



# HHS Public Access

Author manuscript

*Pediatr Pulmonol.* Author manuscript; available in PMC 2017 June 01.

Published in final edited form as:

*Pediatr Pulmonol.* 2016 June ; 51(6): 633–642. doi:10.1002/ppul.23332.

## Spatial Clusters of Child Lower Respiratory Illnesses Associated with Community-Level Risk Factors

Paloma I. Beamer, Ph.D.<sup>1,2,3</sup>, Nathan Lothrop, M.S.<sup>2</sup>, Zhenqiang Lu, Ph.D.<sup>3</sup>, Rebecca Ascher, M.S.<sup>2</sup>, Kacey Ernst, Ph.D.<sup>2</sup>, Debra A. Stern, M.S.<sup>1</sup>, Dean Billheimer, Ph.D.<sup>2,3</sup>, Anne L. Wright, Ph.D.<sup>1</sup>, and Fernando D. Martinez, M.D.<sup>1,3</sup>

<sup>1</sup>Arizona Respiratory Center, University of Arizona

<sup>2</sup>Mel and Enid Zuckerman College of Public Health, University of Arizona

<sup>3</sup>Bio5 Institute, University of Arizona

### Abstract

Identifying geographic areas with increased incidence of disease may elucidate community-level risk factors for intervention development. Lower respiratory illnesses (LRIs) are the leading cause of death in children and are associated with other morbidities. We assessed geographic clustering of LRIs and evaluated if these spatial patterns and associated risk factors differed by phenotype.

Participants enrolled at birth in the Tucson Children's Respiratory Study were followed through age three for physician diagnosed LRIs. Spatial clustering analysis, based upon each participant's birth address, was performed for four LRI phenotypes. We conducted principal component analysis at the census tract level to generate indices for lower socioeconomic status (SES), poorer housing conditions, and increased air pollution.

Enrollment addresses were mapped for 812 subjects, of whom 58.4%, 33.5%, 34.2% and 23.4% had any LRI, a wheezing LRI, a viral LRI and a respiratory syncytial virus (RSV) LRI, respectively. Patterns of spatial clustering and associated risk factors differed by LRI phenotype. Multivariable regression analyses showed that wheezing LRI clusters were associated with increased air pollution (OR=1.18, p=0.01). Being in a viral cluster was associated with poorer housing conditions (OR=1.28, p=0.01), while being in a RSV cluster was associated with increased air pollution (OR=1.14, p=0.006), poorer housing conditions (OR=1.54, p=0.003), and higher SES (OR=0.77, p=0.001).

Our use of social and environmental indices allowed us to identify broad contextual factors that may contribute to increased incidence of LRIs in specific geographic regions. To reduce LRI incidence, multifaceted interventions should be developed at the community level.

---

Corresponding Author: Paloma Beamer, Ph.D., Associate Professor, Mel and Enid Zuckerman College of Public Health, University of Arizona, 1295 N. Martin Ave., PO 245210, Tucson, AZ 85724. Telephone: 520-626-0006; Fax: 520-626-8009; pbeamer@email.arizona.edu.

This manuscript was presented in part at the American Thoracic Society in San Francisco, CA (2012).

**Conflict of Interest:** P.I.B. received support from Kimberly Clark. F.D.M. received support from Abbott and Merck. This support was outside of the submitted work and the authors certify that their freedom to design, conduct, interpret, and publish this analysis was not compromised by any sponsor.

## Keywords

geographic cluster; socioeconomic status; housing conditions; air pollution; neighborhood

---

## Introduction

Acute lower respiratory illnesses (LRIs) are a leading cause of emergency department visits, hospitalizations, and death in children<sup>1, 2</sup> and are associated with subsequent development of asthma<sup>3, 4</sup>. The Tucson Children's Respiratory Study (CRS), a longitudinal birth cohort, was established in 1980 to study risk factors for acute LRIs in the first three years of life, and the association of these illnesses with subsequent respiratory disease<sup>5, 6</sup>. Prospective analyses from this cohort have demonstrated that there are distinct LRI phenotypes in early life (<3 years) that are associated with different future respiratory outcomes<sup>7</sup>. For example, physician confirmed wheezing LRIs are associated with diminished lung function and recurrent cough<sup>3, 8, 9</sup>. At the time of each LRI in this cohort, virus cultures were conducted. In particular, early-life respiratory syncytial virus (RSV) LRIs (independent of wheeze status) in this cohort are associated with wheeze and bronchodilator responsiveness later in childhood<sup>10</sup> and with active asthma in adulthood among smokers<sup>11</sup>.

In our cohort, these distinct clinical presentations are associated with different risk factors. Increased incidence of wheezing LRIs is associated with parental smoking, evaporative cooling, having older siblings, younger maternal age and day care attendance<sup>12-15</sup>. Similarly, RSV LRI incidence is associated with more than 1 person sharing the child's bedroom and negatively associated with breastfeeding among children born to mothers with less education<sup>16</sup>. While these associations are important for identifying high risk children in clinical settings, discovering specific geographic regions with increased LRI incidence provides an opportunity to identify community-level risk factors that can be targeted for public health interventions.

Spatial analysis can identify geographic regions with significantly higher incidence of disease or a "cluster" of diseased individuals grouped nearer to each other in space that has a low probability to have occurred by chance<sup>17</sup>. Most studies that have identified LRI spatial clusters have focused on disease surveillance, relying on medical or demographic records<sup>18-20</sup>. However, these types of studies are inherently biased and may be more associated with differential health care seeking behavior and subsequent disease reporting than actual higher disease incidence<sup>21</sup>. In previous studies, associations were examined between LRI clusters and either environmental or social risk factors but not together.

The objectives of the current study were to identify LRI spatial clusters in the CRS<sup>5</sup>, a well characterized birth cohort where LRI incidence was prospectively and clinically documented, and to determine if these spatial patterns and associated risk factors differed by LRI phenotype. Given the prevalence and burden of childhood LRIs<sup>2-4</sup> determining their spatial distribution in relation to social and environmental risk factors could help elucidate potential community-level interventions in early life.

## Materials and Methods

### Study Population

The CRS is an unselected longitudinal birth cohort recruited between 1980 and 1984. Healthy infants born to mothers that were patients of the largest health maintenance organization in Tucson, AZ, were eligible for recruitment. A study nurse contacted the 1,596 eligible families, of which 1,246 infants (78%) were enrolled. A detailed description of the CRS cohort design, enrollment, and data collection has been published<sup>5,6</sup>. Informed consent was obtained from each child's parent. The study, including the spatial cluster analysis, was approved by the University of Arizona Human Subjects Protection Program.

### Lower Respiratory Tract Illnesses

During the first 3 years of life participants had extensive evaluations of every LRI (see Online Supplement)<sup>5,6</sup>. At the time of a LRI, study physicians completed a standardized form on the presence of symptoms including wheeze. Nasopharyngeal and throat swabs for virus culture were also taken.

The incidence of any LRI, any wheezing LRI (i.e., 1 LRI where wheeze was documented by physician exam), any viral LRI (i.e., 1 LRI with positive viral culture), and any RSV LRI (i.e., 1 LRI with positive RSV culture) was computed as a dichotomous variable. Children were considered cases if they 1 LRI episode, but were not counted multiple times for additional episodes. These LRI phenotype categories were not mutually exclusive. Thus, if a child had a wheezing RSV LRI they would be included as a case in all three. There were 113 children who either did not have a wheezing status recorded or who only had wheeze from patient history but not by exam, and there were 107 children that were missing virus culture results that were excluded from the wheezing and viral outcomes, respectively.

### Cluster Detection

Participants' enrollment addresses were geocoded in ArcGIS 10 (ESRI, Redlands, CA). We used the "SPATCLUS" package in R 2.10.0 (R Foundation for Statistical Computing, Vienna, Austria) for detecting multiple arbitrarily-shaped spatial clusters for case event data<sup>17,22</sup>. Briefly, one address is randomly selected and the distance to the nearest address (i.e., nearest neighbor) is computed. Subsequently, the distance from that second address to the next nearest neighbor (that does not include the previous point) is computed. This continues until all addresses have been assigned a nearest neighbor distance and a selection order (i.e., the order in which they were assigned the nearest neighbor distance). These distances were weighted by their expected distance under a uniform distribution to account for potential inhomogeneity in the underlying population density. The clusters were detected by modeling multiple structural changes of the distances on the selection order, with the best model selected with the double maximum test<sup>23</sup>. Since point density is higher within a cluster, the points in the cluster should have consecutive selection orders and shorter associated nearest-neighbor distances than points outside the cluster. Results are not affected by selection of the initial address<sup>17,22</sup>. Statistical significance of the clusters is determined through Monte Carlo simulations of the coordinates for the total study population.

We completed the spatial cluster analysis for each LRI phenotype (i.e., any LRI, any wheezing LRI, any viral LRI, any RSV LRI). Cases had 1 LRI corresponding to the phenotype. After identifying clusters for each LRI phenotype, we computed four dichotomous variables, whereby for each phenotype: 1=case and in a cluster; and 0=case and not in a cluster or not a case.

### Risk Factor Analysis

Socioeconomic, housing characteristic, and air pollution data were obtained to determine if community-level risk factors were associated with LRI clusters. Data from the 1980 United States Census were acquired from the National Historical Geographic Information System<sup>24</sup>. A description of the socioeconomic and housing characteristic variables is provided in E-Table 1. We obtained estimated average air pollutant exposure concentrations for each census tract from the United States Environmental Protection Agency (US EPA) National-scale Air Toxics Assessment (NATA) database<sup>25,26</sup> (see Online Supplement). We used the 1996 database, which provides information for 34 pollutants, because this is the earliest year available. We mapped each of the risk factor variables by census tract and overlaid the enrollment addresses to obtain values for each participant.

For each LRI phenotype, we conducted simple logistic regression analyses to assess the relationship between having an LRI and being in a corresponding LRI phenotype cluster and each of the socioeconomic, housing characteristic, and air pollution variables. As many of these variables are highly correlated, we conducted separate principal component analyses (PCA) for categories of geographic risk factors<sup>27</sup>: socioeconomic status (SES), housing characteristics, and air pollution (see Online Supplement). All variables were standardized prior to the PCA. We used the value of the first principal component as an index for SES, housing characteristics, and air pollution levels, respectively. We assessed if these indices were associated with being in clusters of each LRI phenotype using simple and multivariable logistic regression.

STATA 12.0 (StataCorp, College Station, TX) was used for risk factor analyses. We used the robust estimator of variance with the cluster option for multi-level modeling by census tract. An alpha level of 0.05 was used for statistical significance. A Bonferroni correction was used to account for multiple comparisons within each of the geographic risk factor categories.

### Results

Of the 1246 infants, 888 (71%) had complete follow-up through age three and enrollment addresses were mapped for 812 (Figure 1). SES, housing characteristics, and air pollution variable summaries for participants are provided in Table 1. Participants resided in 95% Tucson-area census tracts, and were representative of the overall region (E-Tables 2 and 3). Underrepresented census tracts had a high density of senior citizen communities or consisted of Native American nations. Thus, they had a higher median age and percentage of minority race and lower population density and air pollution levels than the overall region.

## Geographic Distribution of Clustering

Different patterns of significant spatial clusters ( $p < 0.05$ ) were detected for each LRI phenotype (Figure 1). A complete tabulation of each cluster is provided in E-Table 4. Of the 474 children with 1 LRI, 84% were in one large cluster, and further analyses were not meaningful (data not presented). Four unique clusters were identified for wheezing LRIs (Figure 1), which bordered Tucson's two highways. Four unique clusters were identified for viral LRIs in Tucson: two central, one eastern, and one southern. Three unique clusters were detected for the RSV LRI phenotype. One spanned the main highway, another was central, and the third was located in eastern Tucson. These clusters each contained 26 to 40 children.

## Simple Analyses of Census Tract and Air Pollution Risk Factors

We assessed if specific census tract risk factors and air pollutant concentrations were associated with being in a LRI cluster by phenotype. Following the Bonferroni correction none of these variables were commonly associated with clusters for all three remaining LRI phenotypes (Tables 2 and 3). Being in a wheezing LRI cluster was associated with increased exposure to 6 (25%) of the air pollutants following Bonferroni correction (Table 3). Participants were less likely to be in a wheezing LRI cluster if they lived in a census tract with a higher median age (Table 2). Conversely, being in a viral LRI cluster was associated only with living in a census tract with a higher proportion of homes that use gas for household or water heating, and none of the SES or air pollutant variables following Bonferroni correction.

Being in a RSV LRI cluster was associated with multiple housing characteristics and multiple air pollutants even after Bonferroni correction. Participants in an RSV LRI cluster were more likely to live in a census tract with a greater proportion of homes that were attached, or used gas heating (Table 2). Being in a RSV LRI cluster was associated with increased exposure to 10 (42%) air pollutants (Table 3), all of which were unique to this LRI phenotype.

## Principal Component Analyses

We completed three separate PCAs to develop an index for each risk factor category (Table 1). All of our socioeconomic variables were positively correlated with their first principal component, which represented 56% of the variance and was interpreted as an index of lower SES. Similarly, all of our housing characteristic variables were positively correlated with their first principal component, which accounted for 36% of the variance, and was interpreted as an index of poorer housing conditions. Six of the 30 air pollutants demonstrated a modest negative correlation with their first principal component. We excluded these air pollutants from our analyses, and the remaining 24 pollutants were all positively correlated with their first principal component, which accounted for 45% of the variance and was considered to be an index of increased air pollution. Maps of the indices are provided in Figure 2.

## Associations between Risk Factors Indices and Cluster by LRI phenotype

Each LRI phenotype was associated with different risk factor indices (Table 4). Being in a wheezing LRI cluster was associated with the increased air pollution index in both simple

and multivariable regression analyses. Although in the simple analyses, being in a viral LRI cluster was associated with both the lower SES and poorer housing conditions indices, only the association with poorer housing remained significant in the multivariable regression model. Participants in a RSV LRI cluster were more likely to live in a census tract with poorer housing conditions and increased air pollution, but less likely to live in a census tract with lower SES after controlling for housing conditions and air pollution.

## Discussion

To our knowledge this is the first study to examine the spatial distribution of physician-ascertained LRIs in early life with viral culture and wheeze status information. Our results demonstrate that LRI phenotypes cluster in different spatial patterns and that social and environmental determinants of these clusters differ by phenotype. Wheezing LRI clusters were associated with increased air pollution, while viral LRI clusters were associated with poorer housing conditions. RSV LRI clusters were associated with poorer housing, increased air pollution, and higher SES. Previous studies have only assessed environmental or social risk factors separately, but not in combination, despite these factors often being correlated<sup>20, 28, 29</sup>. Our results have implications for targeting interventions, such as RSV vaccination, for high-risk children in specific geographic regions.

Many studies have demonstrated an association between lower SES and increased LRI incidence<sup>15, 29–31</sup>. There are multiple pathways by which SES may affect LRI incidence, including: increased exposure to infectious agents from household crowding; increased susceptibility because of a weakened immune system from high stress or a poor diet; or greater air pollution exposure<sup>30</sup>. Understanding the relative impact of each of these factors is difficult as most studies only examine a specific pathway<sup>1, 29</sup>. Furthermore, these pathways are often intertwined, which makes disentangling their relative impacts difficult.

Messer et al. demonstrated that a neighborhood deprivation index derived from PCA was more useful than individual census variables for assessing neighborhood effects on health outcomes because of high correlations between census variables and the multi-dimensionality of constructs like neighborhood deprivation<sup>27</sup>. Similarly, we used PCA to develop three distinct indices to represent lower SES, poorer housing conditions, and increased air pollution. Although our indices include many factors, we separated variables associated with the physical environment from those purely demographic or economic to determine the relative impact of SES, housing quality, and air pollution on childhood LRIs incidence.

We demonstrated that being in a viral LRI cluster was associated with poorer housing conditions independent of SES (Table 4). We also demonstrated that being in a RSV LRI cluster was positively associated with poorer housing conditions but negatively associated with lower SES. Although other studies have indicated associations of RSV LRIs with SES and poorer housing separately<sup>31</sup>, these predictors were assessed separately. In our study, the poorer housing conditions index was partly based on the proportion of attached dwellings, houses lacking complete plumbing facilities and household crowding in the census tract. Our

results suggest that a primary driver of viral and RSV transmission is higher exposure to infectious agents through crowding and unsanitary conditions in these geographic regions.

Increased air pollution was associated with wheezing and RSV LRI clusters (Table 4), which confirms previously reported associations<sup>32,33</sup>. Wheezing LRIs have also been associated with lower SES<sup>16,34,35</sup>, but these studies did not assess air pollution. Our analysis uniquely demonstrates that the association of air pollution with wheezing LRIs may be independent of SES and housing characteristics. Exposure to air pollution has been associated with increased hospital admissions for RSV bronchiolitis<sup>32</sup> and with increased viral receptors in blood<sup>36</sup>. Higher air pollution may increase the susceptibility of children in those regions to respiratory viruses.

We used the US EPA 1996 NATA database to assess air pollution exposure, as there is no database for multiple air pollutants corresponding to our enrollment period (1980–1984). However, the primary air pollution sources in Tucson (i.e., the freeways and other mobile sources)<sup>25</sup>, were in place prior to enrollment, and the relative concentrations between census tracts are unlikely to have substantially changed. For example, diesel particulate matter exposures are highly correlated between the 1996 and 2002 NATA databases ( $p=0.006$ ) and the traffic volumes between 1980 and 1996 are also highly correlated ( $p<0.0001$ ). The first principal component of air pollutant exposure concentrations is highly correlated between the 1996 and 2005 NATA databases ( $p<0.0001$ ). It is likely that we have identified children who were chronically exposed to higher levels of air pollution at their home. The NATA database provides modeled estimates based on reported emission inventories. Future studies should be performed to determine if these associations can be confirmed with air pollution measurements. However, the 1996 NATA database allowed us to examine both social and environment risk factors in this unique cohort. Air pollution exposure was determined independently of enrollment and study design. Although ideally more specific exposure estimates based on air pollution measurements should be used, this would likely lead to stronger associations than we have reported.

Children were not recruited from 5% of the census tracts in the Tucson region (E-Table 2). These census tracts consisted of Native American nations or had a high density of senior communities. Thus, our analyses do not consider these geographic regions, which may have their own unique risk factors. Infants were recruited through an HMO, which may have biased our results. Although these pediatric offices also cared for non-insured patients who paid for services received, it is likely that very low SES children are underrepresented in this study<sup>5</sup>. Although we recruited children from 95% of the census tracts, we may have underestimated the magnitude of the association of SES and LRI clusters. Similarly, viral culture methods for rhinovirus at the time of our data collection were less accurate than current methods, thus, we may have underestimated viral LRI incidence.

Despite these limitations, our study has numerous strengths. All LRIs were diagnosed prospectively by study physicians using predefined criteria in a non-selected cohort. Wheezing was confirmed by a physician, and viral cultures were obtained at the time of each LRI. Only children followed by study physicians for all three years were include in the analysis. Most studies of LRIs and social and environmental risk factors have relied on

hospital records or parental recall, which may result in misclassification of respiratory disease particularly in children<sup>37</sup>. It has also been hypothesized in previous studies that the association between LRIs and lower SES might result from poorer families postponing a doctor's office visit until the disease is more severe, making it more likely that the child will be hospitalized<sup>38</sup>. By having study physicians diagnose LRIs prospectively we have removed this potential bias. Furthermore, our study included multiple LRI diagnoses, unlike other studies, allowing for a more holistic assessment of LRI clustering by phenotype. The standardized comprehensive method of ascertaining LRIs in this cohort allowed us to elucidate the spatial patterns and associated risk factors for each LRI phenotype.

Another strength was our use of "SPATCLUS", a method that can detect arbitrarily-shaped spatial clusters<sup>22</sup>, making it more powerful than traditional methods that utilize a predefined shape (e.g., circle). Furthermore, the data transformation process takes into account potential inhomogeneity in the underlying population distribution. Both these features are important for communities like Tucson given the substantial inhomogeneity in the underlying population distribution resulting from unique geographic characteristics (i.e., four mountain ranges and two highways) (Figure 1). Unfortunately, this method cannot handle repeated events and we are unaware of any that can while providing these same benefits. Methods should be developed for recurrent illnesses such as LRIs, to strengthen these types of spatial analyses and better specify areas of increased risk.

We examined associations between LRI clusters and risk factors at the census tract level rather than the individual level to identify targeted areas for community-based interventions. However, depending upon the LRI phenotype, being in a cluster was also associated with several individual risk factors we previously identified in this cohort (data not shown), including: the number of persons sharing the child's bedroom, low maternal age, low maternal education, and use of evaporative coolers<sup>13, 15, 16, 39</sup>. Yet, these individual level risk factors also cluster in space and are likely indicators of larger-scale risk factors such as the multi-dimensional census tract level indices in our current analyses. While understanding individual-level risk factors is important for clinical recommendations, it does not facilitate targeting programs and interventions to geographic areas. Our work has demonstrated that specific characteristics are associated with neighborhoods at higher risk for childhood LRIs where intervention efforts could be directed. Particularly in communities with limited health care access<sup>40</sup>, this may be a method for delivering disease prevention programs such as outreach for RSV vaccination or pollution prevention.

In conclusion, we identified statistically significant geographic clusters for LRI incidence, which differed spatially by LRI phenotype. We demonstrated that increased air pollution exposure is associated with being in wheezing or RSV LRI cluster and that poorer housing conditions are associated with being in a viral or RSV LRI cluster. Rather than focusing on individual level characteristics, our use of community-level social and environmental indices allowed us to examine broad contextual factors that may affect LRI incidence. As LRIs are the leading cause of death in children globally and are associated with significant morbidities, studies like this can help elucidate environmental and social risk factors to target for community-level interventions in specific geographic regions.



## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

**Financial Support:** This work was supported by NIH grants ES006694, HL56177 and HL103970. The publication's contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

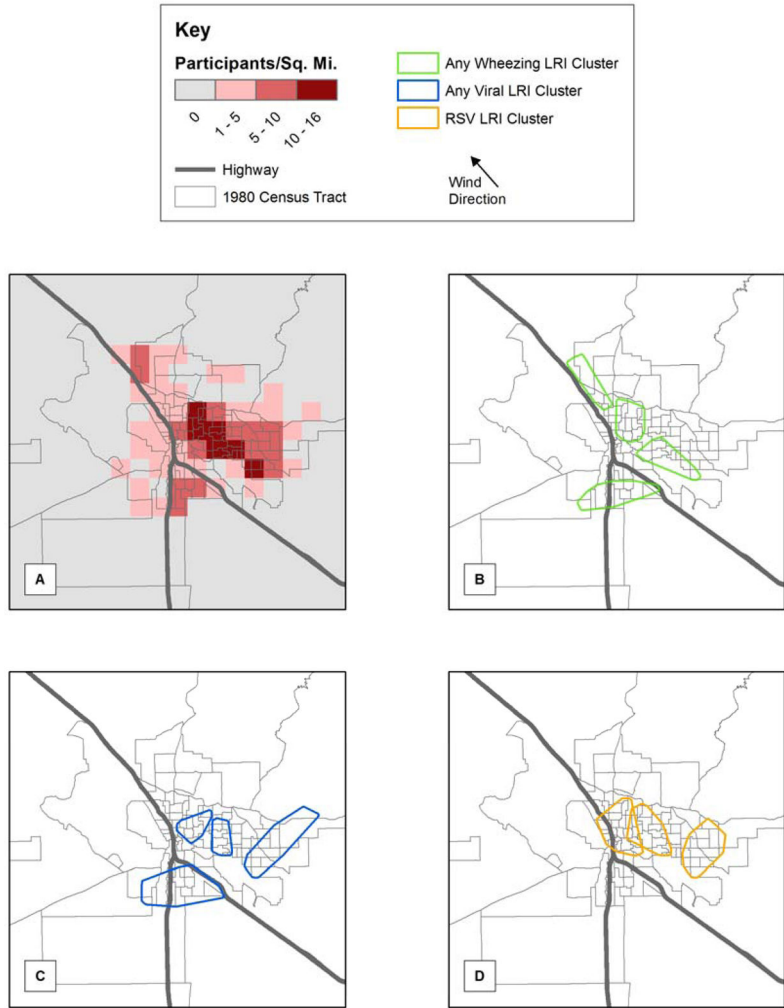
We gratefully acknowledge the contributions of Lynn Taussig who started the Tucson Children's Respiratory Study in 1980. We thank Catherine J. Holberg and Bruce Saul for data management and our study nurses, Marilyn Lindell, Nicole Pargas, and Lydia de la Ossa, for data collection and participant follow-up. We also thank Andrew Comrie, Eric Beternton, Lynn Gerald, and Duane Sherrill for their mentorship and insights on this manuscript.

## References

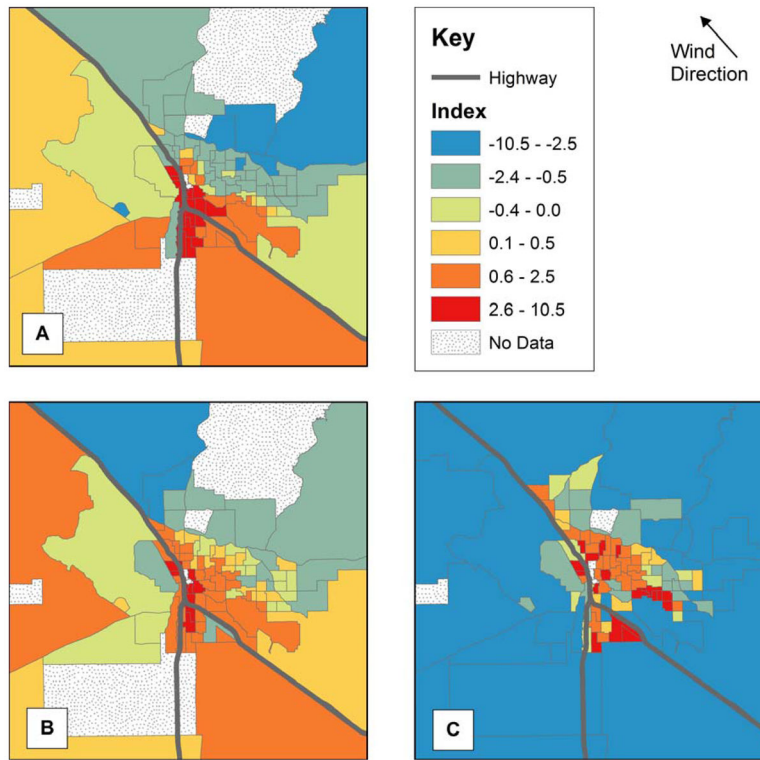
1. Mehta S, Shin H, Burnett R, North T, Cohen AJ. Ambient particulate air pollution and acute lower respiratory infections: A systematic review and implications for estimating the global burden of disease. *Air Quality, Atmosphere & Health*. 2011:1–15.
2. Mansbach JM, Emond JA, Camargo CA Jr. Bronchiolitis in US emergency departments 1992 to 2000: Epidemiology and practice variation. *Pediatr Emerg Care*. 2005; 21(4):242–7. [PubMed: 15824683]
3. Morgan WJ, Stern DA, Sherrill DL, Guerra S, Holberg CJ, Guilbert TW, Taussig LM, Wright AL, Martinez FD. Outcome of asthma and wheezing in the first six years of life: Follow-up through adolescence. *Am J Respir Crit Care Med*. 2005; 172(10):1253–4. [PubMed: 16109980]
4. Saglani S, Payne DN, Zhu J, Wang Z, Nicholson AG, Bush A, Jeffery PK. Early detection of airway wall remodeling and eosinophilic inflammation in preschool wheezers. *Am J Respir Crit Care Med*. 2007; 176(9):858–64. [PubMed: 17702968]
5. Taussig LM, Wright AL, Morgan WJ, Harrison HR, Ray CG. The Tucson Children's Respiratory Study I. Design and implementation of a prospective study of acute and chronic respiratory illness in children. *Am J Epidemiol*. 1989; 129(6):1219–31. [PubMed: 2729258]
6. Wright AL, Taussig LM, Ray CG, Harrison HR, Holberg CJ. The Tucson Children's Respiratory Study II. Lower respiratory tract illness in the first year of life. *Am J Epidemiol*. 1989; 129(6): 1232–1246. [PubMed: 2729259]
7. Taussig LM, Wright AL, Holberg CJ, Halonen M, Morgan WJ, Martinez FD. Tucson children's respiratory study: 1980 to present. *J Allergy Clin Immunol*. 2003; 111(4):661–75. [PubMed: 12704342]
8. Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and wheezing in the first six years of life. *N Engl J Med*. 1995; 332(3):133–8. [PubMed: 7800004]
9. Wright AL. Epidemiology of asthma and recurrent wheeze in childhood. *Clinical Reviews in Allergy and Immunology*. 2002; 22(1):33–44. [PubMed: 11803801]
10. Stein RT, Sherrill D, Morgan WJ, Holberg CJ, Halonen M, Taussig LM, Wright AL, Martinez FD. Respiratory syncytial virus in early life and risk of wheeze and allergy by age 13 years. *The Lancet*. 1999; 354(9178):541–5.
11. Voraphani N, Stern DA, Wright AL, Guerra S, Morgan WJ, Martinez FD. Risk of current asthma among adult smokers with respiratory syncytial virus illnesses in early life. *Am J Respir Crit Care Med*. 2014; 190(4):392–8. [PubMed: 24927374]
12. Wright AL, Holberg C, Martinez FD, Taussig LM. Relationship of parental smoking to wheezing and nonwheezing lower respiratory tract illnesses in infancy. *J Pediatr*. 1991 Feb; 118(2):207–14. [PubMed: 1993946]
13. Aldous MB, Holberg CJ, Wright AL, Martinez FD, Taussig LM. Evaporative cooling and other home factors and lower respiratory tract illness during the first year of life. *Am J Epidemiol*. 1996; 143(5):423–30. [PubMed: 8610657]

14. Ball TM, Castro-Rodriguez JA, Griffith KA, Holberg CJ, Martinez FD, Wright AL. Siblings, day-care attendance, and the risk of asthma and wheezing during childhood. *N Engl J Med*. 2000; 343(8):538–43. [PubMed: 10954761]
15. Martinez FD, Wright AL, Holberg CJ, Morgan WJ, Taussig LM. Maternal age as a risk factor for wheezing lower respiratory illnesses in the first year of life. *Am J Epidemiol*. 1992; 136(10):1258–68. [PubMed: 1476148]
16. Holberg CJ, Wright AL, Martinez FD, Ray CG, Taussig LM, Lebowitz MD. Risk factors for respiratory syncytial virus-associated lower respiratory illnesses in the first year of life. *Am J Epidemiol*. 1991; 133(11):1135–51. [PubMed: 2035517]
17. Demattei C, Molinari N, Daurès J. Arbitrarily shaped multiple spatial cluster detection for case event data. *Comput Stat Data Anal*. 2007; 51(8):3931–45.
18. Andrade ALSS, Martelli CMT, Oliveira RMD, Neto Morais, Libânio de Otaliba, Siqueira JB Júnior, Melo LK, Di Fábio JL. Population-based surveillance of pediatric pneumonia: Use of spatial analysis in an urban area of Central Brazil. *Cadernos De Saúde Pública*. 2004; 20(2):411–21. [PubMed: 15073620]
19. Sloan CD, Gebretsadik T, Wu P, Carroll KN, Mitchel E, Hartert TV. Spatiotemporal patterns of infant bronchiolitis in a Tennessee medicaid population. *Spat Spatiotemporal Epidemiol*. 2013; 6:17–23. [PubMed: 23956809]
20. Ali M, Emch M, Tofail F, Baqui AH. Implications of health care provision on acute lower respiratory infection mortality in Bangladeshi children. *Soc Sci Med*. 2001; 52(2):267–77. [PubMed: 11144783]
21. Elliott P, Wartenberg D. Spatial epidemiology: Current approaches and future challenges. *Environ Health Perspect*. 2004:998–1006. [PubMed: 15198920]
22. Demattei C, Molinari N, Daurès J. SPATCLUS: An R package for arbitrarily shaped multiple spatial cluster detection for case event data. *Comput Methods Programs Biomed*. 2006; 84(1):42–9. [PubMed: 16963154]
23. Bai J, Perron P. Estimating and testing linear models with multiple structural changes. *Econometrica*. 1998:47–78.
24. Minnesota Population Center. National Historical Geographic Information System: Version 2.0. Minneapolis, MN: University of Minnesota; 2011. Available at: [www.nhgis.org](http://www.nhgis.org) [Accessed: January 17, 2013]
25. US EPA. National-Scale Air Toxics Assessment. United States Environmental Protection Agency; 1996. Available at: <http://www.epa.gov/ttn/atw/nata/> [Accessed: August 1, 2012]
26. Windham GC, Zhang L, Gunier R, Croen LA, Grether JK. Autism spectrum disorders in relation to distribution of hazardous air pollutants in the San Francisco Bay Area. *Environ Health Perspect*. 2006; 114(9):1438–44. [PubMed: 16966102]
27. Messer LC, Laraia BA, Kaufman JS, Eyster J, Holzman C, Culhane J, Elo I, Burke JG, O'Campo P. The development of a standardized neighborhood deprivation index. *Journal of Urban Health*. 2006; 83(6):1041–62. [PubMed: 17031568]
28. Omer S, Sutanto A, Sarwo H, Linehan M, Djelantik I, Mercer D, Moniaga V, Moulton L, Widjaya A, Muljati P. Climatic, temporal, and geographic characteristics of respiratory syncytial virus disease in a tropical island population. *Epidemiol Infect*. 2008; 136(10):1319. [PubMed: 18177515]
29. Tupasi TE, Velmonte MA, Sanvictores MEG, Abraham L, De Leon LE, Tan SA, Miguel CA, Saniel MC. Determinants of morbidity and mortality due to acute respiratory infections: Implications for intervention. *J Infect Dis*. 1988; 157(4):615–23. [PubMed: 3346561]
30. Cohen S. Social status and susceptibility to respiratory infections. *Ann NY Acad Sci*. 1999; 896(1): 246–53. [PubMed: 10681901]
31. Simoes EA. Environmental and demographic risk factors for respiratory syncytial virus lower respiratory tract disease. *J Pediatr*. 2003; 143(5):118–26.
32. Karr CJ, Rudra CB, Miller KA, Gould TR, Larson T, Sathyanarayana S, Koenig JQ. Infant exposure to fine particulate matter and traffic and risk of hospitalization for RSV bronchiolitis in a region with lower ambient air pollution. *Environ Res*. 2009; 109(3):321. [PubMed: 19211100]

33. Pershagen G, Rylander E, Norberg S, Eriksson M, Nordvall SL. Air pollution involving nitrogen dioxide exposure and wheezing bronchitis in children. *Int J Epidemiol.* 1995; 24(6):1147–53. [PubMed: 8824856]
34. Gold DR, Burge HA, Carey V, Milton DK, Platts-Mills T, Weiss ST. Predictors of repeated wheeze in the first year of life: The relative roles of cockroach, birth weight, acute lower respiratory illness, and maternal smoking. *Am J Respir Crit Care Med.* 1999; 160(1):227–36. [PubMed: 10390405]
35. Wright RJ, Fisher K, Chiu YM, Wright RO, Fein R, Cohen S, Coull BA. Disrupted prenatal maternal cortisol, maternal obesity, and childhood wheeze: Insights into prenatal programming. *Am J Respir Crit Care Med.* 2013; 187(11):1186–93. [PubMed: 23590260]
36. Ando M, Shima M, Adachi M, Tsunetoshi Y. The role of intercellular adhesion molecule-1 (ICAM-1), vascular cell adhesion molecule-1 (VCAM-1), and regulated on activation, normal T-cell expressed and secreted (RANTES) in the relationship between air pollution and asthma among children. *Arch Environ Health.* 2001; 56(3):227–233. [PubMed: 11480498]
37. Bråbäck L, Björ O, Nordahl G. Early determinants of first hospital admissions for asthma and acute bronchitis among Swedish children. *Acta Paediatrica.* 2003; 92(1):27–33. [PubMed: 12650295]
38. Thörn L, Minamisava R, Nouer S, Ribeiro L, Andrade A. Pneumonia and poverty: A prospective population-based study among children in Brazil. *BMC Infectious Diseases.* 2011; 11(1):180. [PubMed: 21696610]
39. Wright AL, Holberg CJ, Martinez FD, Morgan WJ, Taussig LM. Breast feeding and lower respiratory tract illness in the first year of life. *Br Med J.* 1989; 299(6705):946–9. [PubMed: 2508946]
40. Crighton EJ, Elliott SJ, Moineddin R, Kanaroglou P, Upshur R. A spatial analysis of the determinants of pneumonia and influenza hospitalizations in Ontario (1992–2001). *Soc Sci Med.* 2007; 64(8):1636–50. [PubMed: 17250939]



**Figure 1.** (A) Map of enrollment address of all CRS participants and statistically significant clusters of lower respiratory illness (LRI) for: (B) any wheezing LRI, (C) any viral LRI, and (D) any RSV LRI. Note that children in the cluster reside within the outlines, however there may be children who live within the outlines that are not part of the disease cluster. Individual cases within the clusters are not shown to protect study subjects confidentiality.



**Figure 2.** Map of first principal component for (A) lower socioeconomic status (SES), (B) poorer housing characteristics, and (C) increased air pollution.

**Table 1**

Summary and first principal component loading coefficients of socioeconomic, housing characteristic, and air pollution variables by census tract for CRS participants in Tucson, AZ (n=812). Proportion of variance explained by first principal component is indicated for each category.

Variable	Median	Range	Loading Coefficient	Proportion of Variance
<i>Socioeconomic Characteristics<sup>a</sup></i>				56%
Median Age (yrs) <sup>b</sup>	28.7	20.9–63.8	0.28	
Foreign Born (%)	5.5	1.7–21.0	0.32	
Male (%)	48.8	41.7–53.4	0.07	
Minority Race (%)	9.3	1.0–69.4	0.38	
Spanish Origin (%)	10.7	1.6–93.2	0.39	
Per capita income (\$) <sup>b</sup>	7,040	2,080–15,710	0.37	
Family with single female head of house (%)	8.7	3.4–39.2	0.28	
Do not speak English at home (%)	14.2	2.0–91.6	0.38	
Unemployed 15 weeks (%)	27.8	10.3–48.4	0.20	
12 years of school (%)	57.0	24.0–89.9	0.33	
Population Density (people/sq. mile)	36,700	100–103,200	0.10	
<i>Housing Characteristics<sup>c</sup></i>				36%
Mobile Homes (%)	1.4	0.0–78.1	0.05	
Attached Homes (%)	23.0	1.7–70.8	0.19	
Gas Heating (%)	81.4	50.8–96.2	0.49	
Gas Water Heating (%)	83.7	52.5–99.6	0.53	
Lacking Complete Plumbing Facilities (%)	0.5	0.0–9.9	0.38	
Built before 1940 (%)	1.1	0.0–57.2	0.39	
> 1 person per room (%)	3.4	0.3–26.6	0.38	
<i>Air Pollutant Concentrations (<math>\mu\text{g}/\text{m}^3</math>)</i>				45%
Acetaldehyde	$4.5 \times 10^{-1}$	$9.0 \times 10^{-2}$ –1.2	0.12	
Acrolein	$5.5 \times 10^{-2}$	$9.5 \times 10^{-4}$ – $8.7 \times 10^{-2}$	0.20	
Arsenic	$6.9 \times 10^{-5}$	$2.0 \times 10^{-6}$ – $6.8 \times 10^{-4}$	0.11	
Benzene	1.1	$5.5 \times 10^{-1}$ –1.5	0.26	
Butadiene	$4.6 \times 10^{-2}$	$8.0 \times 10^{-3}$ – $8.0 \times 10^{-2}$	0.26	
Chromium	$3.8 \times 10^{-4}$	$1.7 \times 10^{-5}$ – $2.1 \times 10^{-3}$	0.14	
Diesel PM <sup>d</sup>	$5.4 \times 10^{-1}$	$2.1 \times 10^{-1}$ –1.3	0.18	
Formaldehyde	$4.9 \times 10^{-1}$	$2.5 \times 10^{-1}$ – $8.3 \times 10^{-1}$	0.19	
Lead	$1.3 \times 10^{-3}$	$5.1 \times 10^{-5}$ – $2.3 \times 10^{-2}$	0.07	
Manganese	$1.0 \times 10^{-3}$	$7.0 \times 10^{-5}$ – $1.8 \times 10^{-3}$	0.26	
Mercury	$1.2 \times 10^{-3}$	$1.2 \times 10^{-3}$ – $1.3 \times 10^{-3}$	0.22	
Nickel	$5.8 \times 10^{-4}$	$1.7 \times 10^{-5}$ – $2.2 \times 10^{-3}$	0.21	
POM <sup>e</sup>	$5.5 \times 10^{-2}$	$4.7 \times 10^{-3}$ – $9.8 \times 10^{-2}$	0.26	

Variable	Median	Range	Loading Coefficient	Proportion of Variance
Cadmium	$2.5 \times 10^{-5}$	$1.0 \times 10^{-6}$ - $2.0 \times 10^{-4}$	0.15	
Beryllium	$9.8 \times 10^{-6}$	$4.9 \times 10^{-7}$ - $2.5 \times 10^{-5}$	0.26	
Trichloroethylene	$8.9 \times 10^{-2}$	$7.0 \times 10^{-2}$ - $1.4 \times 10^{-1}$	0.25	
Perchloroethylene	$1.7 \times 10^{-1}$	$1.1 \times 10^{-1}$ - $2.3 \times 10^{-1}$	0.30	
Ethylene dichloride	$5.3 \times 10^{-2}$	$5.2 \times 10^{-2}$ - $5.4 \times 10^{-2}$	0.05	
Chloroform	$6.8 \times 10^{-2}$	$6.8 \times 10^{-2}$ - $6.9 \times 10^{-2}$	0.15	
Methylene Chloride	$2.4 \times 10^{-1}$	$1.3 \times 10^{-1}$ - $4.7 \times 10^{-1}$	0.27	
1,2-dichloropropene	$6.8 \times 10^{-2}$	$7.6 \times 10^{-3}$ - $1.3 \times 10^{-1}$	0.26	
Ethylene Oxide	$1.6 \times 10^{-3}$	$1.2 \times 10^{-4}$ - $1.1 \times 10^{-2}$	0.19	
Quinoline	$2.1 \times 10^{-8}$	0.0- $1.3 \times 10^{-7}$	0.14	
Hydrazine	$5.0 \times 10^{-9}$	0.0- $5.7 \times 10^{-8}$	0.13	

<sup>a</sup> denominator is number of people responding to question;

<sup>b</sup> inverse was used for principal component;

<sup>c</sup> denominator is number of housing units responding to question;

<sup>d</sup> PM is particulate matter;

<sup>e</sup> POM is polycyclic organic matter

Simple logistic regression analyses for the association of being in a cluster by phenotype of lower respiratory illness and socioeconomic and housing characteristics at the census tract obtained from the 1980 US Census of Population and Housing.

**Table 2**

Characteristic	Wheezing LRI Cluster		Viral LRI Cluster		RSV LRI Cluster	
	OR	P	OR	P	OR	P
<i>Socioeconomic Characteristics</i>						
Median Age (yrs)	0.89	0.002*	0.93	0.04	0.98	0.53
Foreign Born (%)	1.03	0.45	1.06	0.05	0.99	0.77
Male (%)	1.19	0.05	1.08	0.40	0.91	0.24
Minority Race (%)	1.01	0.22	1.02	0.06	0.99	0.36
Spanish Origin (%)	1.01	0.23	1.01	0.18	0.99	0.20
Per capita income (per \$1000)	0.81	0.02	0.82	0.007	0.94	0.41
Family with single female head of house (%)	1.02	0.65	1.06	0.09	1.08	0.03
Do not speak English at home (%)	1.01	0.25	1.01	0.18	0.99	0.17
Unemployed 15 or more weeks (%)	0.97	0.32	1.01	0.58	0.94	0.01
12 years of school (%)	1.02	0.15	1.01	0.11	0.99	0.27
Population Density (1000 people/sq. mile)	1.00	0.64	1.01	0.08	1.03	0.007
<i>Housing Characteristics</i>						
Mobile Homes	1.01	0.07	1.00	0.83	0.97	0.01
Attached Homes	1.00	0.55	1.00	0.76	1.03	<0.001*
Gas Heating	1.00	0.76	1.04	0.002*	1.04	0.006*
Gas Water Heating	1.00	0.93	1.04	0.003*	1.04	0.002*
Lacking complete plumbing facilities	0.98	0.85	1.03	0.79	1.17	0.05
Built before 1940	1.02	0.17	1.02	0.04	1.03	0.04
> 1 person per room	1.04	0.04	1.05	0.008	0.97	0.27

Note. OR = odds ratio;

\* indicates significance after Bonferroni correction ( $p < 0.005$  for SES and  $p < 0.007$  of housing characteristics).



**Table 3**

Simple logistic regression analyses for the association of being in a cluster by phenotype of lower respiratory illness and exposure concentration of air pollutant ( $\mu\text{g}/\text{m}^3$ ) in the census tract (per interquartile range of air pollutant).

Pollutant	Wheezing LRI Cluster		Viral LRI Cluster		RSV LRI Cluster	
	OR	P	OR	P	OR	P
Acetaldehyde	1.45	0.004	1.02	0.91	0.76	0.12
Acrolein	1.15	0.0	1.04	0.09	1.19	0.28
Arsenic	1.22	0.35	0.76	0.26	1.09	0.41
Benzene	1.51	0.09	0.96	0.85	2.08	0.001*
Butadiene	1.33	0.17	1.45	0.0	2.05	<0.001*
Chromium	1.41	<0.001*	0.98	0.86	0.79	0.07
Diesel PM	1.41	<0.001*	1.10	0.19	0.89	0.32
Formaldehyde	1.29	0.01	0.12	0.22	0.96	0.73
Lead	1.08	<0.001*	1.08	0.002	0.81	0.003
Manganese	1.49	0.11	0.95	0.79	2.22	<0.001*
Mercury	1.31	0.02	1.30	0.008	1.56	<0.001*
Nickel	1.62	<0.001*	1.07	0.56	1.02	0.85
POM	1.46	0.14	0.94	0.77	2.28	<0.001*
Cadmium	1.28	0.36	0.86	0.39	1.13	0.32
Beryllium	1.23	0.04	1.12	0.23	1.53	<0.001*
Trichloroethylene	1.57	0.002	1.34	0.02	1.47	0.002
Perchloroethylene	1.72	0.008	1.25	0.18	1.96	<0.001*
Ethylene dichloride	1.43	0.10	1.17	0.40	1.05	0.80
Chloroform	1.63	0.03	1.35	0.13	1.31	0.16
Methylene chloride	1.64	0.006	1.35	0.05	1.77	<0.001*
1,2-dichloropropene	1.51	0.10	0.96	0.86	2.17	<0.001*
Ethylene oxide	1.23	0.08	1.27	0.04	1.61	<0.001*
Quinolone	1.19	<0.001*	1.04	0.41	0.91	0.20

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Pollutant	Wheezing LRI Cluster		Viral LRI Cluster		RSV LRI Cluster	
	OR	P	OR	P	OR	P
Hydrazine	1.12	<0.001*	1.05	0.10	0.93	0.17

Note. PM=particulate matter; POM= polycyclic organic matter; CI = confidence interval; OR = odds ratio;

\* indicates significance after Bonferroni correction ( $p < 0.002$  for air pollutants).

**Table 4**

Simple and multivariable logistic regression analyses of being in a cluster by lower respiratory illness (LRI) phenotype and first principal component of community variables

Community variables	Wheezing LRI Cluster	Viral LRI Cluster	RSV LRI Cluster
	OR (95% CI)	OR (95% CI)	OR (95% CI)
<i>Simple Analyses</i>			
Lower Socioeconomic Status	1.09 (0.99 – 1.20)	1.11 (1.02 – 1.21) *	0.97 (0.87 – 1.09)
Poorer Housing Characteristics	1.08(0.88 – 1.32)	1.28 (1.11 – 1.48) **	1.30 (1.09 – 1.53) **
Increased Air Pollution	1.18(1.05 – 1.32) **	1.05 (0.96 – 1.15)	1.16 (1.06 – 1.27) **
<i>Multivariable Analysis</i>			
Lower Socioeconomic Status	1.10 (0.92 – 1.32)	1.00 (0.88 – 1.14)	0.76 (0.66 – 0.89) **
Poorer Housing Characteristics	0.86 (0.63 – 1.18)	1.28 (1.05 – 1.56) *	1.54 (1.16 – 2.06) **
Increased Air Pollution	1.18 (1.04 – 1.35) *	1.00 (0.91 – 1.10)	1.14 (1.04 – 1.26) *

Note. CI = confidence interval; OR = odds ratio;

\*  $P < 0.05$ ;

\*\*  $P < 0.01$