

반복측정 분석을 통한 미세먼지(PM₁₀)와 고혈압 유병률의 상관관계 분석: 2009-2013 지역사회 건강조사 자료를 이용하여

이환희¹, 김호^{1*}

¹서울대학교 보건대학원 보건학과, 서울대학교 보건환경연구소

Association between PM₁₀ and Prevalence of Hypertension: Using Longitudinal Analysis with 2009-2013 Community Health Survey Data

Whanhee Lee¹ and Ho Kim^{1*}

¹Department of Biostatistics and Epidemiology, Graduated School of Public Health, & Institute of Health and Environment, Seoul National University.

Abstract

Particular matter has been linked to hypertension, but most previous studies have reported on the association between hypertension incidence or death and particular matter. This study aimed to estimate the epidemiological association between yearly concentration of particular matter less than 10 μ m (PM₁₀) and the regional prevalence of hypertension in South Korea. This cross-sectional study used data on the regional prevalence of hypertension from the Korea Community Health Survey conducted in 7 major cities (with 70 communities) of South Korea in 2009-2013. PM₁₀ data were collected from the National Institute of Environmental Research. We used generalized linear mixed model to represent the correlation among repeated observation. We found that a significant odds ratio of hypertension of 1 μ g/m³ in PM₁₀ was 1.006 (95% C.I. 1.003, 1.009). And also found the significant lagged effect of PM₁₀ on hypertension prevalence during 5 years and the highest association at 2 years after exposure.

keywords: Particular matter, Hypertension, Air pollution, Longitudinal analysis, Generalized linear mixed model, Community health survey

Introduction

고혈압은 현재 한국의 주요 만성 질병이다. 고혈압은 일반적으로 나이, 비만, 흡연, 스트레스 등 다양한 원인으로 인하여 발생하며, 최근 연구 결과에 따르면 만성 신부전증의 16%가 고혈압에 의한 것으로 나타날 만큼 다른 만성 질환의 원인이 되는 질환이다[1]. 지역사회 건강조사 조사결과에 따르면(Table 1), 2009년의 한국의 7개

광역시(서울특별시, 부산광역시, 인천광역시, 대전광역시, 대구광역시, 광주광역시, 울산광역시)의 고혈압 유병률의 평균은 13.78%로 다른 질환에 비하여 유병률이 매우 높으며, 2013년 유병률은 14.90%로 매년 증가 추세에 있다. 따라서 고혈압의 원인을 밝혀내고, 그에 따른 원인 인자 제거를 통하여 고혈압 유병률 낮춰야 할 필요성이 강하게 제기되고 있다. 본 연구에서는 다양한 고혈압의 위험 인자 중 하나인 직경 10 μ m 이하

* Corresponding author: Ho Kim, Ph.D. (hokim@snu.ac.kr., 02-880-2702)

Department of Biostatistics and Epidemiology, School of Public Health, Seoul National University, 1 Gwanak-Ro, Gwanak-gu, Seoul 151-752, South Korea.

의 미세먼지(PM₁₀)를 분석하였다.

많은 연구들은 직경 10 μm 이하의 미세먼지 (PM₁₀)과 심장질환, 특히 고혈압과의 연관성을 보여주고 있다. 1999년 Pope의 연구 결과에서는 PM₁₀이 100μm 증가할 때 분당 심장 박동수가 평균 0.8회 이상 증가함을 보여주고 있다[2]. 또한 미세 먼지가 심박변화율(Heart Rate Variability, HRV)에 영향을 준다는 연구 결과도 존재한다. Gold et al.[3]은 연구에서 미세먼지가 심장의 자율기능(autonomic function)에 연관성이 있음을 보여준다. 또한 PM₁₀보다 작은 PM_{2.5}가 HRV에 유의한 영향을 주는 연구 결과도 있다[4].

PM₁₀은 이질적인 고체와 액체 먼지의 조합으로 구성되는데, 이는 자동차의 분진이나, 길의 먼지 또는 공장의 오염물질 배출, 철강 재련, 농업, 화전농업 등에서 생긴다[5]. 특히 자동차가 많고, 공장이나 생활에서 오염물질이 많은 광역도시에서는 PM₁₀의 영향이 다른 도시에 비하여 크고 중요하다. 그에 따라 본 연구에서 사용된 PM₁₀ 자료는 2009년 1월 1일부터 2009년 12월 31일까지 한국의 7개 광역도시(서울특별시, 부산광역시, 인천광역시, 대전광역시, 대구광역시, 광주광역시, 울산광역시)의 지역구 중 70개 구의 연평균 PM₁₀과 2009년부터 2013까지 매년 반복해서 측정된 고혈압 유병률과의 연관성을 분석하였다.

Method

Data

PM₁₀ 자료는 7개 광역시 70개 구에 있는 대기오염측정소에서 측정한 자료를 사용하였다. 총 8760번의 측정에서 7000번 이하의 측정을 보인 측정소는 제외하였으며, 구 별로 2개 이상의 측정소가 있는 경우에는 산술평균을 내어 사용하였다. 보정변수로 사용된 구 단위 연평균 온도 자료는 한국 기상청(Korea Meteorological Administration) 자료를 사용하였다.

구별 고혈압 유병률은 지역사회 건강조사(Community health survey)를 통하여 측정된 자료를 사용하였다. 지역사회 건강조사는 2008년부터 매년 시행되고 있으며, 지역사회 단위의 주민의 건강 수준을 측정하는 사업이며, 복합표본설계를 통하여 조사과정에서 발생할 수 있는 편차를 조정하고 있다. 본 연구에서 사용된 구별 고혈압 유병률은 1개 구 당 약 900명 정도의 표본조사자 중 고혈압 평생 의사진단 경험에서 ‘예’를 택한 조사자의 비율을 복합표본설계를 통한 가중치를 통하여 구한 것으로, 성과 연령이 보정된 표준화율이다. 그리고 보정변수로 사용된 월간음주율, 현재흡연율, BMI(Body Mass Index) 또한 지역사회 건강조사를 통하여 조사된 구 별 표준화율을 사용하였다(Table 1).

Table 1. Descriptive statistics of variables.

Variables	Time	N	Mean	S.D.	Min	Max
PM ₁₀ (μg/m ³)	2009	70	51.2	6.2	37.5	11.5
Annual Average temperature (%)	2009	70	13.4	0.9	11.6	15.0
Body Mass Index (BMI) (%)	2009	70	23.0	0.6	22.5	26.2
Rate of Drink Month (%)	2009	70	57.6	3.7	49.2	66.0
Rate of Smoking now (%)	2009	70	26.0	2.4	20.4	31.6
Prevalence of rate of hypertension (%)	2009	70	13.7	1.5	10.2	17.7
	2010	70	14.5	1.6	10.4	17.4
	2011	70	14.7	1.6	11.5	18.4
	2012	70	14.9	1.5	11.9	19.4
	2013	70	14.9	1.5	11.6	19.1

Statistical Analysis

본 연구에서는 종단형 반복측정 자료분석을 위해 일반화 선형 혼합 모형(Generalized Linear Mixed Model, GLMM)을 사용하였다. 일반적인 일반화 선형 모형(Generalized Linear Model)에서는 관측치 간의 독립성을 가정하지만, 반복측정 자료는 최소 개인 단위의 상관성이 존재하기 때문에 독립성 가정에 위배된다. 그렇기 때문에 반복 측정된 자료분석을 할 때는 반드시 개별 연관성에 대한 상관성을 고려하여야 한다[6]. 본 연구에서의 70개 구의 개별 연관성을 고려하여 분석하였다(Figure 1).

이러한 상관성을 고려하기 위해서는 반복 측정된 시간에 따른 분산-공분산 구조(Working Correlation Matrix)를 결정하여야 한다. 본 연구에서는 반복측정 자료에서 많이 사용되는 3가지 Ante-dependence (ante(1)), Heterogeneous Autoregressive(1) (Arh(1)), Autogressive(1) (Ar(1)) 분산-공분산 구조를 사용하였으며, pseudo-AICc 값을 기준으로 Ar(1) 구조를 채택하여 사용하였다.

Autoregressive structure

$$cov(\xi_i, \xi_j) = \sigma^2 \rho^{|i-j|}, \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix}$$

분산 공분산 구조를 고려한 PM_{10} 과 고혈압 유병률과의 연관성을 분석하기 위하여 사용한 선형 식은 다음과 같다.

$$\text{logit}(y) = PM_{10} + time_k + Covs + s_{ik},$$

$$\text{logit}(y) = PM_{10}|time_k + Covs + s_{ik}$$

$$y|s_{ik} \sim B(n, \pi), s_{ik} \sim MVN(0, \sigma^2), \text{var}(y|s_{ik}) = V^{1/2} A_w V^{1/2}$$

y = 고혈압 유병률 ($0 \leq y \leq 1$), i = 70개 구, k = time(2009 - 2013), A_w = Working Correlation matrix,

Covs = 보정변수(연평균 온도, BMI, 월간음주율, 현재흡연율)

본 연구에서는 Conditional distribution을 이용한 GEE(Generalized Estimating Equation)-type의 GLMM분석을 수행 하였으며, 유병률은 0에서 1사이의 확률로 하여 logit함수를 연결함수로 사용하였다. 이를 통하여 PM_{10} 의 1_{μm} 증가에 따른 Odds Ratio(OR)와 OR의 95% 신뢰구간(95% CI)를 추정하였다. Likelihood는 분산의 과소 추정을 방지하여 위하여 pseudo likelihood를 기반으로 한 Restricted Maximum Likelihood(REML)을 사용하였으며, 분산의 편의성을 보정하기 위하여 Kenward-Rodger 보정을 사용하였다. 또한 Random effect의 필요성을 검정하기 위해서 pseudo-likelihood test를 사용하였다.

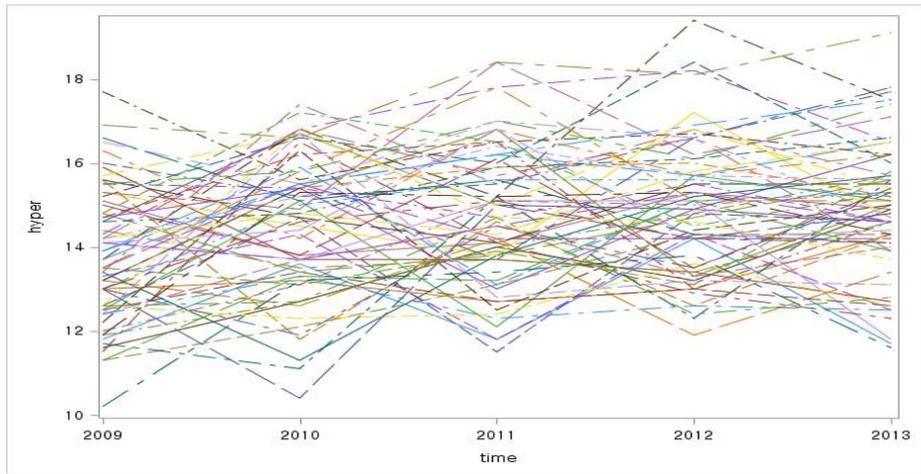


Figure 1. Profile plots of prevalence rate of hypertension

Results

GLMM에 따른 분석 결과, 시간과 PM₁₀의 교호작용은 고려하지 않고 측정된 PM₁₀의 OR은 1.003(95% CI [1.001,1.006])으로 유의하게(p-value=0.014 <0.05) 추정되었다(Table3). 다시 말해 PM₁₀이 1 μ m 증가할 때마다 고혈압 유병률의 OR가 1.003만큼 증가함을 의미한다. Random Effect (s_{ik})의 필요성을 검정하기 위해 실행한 pseudo-LR test 결과 random effect는 필요한 것으로 나타났다($\chi^2 = 17.83, df = 1, p - value = <.001$).

Table 2. Odds Ratio (OR) of hypertension associated with 1 μ m increase of PM₁₀

	Odds Ratio	95% C.I.	
Unadjusted model	1.006	1.003	1.009
Adjusted* Model	1.003	1.001	1.006

* Adjusted for temperature, drinking, and smoking

그리고 시간에 따라 PM₁₀이 고혈압 유병률에 미치는 영향이 다른가를 확인하기 위하여 시간(time)과 PM₁₀의 교호작용을 고려한 선형 모형을 분석해본 결과, PM₁₀의 영향은 유의하게 PM₁₀ * time의 교호작용 약하게 유의한 것으로 나타났다(Table4). 특히 2009-2010년에는 PM₁₀의 효과가 유의하지 않다가, 2011년에 급격하게 효과가 유의하게 증가하며(Figure2), 2012-2013에서도 유의하게 PM₁₀의 효과가 나타남을 볼 수 있다(Table5).

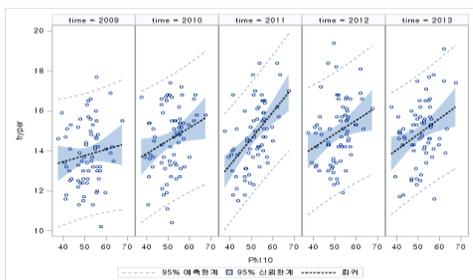


Figure 2. Regression plot by year(2009-2013)

나아가 PM₁₀을 보정하기 위하여 사용한 보정 변수들 또한 고혈압 유병률과 유의한 관계를 나타내고 있다(Figure3). 연평균 온도는 OR= 0.938(95%CI [0.922,0.954]), 월간음주율 OR= 1.005(95%CI [1.001,1.009]), 현재흡연율 OR= 1.007 (95%CI [1.000,1.013])의 관계를 보였다. 이는 온도가 추울수록, 음주와 흡연 비율이 높을수록 고혈압 유병률이 올라감을 보여준다.

Table 3. Odds ratio (OR) between hypertension and PM₁₀ by year

Year	Odds Ratio	DF	95% Confidence Interval	
2009	1.000	291.9	0.996	1.004
2010	1.002	291.5	0.998	1.007
2011	1.008*	291.4	1.004	1.012
2012	1.003*	291.3	0.998	1.007
2013	1.003*	291.3	0.999	1.008

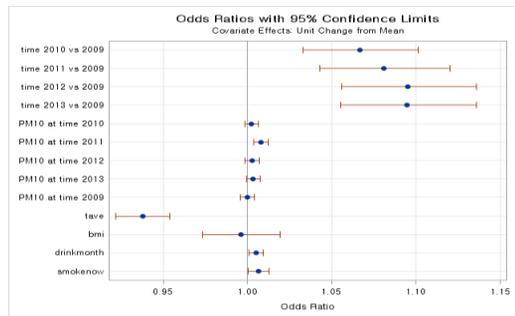


Figure 3. Odds Ratio(OR) for PM₁₀ and covariates

Conclusion

2009년부터 2013년도까지 지역사회 건강조사 자료를 바탕으로 고혈압 유병률과 PM₁₀의 관계에 대한 분석 결과 일반적으로 알려져 있는 관계와 마찬가지로[7], PM₁₀은 고혈압 유병률과 양의 상관관계를 보였다. 주목할 점은 본 연구에서는 PM₁₀의 장기간 효과(Long Term Effect)를

확인했다는 점이다. 2009년에 측정된 PM₁₀ 을 5년간의 고혈압 유병률 자료에 대응하여 분석한 결과 1 μ m의 PM₁₀ 이 증가할 때마다 OR 1.003(95% CI [1.001,1.006])이 됨을 알 수 있었다.

또한 이번 연구에서는 시간과 PM₁₀ 의 교호작용을 확인할 수 있었다. 2009년에 측정된 PM₁₀ 은 2009년과 2010년에는 고혈압 유병률에 유의한 영향을 미치지 않았지만, 2011년에 급격하게 영향이 증가하면서 2013년도까지 영향을 미쳤다. 이 결과는 PM₁₀가 장기간 효과가 있다 라는 결과에서 더 나아가 구체적인 기간적 영향을 보여준다. Figure2에서 보듯 PM₁₀은 2년이 지난 시점에서부터 고혈압 유병률에 영향을 준다. 조금 더 연구가 필요한 부분이겠지만, 이것은 고혈압 유병률에는 PM₁₀이 2년에서 5년 정도의 lag-effect를 가지고 있음을 시사할 수 있다. 일반적인 연구에서는 Daily Time Series를 이용하여 하루라는 단위시간으로 30일 정도의 lag-effect를 추정한다. 이러한 방법은 급성 질환이나 사망 등에는 유용할 수 있어도, 만성질환의 유병률과 같은 변수에서는 적용하기 어렵다. 본 연구는 이러한 단점을 극복하고자 annual 간격으로 반복측정 자료분석을 실행하였으며, 2-5년이라는 PM₁₀과 유병률의 lag-effect가 있을 수 있음을 찾아내었다.

본 연구에서는 70개의 구를 기본 단위로 하여 분석을 진행하였는데, 그 과정에서 각 7개 광역도시(서울특별시, 부산광역시, 인천광역시, 대전광역시, 대구광역시, 광주광역시, 울산광역시)의 차이는 고려하지 않았다. 다시 말해, 7개의 광역도시의 nested 효과를 모수로 추정하지도 않았고, random effect로 넣지도 않았다. ANOVA table과 pseudo likelihood를 이용한 LR test를 해 본 결과 7개 광역시간의 분산 차이는 존재하지 않았다 (pseudo LRT, p-value= 0.6113). 이는 광역도시 간의 고혈압 유병률의 특징이 비슷하기 때문에 생기는 결과라고 추측된다. 차후 연구에서 전국 대상 혹은 도시 농촌 지역간 비교가 필요한 연구에서는 지역 간의 분산이 다를 가능성이 높음으로 구를 기본적인 분석 단위로 하되, 구를 포함하고 있는 시-도의 영향을 추정해주는 것이 필요할 것으로 생각된다.

그리고 본 연구에서 사용한 pseudo-LR test는 원칙적으로는 분산 구조를 선택하거나, random effect의 유무에 사용하는 것이 권장되지 않는다

[8]. Pseudo likelihood란 GLMM 구조에서 계산할 수 없는 marginal function을 정규분포를 따른다는 가정하에 선형 식으로 바꿔서 추정하는 방식을 의미한다. 선형으로 바꾸는 Pseudo 방식과는 달리 몇 적분 지점을 지정하여 Numerical Integration을 하는 Laplace Integration 방식이나 Gauss-Hermite quadrature 방식을 사용하면 marginal function을 계산할 수 있기 때문에 LRT test를 사용할 수 있다. 그렇지만 본 연구에서는 Numerical Integration을 수행할 경우 근이 수렴하지 않는 문제점이 지속적으로 발생하였고, 또한 Beta를 추정하는 데에는 큰 문제점이 없기 때문에 GEE-type의 pseudo likelihood를 사용하였다.

많은 연구에서 드러나듯 PM₁₀은 고혈압과 고혈압과 연관된 질환에 부정적인 연관관계를 가지고 있다. 이제까지 PM₁₀과 고혈압 질환의 발생(Onset) 혹은 사망(Mortality)의 연관성에 대한 분석은 많이 이루어져 왔지만, 방법적 한계와 자료 구조의 문제로 유병률과의 연관성 연구는 많이 존재하지 않았다. 생물학적으로도 PM₁₀과 고혈압 유병과는 연관관계가 존재하기 때문에 더욱 활발하게 PM₁₀과 고혈압 유병률과의 원인적 인과관계에 대해 분석하여야 하고, 그 분석을 토대로 정확한 PM₁₀의 영향을 추정해야 할 필요성이 있다. 나아가 이를 바탕으로 장기간에 걸쳐 PM₁₀이 고혈압 유병률에 미치는 영향을 줄이거나, 혹은 고혈압 예방을 위하여 PM₁₀의 노출을 낮추는 방향으로 과학적-정책적 연구가 필요할 것으로 사료된다.

Acknowledgement

이 논문은 보건연구재단의 지원으로 수행되었습니다.

References

[1]국가건강포털 만성신부전증. [Cited 2014 Oct 15.]; Avliailable from: URL :http://health.mw.go.kr/HealthInfoArea/HealthInfo/View.do?idx=2200

[2] Pope CA, Dockery DW, Kanner RE, Villegas GM, Schwartz J. Oxygen saturation, pulse rate, and particulate air pollution : A daily time-series panel study. *Am J Respir Crit Care Med.* 1999; 159:365-

372.

[3] Gold DR, Litonjua A, Schwartz J, et al. Ambient pollution and heart rate variability. *Circulation*. 2000; 101:1267-1273.

[4] Sung Kyun Park, Marie S. O'Neill, Pantel S. Vokonas, David Sparrow, Joel Schwartz. Effects of air pollution on heart rate variability : The VA normative aging study.

[5] Robert D. Brook. Barry Franklin, et al. Air pollution and cardiovascular disease : A statement for healthcare professionals from the expert panel on population and prevention science of the American heart association. *Circulation*. 2004; 109:2655-2671.

[6] Nan ML, James HW. Random-Effects Models for Longitudinal Data. *BIOMETRICS*. 1982; 38, 963-974.

[7] Angela Ibal-Mulli, Jutta Stieber, H.-Erich Wichmann, Wolfgang Koenig, and Annette Peters. Effect of Air Pollution on Blood Pressure. *Am J Public Health*. 2001;91:571-577)

[8] Walter W. Stroup. Generalized Linear Mixed Models : Modern concepts methods and applications. CRC press ; 2012. 179-189