

# Assessing the Association between Heart Attack, High Blood Pressure, and Heart Disease Mortality Rates and Particulate Matter and Socioeconomic Status Using Multivariate Geostatistical Model

**Faye Anderson**

Center for Advances in Water and Air Quality, Lamar University, Beaumont, TX, USA  
Email: andersonfaye7@gmail.com

Received 14 January 2016; accepted 16 February 2016; published 19 February 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

**This study addresses the public concerns of potential adverse health effects from ambient fine particulate matter as well as socioeconomic factors. Heart attack, high blood pressure, and heart disease mortality rates were investigated against fine particulate matter and socioeconomic status, for all counties in the United States in 2013. Multivariate multiple regressions as well as multivariate geostatistical predictions show that these are significant factors towards assessing the causal inferences between exposure to air pollution and socioeconomic status and the three mortality rates.**

## Keywords

**Heart Attack, High Blood Pressure, Heart Disease, Mortality, United States, Multivariate Geostatistics, Coregionalization, Multivariate Multiple, Linear**

---

## 1. Introduction

The cardiovascular disease of heart disease is the number one cause of death in the World with its mortality rate representing one third of all global deaths. Its documented risk factors include smoking, physical activity, diet, obesity, cholesterol, and diabetes [1]. As for heart attack mortality rates, research shows that despite the progress made in faster hospital care, its rates remain unchanged [2]. The third health outcome of interest is high blood

**How to cite this paper:** Anderson, F. (2016) Assessing the Association between Heart Attack, High Blood Pressure, and Heart Disease Mortality Rates and Particulate Matter and Socioeconomic Status Using Multivariate Geostatistical Model. *Occupational Diseases and Environmental Medicine*, 4, 8-15. <http://dx.doi.org/10.4236/odem.2016.41002>

pressure mortality rate which contributes to thirteen percent of all deaths in the US and is continually increasing [3]. The three health outcomes have been associated with air pollution as a consequence of oxidative stress leading to inflammation, which generates the physiological processes evolving as cardiovascular symptoms like narrowing of airways, shortness of breath, wheezing, cough, and the ability of particles to penetrate the lung wall accumulating in the pulmonary interstitium between the lung and the bloodstream [4]-[6]. The outcomes have also been negatively associated with socioeconomic status (SES) through life style effects [7] [8].

The purpose of this study is to quantify the association between the three health outcomes and exposure to fine particulate matter (PM<sub>2.5</sub>) and the different aspects of SES in the US, assessing causal inferences between exposure to fine particulate matter and mortality rates of heart attacks, high blood pressure, and heart disease provides epidemiological evidence regarding the adverse health effects of air pollution. Studying this association at the county level can address the question of whether the current network of air quality monitoring stations adequately represent the populations and locations with highest risk for respiratory and cardiovascular that has not been fully answered yet. An important underlying reason for this problem is the uniform approach that has been implemented to select monitoring sites locations, which ignores the dissimilarities among sites e.g., Houston, Texas versus Jackson, Wyoming [9] [10]. In addition, the effect of local air pollution versus neighboring levels is not fully explored yet [11] [12], which points to the importance of studying community-related characteristics because of the previously-mentioned factors' contribution to the causal inference between air pollution and its respiratory and cardiovascular health.

## 2. Methods

To achieve the study objective, data was collected from the Centers for Disease Control and Prevention (CDC) [13] and the US Environmental Protection Agency (EPA) [14], for the year 2013. **Table 1** gives the twenty four variables collected and analyzed. Statistical analyses were conducted using R version 3.1.1.

**Table 1.** Data collected from CDC and EPA.

Variable
High Blood Pressure Death Rate per 100,000 All Ages All Race All Gender 2011-2013
Heart Attack Death Rate per 100,000 All Ages All Race All Gender 2011-2013
Heart Disease Death Rate per 100,000 All Ages All Race All Gender 2011-2013
Percentage without High School Diploma Ages 25+ 2009-2013 (5-year)
Percentage without 4+ Years College Ages 25+ 2009-2013 (5-year)
Families with Female Head of Household (%) 2009-2013 (5-year)
Percentage Food Stamp/Supplemental Nutrition Assistance Program Recipients 2011
Median Household Income (in thousands of \$) 2013
Percentage Living in Poverty All Ages 2013
Unemployment Rate Ages 16+ 2014
American Indian/Alaska Native non-Hispanic Population (%) All Ages 2009-2013 (5-year)
Asian and Pacific Islander non-Hispanic Population (%) All Ages 2009-2013 (5-year)
Black non-Hispanic Population (%) All Ages 2009-2013 (5-year)
Hispanic/Latino Population (%) All Ages 2009-2013 (5-year)
White non-Hispanic Population (%) All Ages 2009-2013 (5-year)
Two or More Race non-Hispanic Population (%) All Ages 2009-2013 (5-year)
Other Race non-Hispanic Population (%) All Ages 2009-2013 (5-year)
Population Aged 65 and Older (%) 2009-2013 (5-year)
Total Population 2009-2013 (5-year)
Percentage of Population Living Within Half a Mile of a Park 2010
Annual Average Ambient Concentrations of PM <sub>2.5</sub> 2011
Percentage of Households Living with Severe Housing Problems 2007-2011 (5-year)
Urban-Rural Status 2013
Hospitals 2011

### 3. Results

**Table 2** gives some of the descriptive statistics for the main variables. The high departures from skewness and kurtosis from 0 and 3, respectively, which are the characteristic values of normal distribution show non symmetric nature of heart attack mortality, high blood pressure mortality, heart disease mortality, educational attainment, poverty, unemployment rate, percentage of population aged 65 years and older, percentage of population within a half a mile of a park, annual average ambient PM2.5, and percentages of households living with severe housing problems. **Figure 1** presents average mortality rates for heart attacks, heart disease and high blood pressure per type of county. Heart disease mortality rates are lowest in rural areas compared to urban and metro counties. High blood pressure mortality rates follow a somewhat opposite pattern whereas heart attack mortality rates are lowest in metro counties and highest in urban counties. As for pairwise associations (**Table 3**), the percentage of households living with severe housing problems was positively associated with poverty, percentage of food stamp/supplemental nutrition assistance program recipients, and percentages of Native Americans, Hispanics, and African Americans. It was negatively associated with percentage of population aged 65 years and older and represented about third of the population. Ease of access to nearby parks (Percentage of Population Living within Half a Mile of a Park) was negatively associated with both heart attack and heart disease mortalities. Heart attack mortality rate was positively associated with heart disease mortality rates, educational attainment, and poverty. High blood pressure mortality rate was positively associated with heart disease, poverty, educational attainment, and percentage of African Americans. Heart disease mortality rate was positively associated with educational attainment, poverty, and percentage of African Americans. Understandingly, percentage of population in poverty was positively associated with the percentage of food stamps recipients, median income and unemployment.

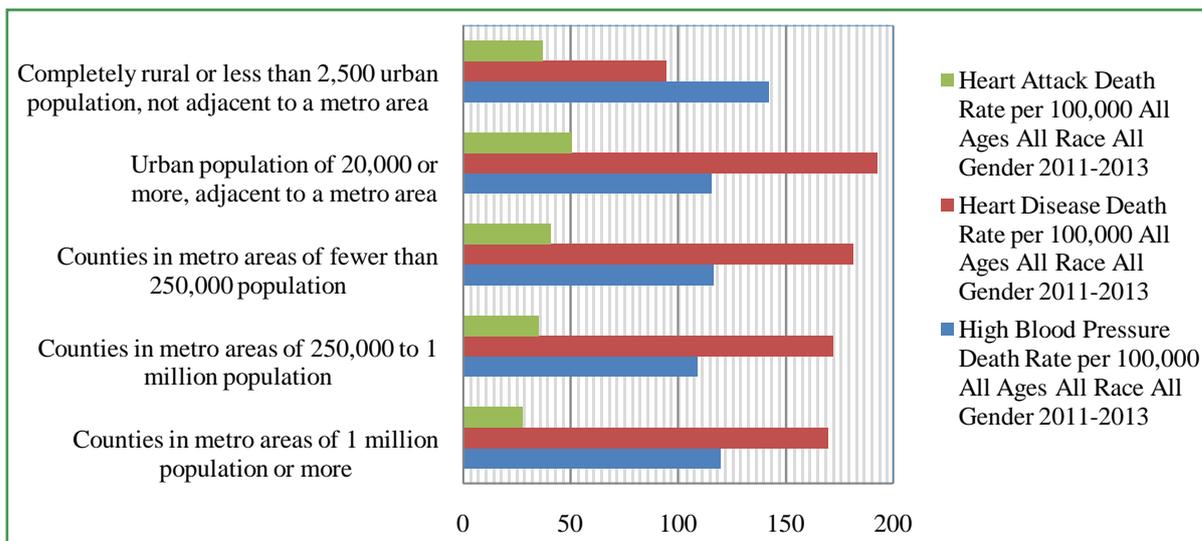
Multivariate analysis of covariance (MANOVA) tests with the three dependent variables heart attack, high blood pressure, and heart disease mortality rates using Pillai Trace, Wilk's Lambda, and Hotelling Trace gave p-values of zeros, which reject the null hypothesis that the population has no effect on the three different health outcomes at the significance level of 0.05. Three univariate (one dependent variable) linear regressions were performed after applying the logarithmic transformation function on the independent variables of the most significant variables of Percentage without high school diploma, families with female head of household, percentage living in poverty, percentage of Asians, percentage of African Americans, percentage of Hispanics, percentage

**Table 2.** Some descriptive statistics.

Variable	Mean	Std. Dev.	Kurtosis	Skewness
Heart Attack Death Rate per 100,000 All Ages All Race All Gender 2011-2013	45.88	26.60	5.99	1.90
High Blood Pressure Death Rate per 100,000 All Ages All Race All Gender 2011-2013	115.78	48.38	5.16	1.39
Heart Disease Death Rate per 100,000 All Ages All Race All Gender 2011-2013	184.64	46.72	3.37	0.31
Percentage without High School Diploma Ages 25+ 2009-2013 (5-year)	15.77	7.39	0.91	0.91
Percentage without 4+ Years College Ages 25+ 2009-2013 (5-year)	79.98	9.72	13.76	-2.62
Families with Female Head of Household (%) 2009-2013 (5-year)	17.10	6.89	2.00	1.08
Percentage Food Stamp/Supplemental Nutrition Assistance Program Recipients 2011	15.47	8.24	1.26	0.75
Median Household Income (in thousands of \$) 2013	45.35	12.04	3.29	1.24
Percentage Living in Poverty All Ages 2013	16.76	7.09	1.54	0.60
Unemployment Rate Ages 16+ 2014	6.49	2.89	8.22	2.11
Population Aged 65 and Older (%) 2009-2013 (5-year)	16.22	4.36	1.97	0.54
Percentage of Population Living Within Half a Mile of a Park 2010	32.56	31.86	0.07	1.15
Annual Average Ambient Concentrations of PM2.5 2011	9.35	1.73	-0.22	-0.37
Percentage of Households Living with Severe Housing Problems 2007-2011 (5-year)	14.04	5.12	10.38	1.46

**Table 3.** The most notable pairwise associations.

Variable	Heart Attack Death Rate	High Blood Pressure Death Rate	Heart Disease Death Rate	Percentage without 4+ Years College	Poverty	Black Population (%)	Hispanic Population (%)
Heart Attack Death Rate	1.00						
High Blood Pressure Death Rate	0.12	1.00					
Heart Disease Death Rate	0.58	0.31	1.00				
Percentage without 4+ Years College	0.40	0.23	0.50	1.00			
Poverty	0.37	0.26	0.58	0.42	1.00		
Black Population (%)	0.11	0.31	0.35	0.10	0.51	1.00	
Hispanic Population (%)	-0.08	0.09	-0.32	0.01	-0.21	-0.14	1.00
White Population (%)	0.06	-0.26	0.11	0.06	-0.17	-0.46	-0.71



**Figure 1.** Rates of health outcomes in rural/urban counties.

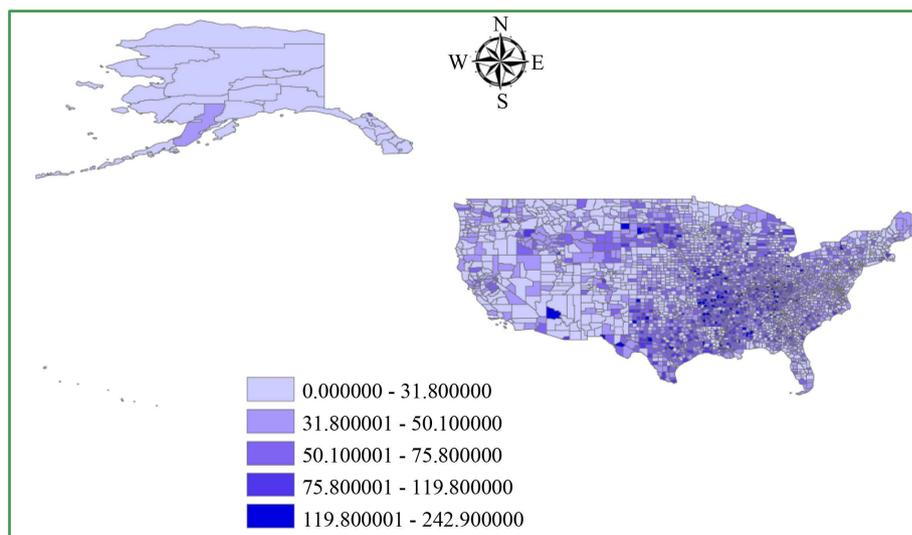
of white population, percentage aged 65 years or older, percentage of population living within half a mile of a park, and annual average for PM2.5. The other variables were excluded because of their minimal association or because of their strong association with the other predictors. **Table 4** presents the results of the three univariate multiple linear regressions for each of the health outcomes whereas **Table 5** presents the results for the multivariate multiple linear regression. The three univariate models are statistically significant. Also, the significant predictors are the same between the three univariate and the multivariate multiple regression model. However, one of the advantages of conducting a multivariate multiple regression is the ability to test the coefficients across the different health outcome variables. In this case, ten different hypotheses were tested simultaneously: testing that the coefficient is zero for each of the ten predictors in all three equations. The residuals mean is near 0.0001 but its variance is as high as 769. Spatial distribution of the residuals is given in **Figure 2**. Positive residuals indicating over-prediction and negative residuals indicating under-prediction indicate the need to improve the

**Table 4.** Univariate multiple linear regression results.

Dependent Variable	Multiple R	Adjusted R Square	Significant Predictors
Heart Attack Mortality Rate	0.52	0.27	Educational attainment, poverty, percentage Asian, percentage African American, percentage white, percentage aged 65 years or older, percentage near parks, and annual PM2.5 concentration.
High Blood Pressure Mortality Rate	0.42	0.17	Educational attainment, poverty, percentage Asian, percentage African American, percentage white, and annual PM2.5 concentration.
Heart Disease Mortality Rate	0.72	0.52	Educational attainment, poverty, percentage Asian, percentage African American, percentage white, and annual PM2.5 concentration.

**Table 5.** Multivariate multiple linear regression results.

Dependent Variable	Multiple R	Adjusted R Square	Significant Predictors
Heart Attack Mortality Rate	0.52	0.27	Educational attainment, poverty, percentage Asian, percentage African American, percentage white, percentage aged 65 years or older, percentage near parks, and annual PM2.5 concentration.
High Blood Pressure Mortality Rate	0.42	0.17	Educational attainment, poverty, percentage Asian, percentage African American, percentage white, and annual PM2.5 concentration.
Heart Disease Mortality Rate	0.72	0.52	Educational attainment, poverty, percentage Asian, percentage African American, percentage white, and annual PM2.5 concentration.



**Figure 2.** Residuals of the multivariate multiple linear regression model.

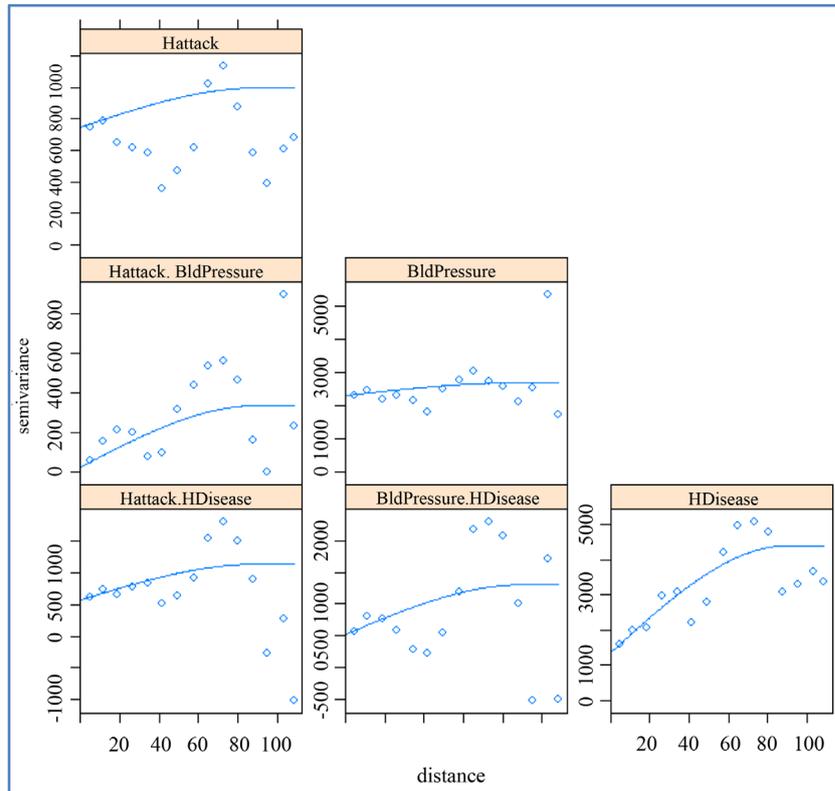
model, which is supported by the low adjusted R square values.

Codispersion coefficients for the three possible combinations of dependent variables (heart attack mortality vs. high blood pressure mortality, heart attack mortality vs. heart disease mortality, and high blood pressure mortality vs. heart disease mortality) were not constant. Thus, the hypothesis of intrinsic correlation could not be accepted. This means that the correlation structure of the three dependent variables is influenced by the spatial scale, which calls for cokriging [15]. **Figure 3** gives the direct variograms (diagonal) and cross variograms (off-diagonal) along with fitted linear model of coregionalization (straight line), with the distance maxed at 90 kilometers. The three mortality rates are predicted to keep rising with the persisting spatial concentrations

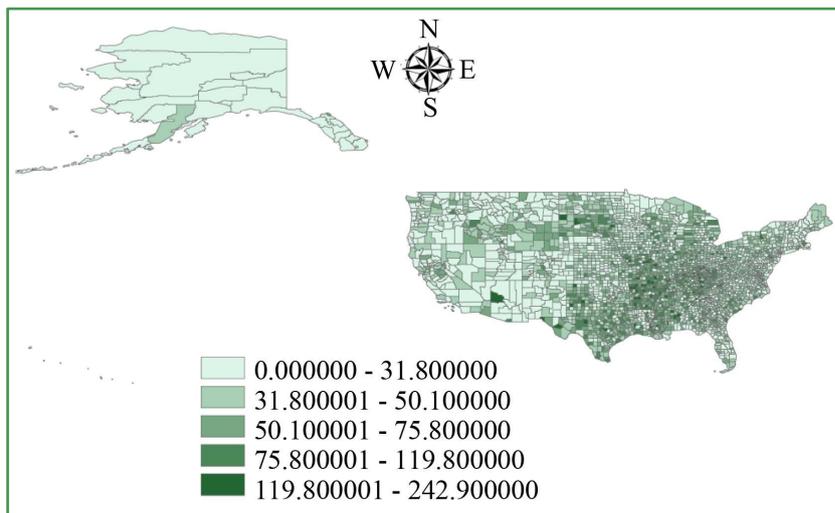
(hotspots) shown in **Figure 4**.

### 4. Discussion

Research has shown associations between air pollution and mortality of respiratory and cardiovascular systems [16]-[19]. Socioeconomic status (SES) has also been associated with these outcomes. Associations between heart attack, high blood pressure, and heart disease mortality rates and between them and SES have been docu-



**Figure 3.** Direct variograms (diagonal) and cross variograms (off-diagonal) along with fitted linear model of coregionalization (straight line).



**Figure 4.** Predicted heart attack mortality rate.

mented as well [20]-[24]. Nevertheless, such associations are seldom investigated in the same context of exposure to air pollution. Application of multivariate geostatistical techniques can assist in this regards because it has the advantage of making use of the spatial correlations between observations pertinent to communities to predict values at unsampled places. These spatial associations can exist even with lack of pair-wise correlations, known as “intrinsic correlation”. Multivariate intrinsic correlations allow simplification of data modeling, which in turn enables us to address previously unidentified or seldom studied associations [25]-[27]. Consequently, new public health findings can be sought that are relevant to each community.

“Hotspots” are geographic areas that form at specific times and places where the mortality or morbidity rate is consistently rising. This kind of clustering reflects high occurrence and frequency of disease, to the point that rates become highly predictable. However, the definition of a hotspot can be ambiguous due to lack of quantitative measures that can be incorporated or agreed upon among other issues. Nevertheless, studying the three mortality rates at the county level shows consistent spatial distribution where certain areas persistently have higher rates than the rest. Paying more attention to such hotspots can lead research to potentially stronger causal inferences [28]-[32].

## 5. Conclusion

This study investigated the three mortality rates of heart attack, high blood pressure, and heart disease in the United States, and their association with fine particulate matter and socioeconomic status represented by educational attainment, poverty, household income, unemployment rate, demographics, access to a nearby park, and urban-rural classification, in 2013. The multivariate geostatistical model used latitude, longitude, educational attainment, poverty, percentage Asian, percentage African American, percentage white population, and annual PM2.5 concentration as the mean covariates. These variables were also significant in univariate and multivariate multiple linear regressions. This study had the limitations of not exploring temporal variations, inherent issues with the data, non-consideration of other air pollutants and demographic factors like gender and age. Nevertheless, it pointed to significant findings including the county level which showed consistent spatial distribution where certain areas persistently had higher rates than the rest. Paying more attention to such hotspots could lead research to potentially stronger causal inferences.

## References

- [1] Go, A.S., Mozaffarian, D., Roger, V.L., Benjamin, E.J., Berry, J.D., Borden, W.B., Bravata, D.M., Dai, S., Ford, E.S., Fox, C.S., Franco, S., Fullerton, H.J., Gillespie, C., Hailpern, S.M., Heit, J.A., Howard, V.J., Huffman, M.D., Kissela, B.M., Kittner, S.J., Lackland, D.T., Lichtman, J.H., Lisabeth, L.D., Magid, D., Marcus, G.M., Marelli, A., Matchar, D.B., McGuire, D.K., Mohler, E.R., Moy, C.S., Mussolino, M.E., Nichol, G., Paynter, N.P., Schreiner, P.J., Sorlie, P.D., Stein, J., Turan, T.N., Virani, S.S., Wong, N.D., Woo, D. and Turner, M.B. (2013) Heart Disease and Stroke Statistics-2013 Update: A Report from the American Heart Association. *Circulation*, **127**, e6-e245. <http://dx.doi.org/10.1161/CIR.0b013e31828124ad>
- [2] Arbor, A. (2015) Heart Attack Mortality Rates Unchanged in Spite of Faster Care at Hospitals|University of Michigan Health System. <http://www.uofmhealth.org/news/archive/201309/heart-attack-mortality-rates-unchanged-spite-faster-care>
- [3] Department of Human Services (Victoria) and Heart Foundation (2011) Blood Pressure. *Better Health Channel*. [http://www.betterhealth.vic.gov.au/bhcv2/bhcarticles.nsf/pages/Blood\\_pressure\\_explained](http://www.betterhealth.vic.gov.au/bhcv2/bhcarticles.nsf/pages/Blood_pressure_explained)
- [4] Donaldson, K., Mills, N., MacNee, W., Robinson, S. and Newby, D. (2005) Role of Inflammation in Cardiopulmonary Health Effects of PM. *Toxicology and Applied Pharmacology*, **207**, 483-488. <http://dx.doi.org/10.1016/j.taap.2005.02.020>
- [5] Duffin, R., Mills, N.L. and Donaldson, K. (2007) Nanoparticles—A Thoracic Toxicology Perspective. *Yonsei Medical Journal*, **48**, 561-572. <http://dx.doi.org/10.3349/ymj.2007.48.4.561>
- [6] Mills, N.L., Miller, M.R., Lucking, A.J., Beveridge, J., Flint, L., Boere, A.J.F., Fokkens, P.H., Boon, N.A., Sandstrom, T., Blomberg, A., Duffin, R., Donaldson, K., Hadoke, P.W.F., Cassee, F.R. and Newby, D.E. (2011) Combustion-Derived Nanoparticulate Induces the Adverse Vascular Effects of Diesel Exhaust Inhalation. *European Heart Journal*, **32**, 2660-2671. <http://dx.doi.org/10.1093/eurheartj/ehr195>
- [7] Rose, G. and Marmot, M.G. (1981) Social Class and Coronary Heart Disease. *British Heart Journal*, **45**, 13-19. <http://dx.doi.org/10.1136/hrt.45.1.13>
- [8] Marmot, M. (2004) Status Syndrome. *Significance*, **1**, 150-154. <http://dx.doi.org/10.1111/j.1740-9713.2004.00058.x>

- [9] Hubbell, B. (2015) Workshop to Define Approaches to Assess the Effectiveness of Policies to Reduce PM2.5. [http://www.cleanairinfo.com/finepmpolicy/Webinar/webinar\\_video\\_P1\\_Hubbell.html](http://www.cleanairinfo.com/finepmpolicy/Webinar/webinar_video_P1_Hubbell.html)
- [10] Zigler, C.M. and Dominici, F. (2014) Point: Clarifying Policy Evidence with Potential-Outcomes Thinking—Beyond Exposure-Response Estimation in Air Pollution Epidemiology. *American Journal of Epidemiology*, **180**, 1133-1140. <http://dx.doi.org/10.1093/aje/kwu263>
- [11] Vanderweele, T.J. and Vansteelandt, S. (2009) Conceptual Issues Concerning Mediation, Interventions and Composition. *Statistics and Its Interface*, **2**, 457-468. <http://dx.doi.org/10.4310/SII.2009.v2.n4.a7>
- [12] Robins, J.M. and Greenland, S. (1992) Identifiability and Exchangeability for Direct and Indirect Effects. *Epidemiology*, **3**, 143-155. <http://dx.doi.org/10.1097/00001648-199203000-00013>
- [13] CDC (2015) Interactive Atlas of Heart Disease and Stroke. [Online]. <http://nccd.cdc.gov/dhdspatlas/>
- [14] EPA (2015) AirData. [Online]. [http://aqsdrl.epa.gov/aqsweb/aqstmp/airdata/download\\_files.html#Daily](http://aqsdrl.epa.gov/aqsweb/aqstmp/airdata/download_files.html#Daily)
- [15] Wackernagel, H. (2013) *Multivariate Geostatistics: An Introduction with Applications*. Springer, Berlin.
- [16] Atkinson, R.W., Bremner, S.A., Anderson, H.R., Strachan, D.P., Bland, J.M. and de Leon, A.P. (1999) Short-Term Associations between Emergency Hospital Admissions for Respiratory and Cardiovascular Disease and Outdoor Air Pollution in London. *Archives of Environmental Health*, **54**, 398-411. <http://dx.doi.org/10.1080/00039899909603371>
- [17] Wong, T.W., Tam, W.S., Yu, T.S. and Wong, A.H.S. (2002) Associations between Daily Mortalities from Respiratory and Cardiovascular Diseases and Air Pollution in Hong Kong, China. *Occupational and Environmental Medicine*, **59**, 30-35. <http://dx.doi.org/10.1136/oem.59.1.30>
- [18] Crapo, J.D., Broaddus, V.C., Brody, A.R., Malindzak, G., Samet, J. and Wright, J.R. (2003) Workshop on Lung Disease and the Environment: Where Do We Go from Here? *American Journal of Respiratory and Critical Care Medicine*, **168**, 250-254. <http://dx.doi.org/10.1164/rccm.168.2.250>
- [19] Pope, C.A., Schwartz, J. and Ransom, M.R. (1992) Daily Mortality and PM10 Pollution in Utah Valley. *Archives of Environmental Health*, **47**, 211-217. <http://dx.doi.org/10.1080/00039896.1992.9938351>
- [20] Jerrett, M., Gale, S. and Kontgis, C. (2010) Spatial Modeling in Environmental and Public Health Research. *International Journal of Environmental Research and Public Health*, **7**, 1302-1329. <http://dx.doi.org/10.3390/ijerph7041302>
- [21] Glass, T.A., Goodman, S.N., Hernán, M.A. and Samet, J.M. (2013) Causal Inference in Public Health. *Annual Review of Public Health*, **34**, 61-75. <http://dx.doi.org/10.1146/annurev-publhealth-031811-124606>
- [22] Noffsinger, M.A., Pfefferbaum, B., Pfefferbaum, R.L., Sherrib, K. and Norris, F.H. (2012) The Burden of Disaster: Part I. Challenges and Opportunities within a Child's Social Ecology. *International Journal of Emergency Mental Health*, **14**, 3-13.
- [23] Morgenstern, H. and Thomas, D. (1993) Principles of Study Design in Environmental Epidemiology. *Environmental Health Perspectives*, **101**, 23-38.
- [24] Anderson, F., Carson, A., Whitehead, L. and Burau, K. (2014) Spatiotemporal Analysis of the Effect of Ozone and Fine Particulate on CVD Emergency Room Visits in Harris County, Texas. *Open Journal of Air Pollution*, **3**, 87-99. <http://dx.doi.org/10.4236/ojap.2014.34009>
- [25] Smyth, J.D. (1988) Multivariate Geostatistical Analysis of Groundwater Contamination by Pesticide and Nitrate.
- [26] Anderson, F. (2014) Multivariate Geostatistical Model for Groundwater Constituents in Texas. *International Journal of Geosciences*, **5**, 1609-1617. <http://dx.doi.org/10.4236/ijg.2014.513132>
- [27] Goovaerts, P. (2006) Geostatistical Analysis of Disease Data: Accounting for Spatial Support and Population Density in the Isopleth Mapping of Cancer Mortality Risk Using Area-to-Point Poisson Kriging. *International Journal of Health Geographics*, **5**, 52. <http://dx.doi.org/10.1186/1476-072X-5-52>
- [28] Kulldorff, M., Tango, T. and Park, P.J. (2003) Power Comparisons for Disease Clustering Tests. *Computational Statistics & Data Analysis*, **42**, 665-684. [http://dx.doi.org/10.1016/S0167-9473\(02\)00160-3](http://dx.doi.org/10.1016/S0167-9473(02)00160-3)
- [29] Mantel, N. (1967) The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research*, **27**, 209-220.
- [30] Rothman, K.J. (1987) Clustering of Disease. *American Journal of Public Health*, **77**, 13-15. <http://dx.doi.org/10.2105/AJPH.77.1.13>
- [31] Song, C. and Kulldorff, M. (2003) Power Evaluation of Disease Clustering Tests. *International Journal of Health Geographics*, **2**, 9. <http://dx.doi.org/10.1186/1476-072X-2-9>
- [32] Lawson, A. (2010) Statistical Methods for Disease Clustering. *Annals of Epidemiology*, **20**, 964. <http://dx.doi.org/10.1016/j.annepidem.2010.07.101>