Environmental Research 159 (2017) 531–538

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Does a lag-structure of temperature confound air pollution-lag-response relation? Simulation and application in 7 major cities, Korea (1998–2013)

Honghyok Kim[a], Michelle L. Bell[b], Jong-Tae Lee[a,c,d,*]

[a] BK21PLUS Program in 'Embodiment: Health –Society Interaction', Department of Public Health Science, Graduate School, Korea University, Seoul, Republic of Korea
[b] School of Forestry and Environmental Studies, Yale University, New Haven, CT, United States
[c] School of Health Policy and Management, College of Health Science, Korea University, Seoul, Republic of Korea
[d] Department of Environmental Health, Korea University, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
PM_{10}
Temperature
Mortality
Lag effects
Simulation study

ABSTRACT

Background: Temperature must be controlled when estimating the associations of short-term exposure to air pollution and mortality. Given that multi-country studies have implied temperature has lagged effects, we aim to explore confounding by temperature-lag-response and investigate PM_{10}-lag-mortality relation in 7 cities, Korea.
Methods: In a simulation study, we compared the performance of different methods to control for: the same day temperature, a lagged temperature and distributed lags of temperature. In a real data study, we explored PM_{10}-lag-mortality relation in 7 cities using these different methods.
Results: We confirmed that a model with insufficient control of temperature offers a biased estimate of PM_{10} risk. The degree of bias was from −82% to 95% in simulation settings. A real data study shows estimates among different models by temperature adjustments and PM_{10} lag variables ranging from −0.3% to 0.4% increase in the risk of all-cause mortality, with a 10 $\mu g/m^3$ increase in PM_{10}. Controlling for temperature as distributed lags for 21 days provided 0.25% (95% CI: 0.1, 0.4) increase in the risk of all-cause mortality.
Conclusions: A lag structure of temperature can confound the air pollution-lag-response relation. Temperature-lag-response relation should be evaluated when estimating air pollution-lag-response relation. As a corollary, air pollution and temperature risk in mortality can be estimated using the same regression model.

1. Introduction

Air pollution time-series studies in APHENA (Air Pollution and Health: a European and North American Approach) and PAPA (Public health and Air Pollution in Asia) have contributed to confirming the relation between short-term exposure to air pollution and mortality. (Anderson et al., 2004; Bell et al., 2004; Dominici et al., 2005; Peng et al., 2013; Samet et al., 2000; Samoli et al., 2008; Wong et al., 2008) A shared protocol of these studies includes adjustment for potential confounding according to daily temperature. To examine whether temperature confounds the air pollutant-mortality association, different adjustment methods for daily temperature of the same day or previous days should be compared.

The relation of short-term exposure to temperature with mortality has also been studied over the past few decades. Recent multi-country studies (Gasparrini et al., 2015; Guo et al., 2014; McMichael et al., 2008) and methodological studies (Armstrong, 2006; Gasparrini et al., 2010) have revealed that effects of short-term exposure to temperature on mortality persist over long lags.

In light of the necessary adjustment of daily temperature in an air pollution time-series study, updated information on temperature-mortality association raises questions on potential confounding effects of a lag-structure of temperature. When estimating the air pollution-mortality association, if the temperature is causally associated with mortality in long lags, it would be desirable to adjust for temperature of such a lag period because the lag-structure of temperature is correlated with air pollution and is not located in an intermediate pathway between air pollution and mortality. In particular, this potential confounding would be more critical when probing for lagged associations of an air pollutant with mortality. Although previous researchers have coped with various adjustment methods for temperature and seasonal confounding in the National Morbidity Mortality Air Pollution Study (Welty and Zeger, 2005), to the best of our knowledge, theoretical and empirical evidence of confounding of a lag-structure of temperature is

---

Abbreviations: DLNM, distributed lag non-linear model; df, degrees of freedom; NCS, natural cubic spline; MMT, minimum mortality temperature; RMSEc, square root of the mean squared error of the cumulative effect estimator

[*] Correspondence to: #373 Hana Science Hall B, Korea University, Seongbuk-gu, Seoul 02841, Republic of Korea.

E-mail address: jtlee@korea.ac.kr (J.-T. Lee).

http://dx.doi.org/10.1016/j.envres.2017.08.047
Received 20 June 2017; Received in revised form 26 July 2017; Accepted 24 August 2017

Table 1
The number of population and daily mortality rates in 7 cities, Korea, 1998–2013.

City	Population size (from Census 2010)	Total of death counts	Crude annual mortality rate (per 100,000)
Seoul	9631,482	552,213	358.3
Busan	3393,191	269,162	495.8
Daegu	2431,774	163,559	420.4
Incheon	2632,036	146,153	347.1
Gwangju	1466,143	85,047	362.5
Daejeon	1490,158	82,787	347.2
Ulsan	1071,673	58,480	341.1
Total	22,116,457	1357,401	383.6

still limited.

In this study, we aimed to investigate confounding by temperature-lag-mortality on the PM_{10}-lag-mortality association in 7 cities in Korea, with theoretical evidence through a simulation study, and empirical evidence of confounding when estimating.

2. Methods

2.1. Data

For 7 major cities in Korea (Seoul, Busan, Incheon, Daegu, Gwangju, Daejeon and Ulsan), we obtained 16 years (1998–2013) of mortality data from Statistics Korea and daily temperature and relative humidity data from Korea Meteorological Administration. Measurement data of particulate matter with aerodynamic diameter less than 10 μm (PM_{10}) was secured from National Institute of Environmental Research. We calculated city-specific daily mortality counts using individual records of mortality data. All-cause mortality was categorized as A00-R99 of International Classification of Disease 10th revision. The population size of the 7 cities from Census 2010 of Statistics Korea and the crude daily mortality rate are listed in Table 1.

City-specific daily 24-h average of PM_{10} was calculated from hourly averages of measurements of different monitoring stations. There are 25, 16, 15, 11, 7, 8 and 13 of monitoring stations in Seoul, Busan, Incheon, Daegu, Gwangju, Daejeon, and Ulsan respectively. Each of district in a city has at least one monitoring station at the center of the district: 25, 16, 10, 8, 5, 5, and 5 of districts respectively. We calculated hourly averages from the available data, then calculated 24-h values for each city. The daily mean temperature and relative humidity were calculated from 3-h interval measurements from a single monitoring station in each city. If the hourly averages of PM_{10} had missing values of more than 25% (6 out of 24), we treated a daily mean PM_{10} calculated from those hourly averages as missing. The missing values were imputed with the annual average. The number of city-specific missing values out of 5844 days was 0 (0%), 7 (0.12%), 3 (0.05%), 28 (0.48%), 10 (0.17%), 141 (2.41%), and 28 (0.48%) in Seoul, Busan, Incheon, Daegu, Gwangju, Daejeon and Ulsan. For temperature and relative humidity, there were no missing values. Since all 7 cities are under the influence of 'Yellow Dust' that is also known as 'Asian Dust' from the Gobi Desert in northern and northwestern China, we changed PM_{10} values over 99th percentile to the value of 99th percentile, to reduce the bias from unrepresentative high concentrations.(Kim et al., 2015). Pearson correlation (p-value) between temperature and PM_{10} over the extreme values was minimal: −0.15 (0.24), −0.25 (0.05), −0.10 (0.46), 0.10 (0.46), −0.12 (0.35), 0.05 (0.71), and −0.06 (0.67) in Seoul, Busan, Incheon, Daegu, Gwangju, Daejeon and Ulsan respectively.

2.2. Simulation settings

We simulated the confounding of a lag-structure of temperature on PM_{10}-mortality association in realistic settings. For this, we first estimated temperature-lag-mortality association in Seoul for only 8 years (2003–2011). Data of the influenza epidemic was available only for that period. Using 16 years of data requires substantial computation time for a simulation. In light of published air pollution time-series studies, 8 years of data seem to be adequate. Since daily temperature is non-linearly associated with daily mortality (Gasparrini et al., 2015; Guo et al., 2014), a distributed lag non-linear model (DLNM) (Gasparrini et al., 2010) with a Poisson distribution was applied. The DLNM contained constrained distributed lag variables, which is the cross-basis for a cubic B-spline of temperature with second degree and city-specific knots of 25th, 50th, 90th percentile of temperature and for a cubic B-spline of lags with third degree and 7 degrees of freedom (df). We tested different maximum lags of temperature: 14, 21 and 30 days. To adjust for a temporal trend, a natural cubic spline (NCS) of time with different df 4–10/year) was used: Hereafter we refer to them as NCS(4), … NCS(10) respectively. The day of the week and holidays were also adjusted. Of the aforementioned 18 models (3 lag types * NCS 4–9) = 18), the temperature-mortality association in Seoul was consistently J-shaped as shown in Fig. 1 [See Fig. S1 for detail]. NCS(10) alone showed smaller associations of temperature with mortality. This model seemed to be over-fitted. Residual (partial) autocorrelation was ignorable in NCS(5, to 10). A confidence interval for temperature range below minimum mortality temperature (MMT) was relatively wide: MMT is temperature where estimated mortality risk is the minimum over temperature distribution. MMT was 26.3 °C, which is the 93rd percentile of the temperature. Fig. 2 shows lagged associations for both heat and cold temperature range. At high temperatures, there were two peaks of positive associations over lags: lag0 and lag4-8. At cold temperature, we found a negative association at lag0 and a peak of positive associations at lag2. We selected T_{dl021} with NCS(6) for the subsequent simulation study.

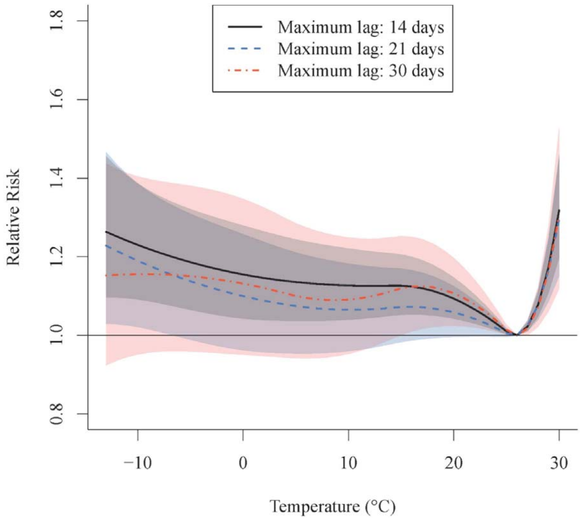

Fig. 1. Cumulative temperature-mortality associations from three different distributed lag non-linear models in Seoul for 8 years (2003–2011).

After fitting a Poisson regression model for temperature-mortality association, we added hypothetical PM_{10} effect onto predicted daily death counts ($\hat{y}_t$). A product between a real PM_{10} measurement at day t and a beta coefficient specified by PM_{10} effect scenarios was added to a logarithm of $\hat{y}_t$. Then, it was exponentiated to a new predicted death counts (y_t^*). Three PM_{10} effect scenarios were specified as follows. The cumulative effect size was similarly specified according to our previous research (i.e. 0.00012–0.00044 per 1 $\mu g/m^3$ increase in all-cause and cardiovascular mortality).(Kim et al., 2015)

Scenario 1. A single effect from the concurrent day: $\beta_{lag0} = 0.0003$

Scenario 2. Effects from the concurrent day to the two previous days with 0.0003 of the cumulative effect ($\bar{\beta}$):