syncytial Virus Lower Respiratory Tract Disease.* The Journal of Pediatrics, 2003. **143**: p. S118-S126.
14. Pitzer, V., Viboud C, Alonso WJ, Wilcox T, Metcalf CJ, Steiner CA, Haynes AK, Grenfell BT, *Environmental Drivers of the Spatiotemporal Dynamics of Respiratory Syncytial Virus in the United States.* PLos Pathogens, 2015. **11**(1): p. e1004591.