EXPLORING THE SHORT-TERM EFFECTS OF PM_{10} EXPOSURE ON NATURAL MORTALITY: A GPS-BASED MATCHING APPROACH IN THREE EUROPEAN CITIES

Marzi Chiara[1], Biscardi Giulio[1], Baccini Michela[1]

[1]*Dipartimento di Statistica, Informatica e Applicazioni "Giuseppe Parenti", Università di Firenze, Firenze*

Introduction

Since the year 2000, several epidemiological studies have examined the short-term impact of air pollution on health, understanding the causal relationship between exposure and response. The potential outcome (PO) approach to causal inference promotes a conceptual framework based on causes and consequences within a formal mathematical framework [1,2]. Although it has gained popularity in various fields, it is relatively novel in studies focusing on evaluating the effects and impacts of air pollution on health. Within a PO approach to causal inference, the propensity score (PS) has been introduced for confounding adjustment in observational studies with binary exposure [3]. A few years later, Hirano and Imbens made further advancements by extending the generalized PS (GPS) to the context of continuous exposure [4]. Subsequently, several studies in the literature have employed methods based on GPS for covariates adjustment, requiring the parametric or semi-parametric specification of a model for treatment and a model for the outcome [5,6]. Recently, Wu and colleagues proposed a GPS matching approach that, under a local ignorability assumption, jointly matches on both the estimated scalar GPS and exposure levels to adjust for confounding bias [7]. They estimated the exposure-response function that characterizes the long-term impact of the exposure to particulate matter with a diameter less than or equal to 2.5 μm ($PM_{2.5}$) on natural mortality in a large observational cohort in the United Stated (2000-2016).

Aims

In this study, we use the GPS-based matching method [7], to estimate the average causal dose-response function (aDRF) describing the short-term effect of the exposure to PM with a diameter less than or equal to 10 μm (PM_{10}) on natural mortality, across three European cities.

Methods

Data

This study examined the impact of short-term exposure to PM_{10} on natural mortality in three European cities: Milan (Italy), Skopje (North Macedonia), and Donostia - San Sebastián (Spain). For Milan, we obtained daily time series data of PM_{10} levels, temperature, and humidity from the Regional Environmental Protection Agency and daily mortality data from the Regional Mortality Register, for the years 2003-2006. For Skopje, our data covers the ten districts within the municipality of Skopje, excluding Sopište, which is physically segregated from the Skopje agglomeration. For the period 2007-2011, daily PM_{10} concentrations were obtained from the Ministry of the Environment and daily mortality data for all causes were collected from the Institute of Public Health. Regarding Donostia-San Sebastián, we obtained the mortality data from administrative registers (Biodonostia Institute) for the period 2010-2015. Environmetal time series have been collected from the monitoring stations for air quality control located in the city area.

Methods

Let $i = 1, \dots, N$ be the indicator of the day, also referred to as the unit. Let $Z_i \in \mathcal{Z}$ be the exposure level in day i, defined as the average level of PM_{10} in the current day i and in the previous one $i - 1$ (lag 0-1 exposure), and let $Y_i \in \mathcal{Y}$ be the number of natural deaths in day i. Finally, let $X_i \in \mathcal{X}$ be a vector of K covariates for day i, which includes meteorological variables, terms, holidays, and influenza epidemics indicators.

Under the Stable Unit Treatment Value Assumption (SUTVA) [2,6], we denote by $Y_i(z)$ the potential number of deaths in day i if z were the exposure level in that day. For each day, a collection of POs is defined, one for each possible level of exposure z, but we only observe the one corresponding to the actual exposure of that day, Z_i, being $Y_i(Z_i) = Y_i$.
To extrapolate information on the unobserved POs across days with similar covariates, we first defined a parametric log-normal model for the exposure Z_i, including different regressors for each city, and we estimated the GPS, according to [4], for each day i. The GPS was evaluated at the level of the observed exposure Z_i and for various hypothetical values z that the pollutant could potentially take. In this study, we define a predetermined set of exposure levels z, which represent the median points of L quantile-based bins. Under an assumption of weak local unconfoundedness, meaning that within a range of exposure levels the assignment of the exposure value is random given the covariates, we imputed the missing POs, for each exposure level z and for each unit i', finding a matched observed unit i such that: (a) unit i had observed exposure Z_i that belonged to the bin of z; and (b) unit i was the nearest neighbor of the unit i' with respect to a two-dimensional Manhattan L1 distance on the exposure level and the estimated GPS, on a standardized scale. Then, on the entire matched set of POs, we constructed the smoothed aDRF, using a flexible regression spline. The 90% confidence intervals of the aDRF were estimated with a bootstrap method [4] repeated 100 times.

Results

In Fig. 1 we report the estimated aDRF for natural mortality, which describes how the average number of deaths in the city changes according to the PM_{10} level at lag 0-1. For all cities the causal relationship between natural mortality and air pollutant level exhibited a certain non-linearity and, although in different ranges of PM_{10}, an increasing pattern. This is in line with previous literature.

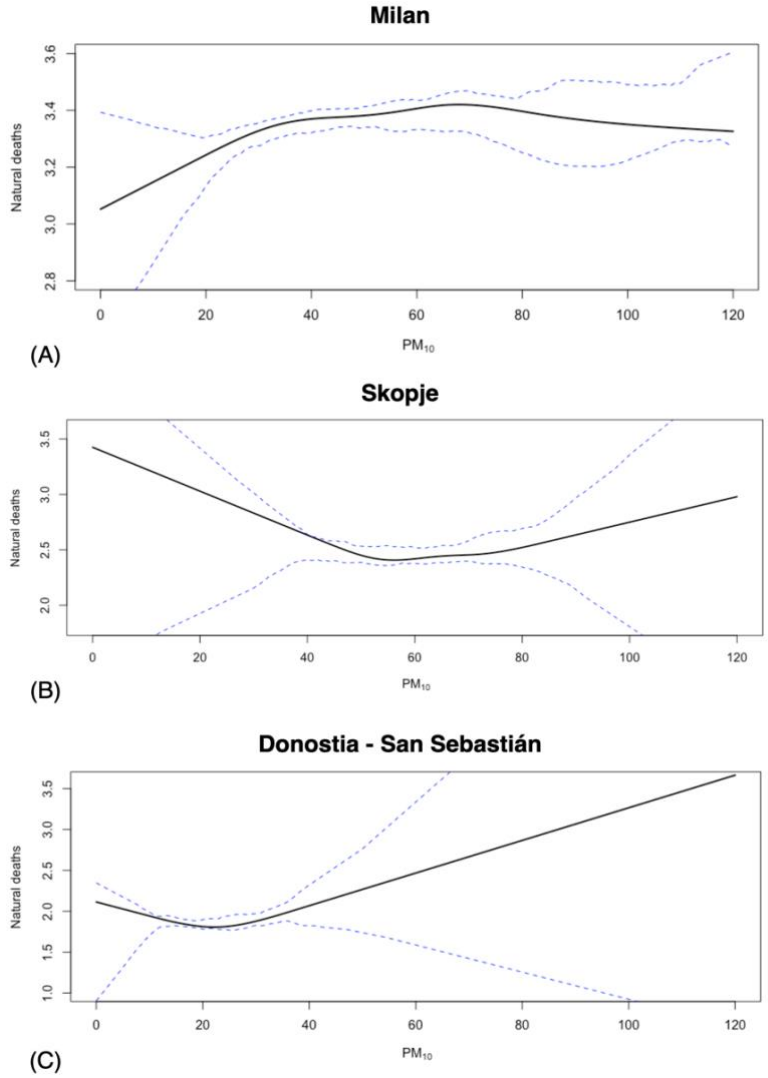

Figure 1. Average dose-response function (90% pointwise confidence band) of the causal relationship between PM_{10} exposure at lag 0-1 and average daily mortality from natural causes in Milan (A), Skopje (B), and Donostia – San Sebastián (C).

Conclusions

This multi-city analysis broadens our perspective on the impact of PM_{10} on mortality, by adopting a GPS-based matching method to obtain the average dose-response function describing the short-term effect of airborne particles on mortality. This approach can be easily extended to other environmental epidemiology contexts.

Bibliography

1. *Holland P.W. Statistics and causal inference. Journal of the American statistical Association. 1986;81(396):945-960.*

2. *Imbens G.W., Rubin D.B., Causal Inference in Statistics, Social, and Biomedical Sciences. Cambridge University Press; 2015.*

3. *Rosenbaum P.R., Rubin D.B., The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70(1):41-55.*

4. *Hirano, K. Imbens, G.W., The Propensity Score with Continuous Treatments. In Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives (eds Shewhart W.A., Wilks S.S., Gelman A. and Meng X.-L.). 2004; chapter 7:73-83.*

5. *Bia M., Flores C.A., Flores-Lagunes A., et al., A Stata package for the application of semiparametric estimators of dose–response functions. The Stata Journal. 2014;14(3):580-604.*

6. *Forastiere L., Carugno M., Baccini M., Assessing short-term impact of PM10 on mortality using a semiparametric generalized propensity score approach. Environ Health. 2020;19(1):46.*

7. *Wu X., Mealli F., Kioumourtzoglou M.A., et al., Matching on Generalized Propensity Scores with Continuous Exposures. Journal of the American Statistical Association. Published online December 12, 2022:1-29.*