Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Scienze Statistiche

Ciclo XXX

Settore Concorsuale: 13/D1
Settore Scientifico Disciplinare: SECS-S/01

CAUSAL INFERENCE METHODS IN ENVIRONMENTAL EPIDEMIOLOGY: DIFFERENT APPROACHES TO EVALUATE THE HEALTH EFFECTS OF INDUSTRIAL AIR POLLUTION.

Presentata da: Ester Rita Alessandrini

Coordinatore Dottorato

Chiar.ma Prof.ssa Alessandra Luati

Supervisore

Chiar.ma Prof.ssa Rossella Miglio

Co-Supervisore

Dr. Francesco Forastiere

Esame finale anno 2018

Acknowledgments

This thesis represents not only the final milestone of my PhD program, but also the perfect conclusion to 7 years of research in the environmental epidemiology. I was lucky during all these years to be supported and helped by several people, that I would like finally to thank.

First and foremost I wish to thank my advisors, Professor Rossella Miglio and MD Francesco Forastiere. I have been Rossella's student since the age of nineteen, when I first started studying at the University of Bologna. She taught me everything I know regarding biostatistics and epidemiology, and guided me in my personal career, always leaving an open door for me to come back. And this is what happened. I couldn't chose anyone else but her to be my PhD advisor. Rossella, I really want to thank you for always being there when I need, competent, kind and supportive.

I also really want to thank Francesco, director of the Unit of Occupational Epidemiology at the Department of Epidemiology and my advisor for this thesis. He has been my mentor, inspiratory, my example of excellence as a researcher. During the time spent working together in Rome he has always been personally touched and involved in all the studies conducted, showing a passion that I have never seen anywhere else. Francesco, I want to thank you for infusing me, I hope, with at least a part of this passion. I can consider myself an epidemiologist thanks to you and your training.

I would like to thank all the colleagues at the Department of Epidemiology in Rome, where I used to work until last year, especially Massimo, Daniela, Carla, Giulia, Chiara, Silvia, Francesca, Lisa, Martina, and many others, for their assistance, help, courage, strong emotions and support for staying sane during this doctoral adventure, and the years working together. I love you and I miss you a lot.

The PhD can be very stressing and push you beyond your limits, especially when you have decided to leave your country and start a new life and a new job. I am lucky to have found amazing people who understood that and did everything to make this challenge easier. These people are my new colleagues in Paris, who, I think, where bored to be listening to me complaining every day because I could never rest after work to finish my thesis, that I had spent another weekend depressed studying...nevertheless, they were so understanding and always cheered me up. Among them I would like to thank in particular my new managers Perrine and Patricia, who gave me their unconditioned support and who believed in me.

I would have never survived the first year of the PhD without my great colleagues of the 30° cycle in Bologna. They are all full of enthusiasm and optimism and remind me of how I used to be when I was their age. I haven't gotten a chance to get to really know them since we've been all far the ones from the others in these past two years, but they've all been so friendly and I wish these younger students the best of luck.

Thanks to my friends Valentina F, Federica, Valentina DM, Chiara, Damia, Betty, Elisa, and Aurélie. You have been patient with me when I was frustrated, you celebrated with me even when the littlest things went right, and you are there whenever I need you to just listen.

I would like to express my gratitude to Zoulikha and Alain for their unfailing emotional support and for giving me the calm and the time for studying.

I would especially like to thank my amazing family for the love, support, and constant encouragement I have gotten over the years. In particular, I would like to thank my parents, my sisters and my aunts. You are the salt of the earth, and I undoubtedly could not have done this without you.

Finally, I thank with love Pierre-Emmanuel, my boyfriend. Understanding me best as a student again himself, Pierre has been my best friend and great companion, who loved, supported, encouraged, entertained, and helped me get through this tough period in the most positive way. Pierre, I am truly thankful for having you in my life.

Index

List of Tables

List of Figures

List of abbreviations

ATE: Average Treatment Effect

ACE: Average Causal Effect

ATT: Average Treatment Effect on Treated

DID: Difference in difference

PS: Propensity score

GPS: Generalized Propensity Score

SUTVA: Stable Unit Treatment Value Assumption

ANCOVA: Analysis of Covariance

OLS: Ordinary Least Squares

IPTW: Inverse Probability of Treatment Weighting

PF: Propensity Function

DRF: Dose-response Function

HI: Hirano and Imbens

IvD: Imai and van Dyk

RISt: Robins Importance Sampling trimmed

PM_{10}: Particulate Matter of less than 10 μm in aerodynamic diameter

$PM_{2.5}$: Particulate Matter of less than 2.5 μm in aerodynamic diameter

SO_2: Sulphur dioxide

SEP: Socioeconomic position

SHR: Standardized Hospitalization Rates

COPD: Chronic Obstructive Pulmonary Disease

HR: Hazard Ratio

PH: Proportional Hazard

Abstract in English

Causal inference in air pollution epidemiology typically relies on the determination of whether estimated associations between exposure to air pollution and disease can be considered as causal. This issue is very important in environmental health sciences, where most of the research is observational in nature and the ability of the investigator to control exposure assignment is limited or non-existent. Up to now little evidence exists in literature on the use of causal inference methods in environmental epidemiology. This study aims at implementing different causal inference approaches for the first time in a longitudinal cohort analysis with a continuous exposure, to assess the causal effect of industrial air pollution on health. A first review of the literature on the addressed causal inference methods is conducted, focusing on the main assumptions and suggested applications. Then the main longitudinal study, from which the causal inference methods originate, is described. A standard time-to-event analysis is performed to assess the relationship between exposure to air pollution (PM_{10} and SO_2 from industrial origin) and mortality, as well as morbidity, in the cohort of residents around a large steel plant in the Taranto area (Apulia region, Italy). The Difference-in difference (DID) approach as well as three methods using the generalized propensity score (Propensity Function-PF of Imai and van Dyk, the Dose-response Function DRF by Hirano and Imbens, and the Robins' Importance sampling-RIS using the GPS) were implemented in a Cox Proportional Hazard model for mortality. The main study demonstrated a negative effect of exposure to industrial air pollution on mortality and morbidity, after controlling for occupation, age, time period, and socioeconomic position index. The health effects were confirmed in all the causal approaches applied to the cohort, and the concentration-response curves showed increasing risk of natural and cause-specific mortality for higher levels of PM_{10} and SO_2. We conclude that the health effects estimated are causal and that the adjustment for socioeconomic index already takes into account other, not measured, individual factors.

Abstract in Italiano

L'inferenza causale nell'epidemiologia dell'inquinamento atmosferico si basa tipicamente sul determinare se le associazioni stimate tra esposizione all'inquinamento atmosferico e la malattia possano essere considerate causali.
Questo problema è molto importante nelle scienze della salute e dell'ambiente, dove la maggior parte della ricerca è di natura osservativa e la capacità dello sperimentatore di controllare l'assegnazione dell'esposizione è limitata o inesistente. Fino ad ora esistono poche prove in letteratura sull'uso dei metodi di inferenza causale nell'epidemiologia ambientale. Questo studio mira a implementare diversi approcci di inferenza causale per la prima volta in un'analisi di coorte longitudinale con un'esposizione continua, al fine di valutare l'effetto causale dell'inquinamento atmosferico industriale sulla salute. E' stata condotta una prima revisione della letteratura sui metodi di inferenza causale affrontati, concentrandosi sulle ipotesi principali e sulle possibili applicazioni. Quindi é stato descritto lo studio longitudinale principale da cui provengono i metodi di inferenza causale. E' stata eseguita un'analisi standard time-to-event per valutare la relazione tra esposizione all'inquinamento atmosferico (PM_{10} e SO_2 di origine industriale) e mortalità, nonché morbilità, nella coorte di residenti intorno a un grande stabilimento siderurgico nell'area di Taranto (Regione Puglia, Italia).
L'approccio Difference-in-differences (DID) e tre metodi che utilizzano il punteggio di propensione generalizzata (Propensity Function-PF di Imai e van Dyk, la funzione Dose-response-DRF di Hirano e Imbens e l''Importance sampling di Robins-RIS utilizzando il GPS) sono stati implementati in un modello a rischi proporzional di Cox per la mortalità.
Lo studio di coorte principale ha mostrato un effetto negativo dell'esposizione ad inquinamento atmosferico industriale sulla mortalità e morbilità, dopo aver controllato per occupazione, età, periodo di calendario e indice di posizione socioeconomica. Gli effetti sulla salute sono stati confermati in tutti gli approcci causali applicati alla coorte e le curve concentrazione-risposta hanno mostrato tassi di mortalità naturali e causa-specifici crescenti per livelli più alti di PM_{10} e SO_2. Possiamo concludere che gli effetti stimati sono causali e che l'aggiustamento per l'indice socioeconomico tiene già conto di altri fattori individuali non misurati.

CHAPTER 1. Causal Inference in Observational Studies

1.1. Introduction

A central issue in epidemiology is the evaluation of the causal nature of reported associations between exposure to defined risk factors or treatments and the occurrence of disease. This issue is even more important in environmental health sciences, where most of the research is observational in nature and the ability of the investigator to control exposure assignment is limited or non-existent. Nonetheless, besides contributing to the understanding of disease causation, etiologic studies are commonly regarded as providing the scientific basis for the adoption of preventive or regulatory actions. Therefore, it becomes necessary to clearly define what is meant by "causal relationship", how to properly design an epidemiological study to detect causal effects, and under which conditions and assumptions such an approach would be feasible.

An internal and external valid statistical association is not sufficient to determine causal association. After assessing the study's validity, the extent of the evidence being supportive of causality should be considered. Hill (1965) has provided his famous 9 viewpoints to determine causality: strength of the association, consistency, specificity, temporality, biological gradient, plausibility, coherence, experiment, and analogy. However, Hill himself mentioned that none of them are essential to infer causality and Bae at al. (2017) pointed out their limits of application in environmental epidemiology studies, where an exposure can be associated with multiple diseases (hence specificity doesn't hold) or can have a non-linear association with the disease (linear biological gradient needs to be reconsidered).

A large body of evidence on the health effects of air pollution exists, indicating that chronic and acute exposure to particulate matter is associated with cardiorespiratory mortality and morbidity (Alessandrini et al., 2013; Analitis et al., 2006; Beelen et al., 2008; Bentayeb et al., 2015; Giulia Cesaroni et al., 2013; Colais et al., 2009; Jerrett et al., 2009; Kloog, Ridgway, Koutrakis, Coull, & Schwartz, 2013; Krall, Anderson, Dominici, Bell, & Peng, 2013; Mataloni et al., 2012; Peng et al., 2013; E Samoli et al., 2006; Evangelia Samoli et al., 2013; Scarinzi et al., 2013; Schwartz, 2004; Stafoggia et al., 2013; Stölzel et al., 2007; Zanobetti, Dominici, Wang, & Schwartz, 2014; Zanobetti & Schwartz, 2009). These were all associational studies, where causality was supported by the consistency of results and by Hill's criteria.
Causal inference in air pollution epidemiology typically relies on the determination of whether estimated associations between exposure to air pollution and disease can be considered as causal (Correia et al., 2013; Dockery et al., 1993; Laden, Schwartz, Speizer, & Dockery, 2006; Pope III, Ezzati, & Dockery, 2009; Zeger, Dominici, McDermott, & Samet, 2008).The potential outcome approach in causal inference is not new in biomedicine (Afendulis, He, Zaslavsky, &

Chernew, 2011; Gillespie et al., 2015; Hernán et al., 2008; Jacob, Sutradhar, Moineddin, Baxter, & Urbach, 2013; B. Lu & Marcus, 2012; M. Lu, 1999; Shimizu et al., 2016), but it's not consolidated in studies of air pollution effects on health.
In support of the use of causal inference approaches in this field, Zigler and colleagues (2016) have provided a textbook on how more systematic approaches to testing of causality (i.e., through use of causal inference frameworks and methods) could be adapted to the assessment of the effects of air pollution interventions on air quality and health.

Up to now applications of causal inference methods in environmental epidemiology for air pollution are sparse (Baccini, Mattei, Mealli, Bertazzi, & Carugno, 2017; Capuno, Tan, & Javier, 2018; Kirby, Nagel, Rosa, Thomas, & Clasen, 2016; Mueller, Pfaff, Peabody, Liu, & Smith, 2011; Neupane et al., 2015; Rosa et al., 2017; Sætterstrom et al., 2012; Schwartz, Bind, & Koutrakis, 2017; Wang et al., 2016; Wylie et al., 2014; Wylie, Coull, et al., 2015; Wylie, Singh, et al., 2015; C. Zigler et al., 2016).

The purpose of this study is to show new applications of different causal inference methods to a longitudinal cohort, aimed at assessing the effects of air pollution on health in the industrial area of Taranto, Italy.

This study apply two methods dealing with non-experimental data, characterizing the identification assumptions necessary to justify their application. One of these approaches for longitudinal data is the difference-in-differences approach (DID), a quasi-experimental design to obtain a more robust estimate of the treatment effect, and the other is the use of the propensity score modelling in the continuous setting. After a first overview of the potential outcome framework in Chapter 1, the wide array of existing literature about the PS methodology will be presented in Chapter 2 and Chapter 3 for extension to the continuous treatment setting, while the general DID approach will be treated in this chapter.

Chapter 4 will introduce the cohort study on the residents living around a large steel plant in the south of Italy, from which the two applications of causal inference methods originated. The cohort study was characterized by an extensive work of exposure assessment and cohort building. The first year of the PhD was dedicated to this aim and to the estimate of the association between exposure measures and health events, like mortality, hospitalizations and cancer incidence. The whole cohort study itself would require a monography to describe it entirely. Here it was presented in one chapter, focusing on the main statistical methods and results, from which the work on the DID and PS approach, presented in their applications in Chapter 5 and 6, have started.

Little evidence exists in this field referring to the DID approach (Wang et al., 2016) and the propensity score for continuous exposure variable (Schwartz, Austin, Bind, Zanobetti, & Koutrakis, 2015).
In Italy in the same field no study until now applied DID approach and only one study by Baccini et al. (2017) implemented the propensity score matching

approach to obtain the causal impact of high (dichotomized over the threshold) daily levels of PM_{10} on mortality in Milan.
Other than Schwartz et al.'s work in a time series setting (Schwartz et al., 2015), no study worldwide has been done on propensity score methods in time-to-event analysis with continuous exposures measurements in air pollution studies.

Further studies are needed on the correct use of causal inference methodologies for air pollution longitudinal studies.
The purpose of the current study is to provide new evidences on the use of the DID approach, with an extension of the existing method, and the generalized propensity score methods in a longitudinal study on the effects of air pollution on health in the Italian context.

There is no intention here to prove which one of the methods implemented is the best one or performed better, since no assessment of bias is possible, unless simulations are performed.

The data and results of the research for this thesis have been developed during the working activity at the Department of Epidemiology of the Lazio Regional health Service in Rome, Italy, as part of the PhD program. The study was conducted in collaboration with the Regional Agency for the Environmental Protection of the Apulia Region (ARPA), in the South of Italy, the Local Health Service of Taranto (ASL), Italy, and the AReS Puglia, Italy.

All the materials and results developed during the main cohort study described in Chapter 4 and Chapter 5 were collected in a final exhaustive report published in August 2016 on the Puglia Salute website of the Apulia Region (https://www.sanita.puglia.it/documents/890301/896208/Relazione+Finale+Studio+di+Coorte+-+2016/ea231c81-e196-4b43-99a4-0882bd60b83b).

Two papers for publication on scientific journals have been prepared respectively for the main cohort study results and for the DID approach, attached in Annex A and Annex B. The research on GPS at the time of the end of the program was not ready for publication.

1.2. Overview

Causal inference aims at quantifying the causal relationship between a specified treatment and a subsequent outcome. Among all type of study designs, randomized controlled trials are considered as the gold standard for studying causal inferences (W.R. Shadish, Cook, & Campbell, 2002)

In experimental studies, the treatment is randomly assigned to the subjects with a known probability. Randomization generally provides an equal chance of being assigned to treated and non-treated groups, and tends to assure that the two groups are homogenous concerning their pre-treatment characteristics. Therefore the

direct comparison between the treatment and control group should lead to a valid estimate of causal effect, since the bias due to both observed and unobserved covariates is generally eliminated by the randomization of treatment assignment.

In many instances, however, conducting a randomized experiment is not feasible or ethical.

Observational studies are proposed and used as alternatives to randomized designs. In these studies the mechanism of treatment assignment is not under the control of the investigator and can be dictated either by the choice of the individual or other factors. Consequently, the treated and control groups may have differed prior to receiving the treatment in ways that are relevant to the outcome of interest. Thus, observed differences post treatment may be attributable either to the treatment effect, these pre-treatment differences or both. Controlling observed biases and addressing hidden biases are the two major tasks of making valid causal inferences in an observational study. (Paul R. Rosenbaum, 1996)

There are different ways to formulate and address the problem of causal inference with observational data. One of them is the framework of potential outcomes (Rubin, 1978; Splawa-Neyman, Dabrowska, & Speed, 1990).

1.3. Potential Outcomes Framework

Potential outcomes are the responses that would be realized if different treatments were given to a unit. Some authors have called them counterfactuals (Greenland, Pearl, & Robins, 1999).

Notation for the potential outcomes was first introduced by Neyman et al. (1990) in the context of randomized experiments. Rubin (1974) extended Neyman's work and formalized the potential outcomes framework for the observational studies and formulated causal inference as a problem of missing data (Little & Rubin, 2000; Rubin, 2005).

I will first start discussing the problem of causal inference when the treatment is binary. Let T_i represent the value of the treatment applied to unit i, having, for example, $T_i=1$ for the treatment group and $T_i=0$ for the control group. Let $Y_i(0)$ and $Y_i(1)$ denote the potential outcomes for unit i under treatments 0 and 1, respectively. Rubin (1978) defined the causal effect of a treatment on unit i as $Y_i(1) - Y_i(0)$.
However, as pointed out by Holland (1986), the "fundamental problem of causal inference" is that for each subject we can observe only one of these potential outcomes, because, at a particular point in time, each subject would only receive either the treatment or the control, not both. For this reason, estimating causal effects can be regarded as a missing data problem, where the unobserved potential outcomes are missing. The observed outcome for unit i can be denoted by $Y_i = Y_i(T_i)$, which can also be written as $T_iY_i(1) + (1-T_i)\ Y_i(0)$.
The variable T_i can be thought of as a missing data indicator that tells us which potential outcome is seen and which one is hidden for the given unit.

The final quantity of interest will be the average of the causal effect (ACE) or average treatment effect (ATE) for a given population,

$$ATE = \mu(1) - \mu(0) = E[Y(1)] - E[Y(0)] \quad (1.1)$$

The notation developed so far can be easily extended to handle more than two treatments. In the case of three treatment levels there would be three potential outcomes for each unit, $Y_i(0)$, $Y_i(1)$, and $Y_i(2)$, and three causal comparisons: $Y_i(1) - Y_i(0)$, $Y_i(2) - Y_i(0)$, $Y_i(2) - Y_i(1)$. The observed outcome can be written as $Y_i = Y_i(T_i) = \sum_{t=0}^{2} \mathbb{1}(T_i = t) \cdot Y_i(t)$.

The framework of potential outcomes can then be easily extended to the situations in which the treatment variable is continuously distributed. Suppose that T_i takes values within a real interval $\mathrm{T} = [t_{\min,}, t_{max}]$. The potential outcomes are now an uncountable infinite set $Y_i = \{Y_i(t): t \in \mathrm{T}\}$. The observed outcome can still be written as $Y_i = Y_i(t)$, and the causal effect for unit i of moving from treatment dose t to t^* is $Y_i(t^*) - Y_i(t)$.
This unit-level effect is unobservable, but under certain conditions, we may be able to reconstruct a reasonable estimate of the population average effect of moving from t to t^*, $E\big(Y_i(t^*)\big) - E\big(Y_i(t)\big)$.

In some situations, it is not the ATE that is of interest, but rather the average treatment effect on treated (ATT). This implies considering the differences in outcomes with and without treatment only for individuals who are treated (Caliendo & Kopeinig, 2008; Holland, 1986; Rubin, 1973a, 1973c).

1.4.Rubin's causal model

Since outcomes for all treatment conditions cannot be observed for all units, Rubin's causal model operates under several key assumptions discussed below (Paul R Rosenbaum & Rubin, 1983, 1984; Rubin, 2001).

Stable unit treatment value assumption (SUTVA) (Cox, 1958; Rubin, 1978)

The outcome of subject *i* to the treatment is independent of the treatment status of subject *j*.
More formally, SUTVA is defined as an "a priori assumption that the value of Y for unit i when exposed to treatment t will be the same, no matter what mechanism is used to assign treatment t to unit i and no matter what treatments the other units receive" (Rubin, 1986, p.961).

Strongly ignorable treatment assignment (Paul R Rosenbaum & Rubin, 1983)

The treatment assignment T and the potential outcomes $\big(Y(1), Y(0)\big)$ are conditionally independent given the observed covariates X, and each subject has a positive probability of receiving the treatment:

$$\{Y(1), Y(0)\} \perp T | X \quad (1.2)$$
$$0 < p\,(T = 1|\mathrm{X}) < 1 \quad (1.3)$$

This implies that conditional on the observed covariates *X*, the treatment is randomly assigned as if in randomized experiments, and there is no systematic pre-treatment difference between the treatment and the control group. This assumption is known as *no unmeasured confounders*. In addition, each subject has a nonzero probability of receiving either treatment. This second condition is also an *overlap* assumption, meaning that given covariates X, the person with the same X values has positive and equal opportunity of being assigned to the treated group or the control group.
If ignorability holds, one can obtain unbiased treatment effect estimates.

This condition is considered to be satisfied in randomized studies where units are assigned randomly to conditions and thus the cause of assignment mechanism T, is statistically independent from the outcomes. However, when randomization is not employed, satisfying this assumption becomes more difficult, as the differences in the observed outcomes may be attributed to alternative or unobserved variables related to the assignment mechanism.

1.5.Causal Inference in Non-experimental Data

In contrast to the experimental analysis, data in non-experimental or observational studies are not derived in a process that is completely under the researcher's control. For example, a government authority might offer a program to a particular area or specific individuals in order to improve their condition or because it is believed that they held favourable expectations regarding the program's impact.
The main objective of any observational study is to use the observable information in an appropriate way to replace the comparability of treatment and control group by an appropriate alternative identification condition. The objective is to use the available information such that in the sub-population defined by these observables, any remaining differences between treated and non-treated might be attributed to the program.
The process of finding information is really about finding counterfactuals for observations in the treatment group. In other words, researchers need to make adjustment and use statistical approaches to obtain groups that are homogeneous in terms of the distribution of covariates.

The natural experiment approach attempts to find a naturally occurring comparison group that can mimic the properties of the control group in the properly designed experiment. The major limitation in inferring causation from natural experiments is the presence of unmeasured confounding (Craig, Katikireddi, Leyland, & Popham, 2017).

The different methodologies for solving the evaluation problem with non-experimental data mainly depend on four factors: the type of information available to the researchers, the parameters of interest, the underlying model and how the counterfactual is structured.

Statistical approaches such as ordinary regression, covariate adjustment analysis, structural equation modelling, selection models, and matching methods can be applied to adjust for differences between the groups in non-randomized designs (W.R. Shadish et al., 2002; Stuart, 2010) and rely on the *untestable assumption of no unmeasured confounding* (Cole & Hernán, 2008; Stampf, Graf, Schmoor, & Schumacher, 2010).

Traditional matching methods, in particular, such as simple mean matching, pair matching (Rubin, 1973a) and multivariate matching (Rubin, 1976, 1979) are statistical methods that aim to equate groups for causal and non-causal effect estimation (Stuart, 2010). As the number of variables used to match increases, the number of combinations for individual matches between groups also increases exponentially (W. G. Cochran & Chambers, 1965). In addition, with traditional matching methods there is a potential to lose a lot of data on the final matched individuals.

In 1983, Rosenbaum and Rubin introduced the propensity score (PS), a major advancement in causal analysis, particularly due to its ability to balance groups using a set of covariates reduced to a single score, thus eliminating the challenges with traditional matching. The ability to include many variables increases the likelihood of satisfying the strong ignorability assumption and yielding then to unbiased treatment effects (William R Shadish & Steiner, 2010; Stuart, 2010).

Given the difficulty of measuring accurately all of the characteristics associated with exposure to an intervention, methods such as difference in differences, instrumental variables and regression discontinuity designs that deal with unobserved factors are a potentially valuable advance on those that only deal with observed factors.

The *Instrumental variable* method is based on the existence of an external factor strongly influencing the selection process, conditioning the individual behaviours, and that has no effect on the outcome. Intuitively the strategy consists in disentangling the treatment variable in two parts: one determined by the external factor out of the individual control, and another one determined by the individual behaviours linked to unobservable factors. So only the part of the treatment not depending on the unobserved factors will be used: this takes the name of "instrumental variable". While useful in a variety of applications, the validity and interpretation of IV estimates depend on strong assumptions, the plausibility of which must be considered with regard to the causal relation in question (Blundell & Dias, 2009; Craig et al., 2012; Craig et al., 2017).

The *discontinuity design* method exploits situations where the probability of enrollment into treatment changes discontinuously with some continuous variable

according to a step change or 'cutoff'. The assumption is that units (individuals, areas, etc.) just below and just above this threshold will otherwise be similar in terms of characteristics that may influence outcomes, so that an estimate of treatment effect can be obtained by comparing regression slopes on either side of the cutoff (Blundell & Dias, 2009; Craig et al., 2012).

Difference in differences (DID) method is usually implemented by comparing the difference in average behaviour before and after the treatment/policy for the eligible group with the before and after contrast for a comparison group. This approach can be a powerful tool in measuring the average effect of the treatment on the treated. It does this by removing *unobservable* individual effects and common macro effects by relying on two critically important identifying assumptions of (i) common time effects across groups, and (ii) no systematic composition changes within each group.
More details of the DID method and one application to the Taranto cohort study will be presented in this thesis. (Blundell & Dias, 2009; Craig et al., 2012; Craig et al., 2017)

1.6. Difference in Differences (DID) Estimator

This method, popular in economics research, evaluates the effect of policy change and public intervention programs. Much literature on DID has arisen since Ashenfelter and Card (1985).

One of the first applications of the DID design was the study by Card and Krueger (1994). The authors used the DID design to assess the effect of an increased minimum wage on fast food employment rates in New Jersey 5 months before and 5 months after increasing the minimum wage. Pennsylvania, where the minimum wage did not change, was used to identify the variation in employment expected in the absence of the raise. The increase in the minimum wage was the random shock.

In the typical set up of using DID, the outcomes are observed for two groups at two-time periods. During the first period, two groups are not exposed to the treatment, but during the second period, one of the group is exposed to the treatment (referred to as treatment group) while the other is not (referred to as control group). The underlying key assumption of this approach is that in absence of the treatment, the average outcomes for both treated and control groups would have changed in parallel paths over time, and all the other factors, other than the treatment, would affect both groups in the same way.

The basic DID framework can be described as follows (Abadie, 2005; Ashenfelter & Card, 1985). Let Y be an outcome over a population of individuals for which

we wish to estimate the impact of a treatment T. Suppose that there exist two groups with treatment status T=0,1, indicated the control and treated group, respectively.

Assume then that we observe individuals in two moments in time, *t=0,1,* indicating the pre-treatment, i.e. the time period before the treatment group receives the treatment, and the post-treatment, i.e. the time period after the treatment group receives the treatment, respectively. For each individual the outcome Y_i is modelled by the following equation:

$$Y_i = \alpha + \beta T_i + \gamma t_i + \delta(T_i \cdot t_i) + \varepsilon_i \qquad (1.4)$$

where the coefficients α, β , γ , δ are all unknown parameters and ε_i is the random error. Considering the constant term α, β can be interpreted as the treatment group specific effect, while γ as the common time trend to treated and not-treated, and δ as the **true effect of treatment** and the quantity of interest.

The DID estimator of δ, $\hat{\delta}$, requires the following assumptions to be true:

1. The model in (1.4) to be correctly specified
2. The error term to have zero mean, i.e. $E[\varepsilon_i] = 0$
3. The error term to be uncorrelated with the other terms in (1.4).

Under these assumptions we can estimate the expected values of the average outcomes using arguments in equation (1.4) as:

$$\begin{aligned} E[Y_0^T] &= \alpha + \beta \\ E[Y_1^T] &= \alpha + \beta + \gamma + \delta \\ E[Y_0^C] &= \alpha \\ E[Y_1^C] &= \alpha + \gamma \end{aligned}$$

with subscripts identifying the time period and the superscript the treatment status. We consider $\bar{Y}_0^T$ and $\bar{Y}_1^T$ as the sample averages of the outcome for the treatment group before and after treatment, respectively, and $\bar{Y}_0^C$ and $\bar{Y}_1^C$ the correspondent in the control group.

First, we build an estimator based on comparing the average difference in outcome before and after treatment in the treatment group alone:

$$\hat{\delta}_1 = \bar{Y}_1^T - \bar{Y}_0^T,$$

For which the expectation is

$$E\left[\hat{\delta}_1\right] = E[\bar{Y}_1^T] - E[\bar{Y}_0^T] = [\alpha + \beta + \gamma + \delta] - [\alpha + \beta] = \gamma + \delta$$

Meaning that the estimator is correct if $\gamma = 0$, i.e. if there is no trend in the outcome Y_i.

If in the same way we consider an estimator based on comparing the average difference in outcome Y_i in the post-treatment period only, between treated and control groups we have:

$$\hat{\delta}_2 = \bar{Y}_1^T - \bar{Y}_1^C.$$

The expectation is

$$E\left[\hat{\delta}_2\right] = E[\bar{Y}_1^T] - E[\bar{Y}_1^C] = [\alpha + \beta + \gamma + \delta] - [\alpha + \gamma] = \beta + \gamma,$$

meaning that the estimator will be biased as long as $\beta \neq 0$; that is,when significant differences in the outcome exist between treated and control groups, due to differences existing between the two groups prior to any treatment.

The DID estimator is a combination of the two estimators and defined as the difference in average outcome in the treatment group before and after treatment minus the difference in average outcome in the control group before and after treatment. It is in this sense a "difference of differences"

$$\hat{\delta}_{DD} = (\bar{Y}_1^T - \bar{Y}_0^T) - (\bar{Y}_1^C - \bar{Y}_0^C) \qquad (1.5)$$

for which the expectation is unbiased:

$E\left[\hat{\delta}_{DD}\right] = E[\bar{Y}_1^T] - E[\bar{Y}_0^T] - (E[\bar{Y}_1^C] - E[\bar{Y}_0^C])$
$= [\alpha + \beta + \gamma + \delta] - [\alpha + \beta] - [\alpha + \gamma - \gamma]$
$= [\gamma + \delta] - \gamma$
$= \delta.$

The DID estimator relies on the three assumptions at p.19, that means that if any of them do not hold the estimator, $\hat{\delta}_{DD}$ will be biased. There are major difficulties in verifying them, since they are made about unobservable quantities.

The third one, implying no correlation of the random errors with the terms of model (1.4), actually imposes one strong condition called the "parallel-trend assumption", that states that $cov(\varepsilon_i, T_i \cdot t_i) = 0$, meaning that Y follows a different trend for the treatment and control groups. The failure of these conditions may be a common problem in many studies and can cause the DID estimators to be biased (Meyer, 1995).

CHAPTER 2. Propensity Score

2.1.Definition

The use of matching methods is seriously limited by the high dimensionality of the vector *X,* because in practice it would be difficult to find control units with similar values of *X* if it is composed by a large vector of variables. A solution can be to match on a function of *X*. The most common and useful is the propensity score, $\pi(x)$, introduced by Rosembaum and Rubin (1983) as a balancing score to group treatment and control units in non-randomized studies. A balancing score for its definition is a "function of the observed covariates, such that the conditional distribution of these covariates is the same for treated and control units" (Paul R Rosenbaum & Rubin, 1983, p. 42).

In the binary treatment setting, the propensity score is the conditional probability $\pi(x)$ of being assigned to a treatment given a set of covariates:

$$\pi(x) = P(T_i = 1|X_i = x). \quad (2.1)$$

The propensity score for sample unit i will be written as $\pi_i = \pi_i(X_i)$. It summarizes in a single value the impact of all the observable pre-interventions units' characteristics that differentiate the treated units from the excluded ones.
It is possible to demonstrate that the PS is the coarsest balancing score. According to this property, the PS balances covariates across the treatment groups, which is: $P\ (X_i = x_i|\pi_i = c, T_i = 1) = P\ (X_i = x_i|\pi_i = c, T_i = 0)$. Units with the same propensities have the same probability of treatment. This means that units randomly receive treatment or control and have the same distribution of covariates *X* on average (Paul R Rosenbaum & Rubin, 1983, 1984).

Rosembaum and Rubin (1983) demonstrated that the conditional independence assumption remains valid conditioning for $\pi(x)$ instead of *X*:

$$\{Y(1), Y(0)\} \perp T|\ \pi(x) \quad (2.2)$$

Conditioning on $\pi(x)$ means that units with equal propensities appear to be allocated to a treatment randomly, then estimating causal effects becomes straightforward.

2.2.Estimation

One of the most important considerations to be made on propensity score is regarding its estimation. In a nonrandomized study, the PS is unknown and must be estimated. The most common way to model the PS in a binary treatment setting

is by logistic regression, although other methods such as probit regression, boosted regression (McCaffrey, Ridgeway, & Morral, 2004; Zhu, Coffman, & Ghosh, 2015), CART, and random forest have been used (Lee, Lessler, & Stuart, 2010).

The selection and inclusion of appropriate covariates is an integral component in PS. The quality of the estimated PS depends upon its ability to remove hidden bias, which is a function of the covariates included in the model. PS relies heavily upon the assumption of ignorable treatment assignment; therefore, in order to satisfy the ignorable assumption all variables related to treatment and outcome need to be included (Rubin & Thomas, 1996; Stuart, 2010). There is no statistical test for this assumption; it is only based on the substantive knowledge of the possible confounding variables in the applied context. Therefore, even if a large number of covariates are included, there is no way to tell whether a confounding variable has been excluded or overlooked. If important covariates related to treatment assignment are omitted or excluded, then there is potential for the PS to be biased, as will be the case for the resulting treatment effect estimates (Paul R Rosenbaum & Rubin, 1983; Rubin, 1997; William R Shadish & Steiner, 2010).

Once the PS has been estimated it is important to check its fit, traditionally through covariate balance. Ideally, the PS balances the distribution of covariates across all treatment levels. In the binary treatment setting, in the subgroup of units with the same values of the propensity score, we will find no major differences between $T_i = 1$ and $T_i = 0$ on any component of X_i or function of X_i. One possible check would divide the sample into subgroups defined for example by the quantiles of the estimated PS, and will calculate standardized mean differences and relative t-tests for each covariate entering the prediction model for the PS. Graphically, this could be checked through histograms or boxplots of the covariate distributions at different treatment levels.

Another condition to be checked for the PS is the overlap to determine if estimating causal effects is even feasible. If there is an insufficient number of graphically determined units with either $T_i = 1$ in regions with low propensity scores or an insufficient number of units with $T_i = 0$ in high-propensity regions, then it isn't advisable to have causal inference in those regions because it may require excessive extrapolation.

A broad region of common support allows for causal effect estimates to be based on the full range of PS in the sample, whereas a small common support region restricts the effect estimates. Often individuals who have estimated PS that fall outside the common support region are dropped from the analysis. These individuals are usually located in the tails of the distribution, with estimated PS close to 1 or 0. If PS yields insufficient overlap then it may be that the model to estimate the PS needs to be adjusted.

Overfitting is not a serious problem when estimating PS, because the main goal is to predict the treatment probabilities rather than to estimate the best model.

In fact, overfitting the PS is not disadvantageous for the estimation of ATE, since prediction and covariate balance are the main criteria for estimating the PS (Brookhart et al., 2006; Rubin, 2004).

Once PS is estimated, the next step is to utilize, or condition on it to balance the treatment and control groups. There are four general conditioning techniques: matching, stratification, covariance adjustment, and weighting. The first three techniques were introduced with the PS in 1983 by Rosenbaum and Rubin, while the last technique was introduced a few years later (Paul R Rosenbaum, 1987).

2.3. Use of Propensity Score

2.3.1. Matching

Propensity score matching involves the formation of new samples of data that contain only those individuals matched. Individuals from treatment and control groups who share similar covariate distributions are matched to balance the groups (Paul R Rosenbaum & Rubin, 1983; Rubin & Thomas, 1996, 2000; Stuart, 2010).
With a binary treatment, once the PS is estimated, for each unit in the smaller group ($T_i = 0$ or $T_i = 1$), an algorithm is used to select a unit in the larger group with a similar propensity score, and perhaps similar values of X_i as well, if the PS is correctly specified.

The similarity of the PS between treated and controls rely on the distance/difference between the values for the selected observations. Each individual is randomly matched to an individual with the closest propensity score. After matches have been found for all units in the smaller group, the excess units in the larger group are discarded. This method is called 1-to-1 matching. The method can provide a causal inference on the effect of a treatment on the outcome while removing biases due to confounding attributed to covariates, since the pairs have similar covariates but different treatments. The ATE is calculated by comparing the mean outcomes in each group. It is also possible to combine model-based regression adjustment with matching (Rubin & Thomas, 2000).

Because of the common support assumption, the ATE may be estimated in a matched sample where the covariate distribution is different from the overall population, thus it may be not generalizable to the ATE for the whole population. However, it may provide a realistic and more accurate estimate for a smaller population.

The matching techniques can be categorized into parametric and non-parametric matching. The near-neighbour, the Mahalanobis distance, the caliper matching and the near-neighbour with caliper are parametric types of greedy matching. Greedy matching considers the individual scores to match on, based on the best control unit matched to each individual treatment unit (Gu & Rosenbaum, 1993). Full and optimal matching belongs to the optimal parametric approach. Optimal

matching has an advantage on the greedy matching, as it does not consider the best matches for individual treatment units; rather, matches are made to minimize the overall distance for all units in the matched sample (Guo & Fraser, 2014; Thoemmes & Kim, 2011). Full matching is a type of optimized matching analysis that allows an observation/individual to be matched to many cases (1 to many). Full matching has the advantage of retaining a sufficient number of observations for the outcome analysis.

Non-parametric matching includes kernel matching that uses weighted averages of all observations in the control group to estimate the counterfactual outcomes. The weight is calculated by the PS distance between a treatment observation and all control observations. The closest control cases are given the greatest weight (Smith & Todd, 2005).

Matching with replacement allows matched controls to be placed back into the control group to be matched to other treatment units. Therefore, control units may be matched with multiple treatment units. Matching without replacement does not allow a control unit previously matched to a treatment unit to be considered for matching again (Austin, 2011).

The extension of matching methods to the continuous setting is not easy, because of the presence of few cases for each value of $t \in T$, and the selection rule, that cannot distinguish between treated and non-treated units. Lu et al. (2001) applied matching on doses, but they dichotomized the dose, reconducting to the binary setting and estimating the effect of high dose vs. low dose, rather than estimating all the dose-response function.

Matching in the continuous setting has not been treated in this monography, as our current focus is on the main generalizations of propensity score methods.

2.3.2. Stratification

Propensity score stratification or subclassification was originally proposed by Rosenbaum and Rubin (1983) to adjust for selection bias. The idea is that subjects classified in the same stratum would have similar propensity scores. Within the subclass, the treatment status is essentially random.

In this method, the PS is first estimated, then the units are divided into groups with similar propensities, usually based on quantiles of the estimated PS values. The number of classes may depend on the size of the data set, but Rosenbaum and Rubin (1984), extending Cochran's work (1968) of stratification on single covariate, found that stratifying the data into quintiles (K=5) of the estimated PS eliminates up to 90% of the bias in each coordinate of variables. Units within the same PS subclasses should be similar in their covariate distributions.

The effects of each subclass are pooled across the strata to estimate the ATE (Paul R Rosenbaum & Rubin, 1984).

$$\widehat{ATE} = \sum_s \frac{N_s}{N} \widehat{\theta}_s, \quad (2.3)$$

where s is the subclass index, N_s is the number of units in subclass s, N is the total sample size, and $\widehat{\theta}_s$ is the ATE estimate within subclass s.

Austin (2011, p. 408) compared this procedure to that of a "meta-analysis of a set of quasi-RCTs," because within each stratum the effect of treatment on outcomes is estimated by comparing the outcomes of the treated and untreated subjects.

2.3.3. Covariance adjustment

Propensity score can also be used as a regressor in regression models for covariates.
This is similar to an Analysis of Covariance (ANCOVA) model; that is, a class of models that allow for both continuous and categorical predictors to be modelled simultaneously and to test whether certain factors have an effect on the outcome of interest after removing the variance for which covariates account.

As shown in Rosenbaum and Rubin (1983), using propensity score in the linear models requires a stronger assumption that the conditional expectation of the outcome is linear in the propensity score. Compared with matching and stratification, covariance adjustment is more sensitive to the nonlinearities in the response model. However, covariance adjustment for propensity scores can be used as a supplement to matching or stratification, which will help to remove the residual bias (Rubin & Thomas, 2000).

The estimation of treatment effect is generally not affected dramatically, but using regression adjustment can decrease the standard errors of the estimated effects.

2.3.4. Weighting

First introduced by Rosenbaum (1987), weighting on the PS is conducted to ensure samples are representative of the population of interest. Weighting is often used in a survey research, test equating, and norming to draw inferences from non-representative samples to a population. In PS analysis, individual units are weighted based on their estimated PS (Robins & Finkelstein, 2000).

In the binary framework the weighting procedure starts by estimating the PS. Each unit in the $T_i = 1$ group is assigned a weight proportional to $1/\pi_i$, and each unit in the $T_i = 0$ group is assigned a weight proportional to $1/(1 - \pi_i)$. This technique is very similar to the Horvitz and Thompson (1952) weighting method used in sample surveys. The weight of a sample unit is generally the inverse of the subject's probability of being included in the sample. Units with lower probabilities of getting selected into their respective treatment group will have more weight allocated to them, whereas units with higher probabilities of selection will be assigned less weight.

Considering $\omega_i(1) = \frac{1}{\pi_i} = 1/P(T_i = 1|\boldsymbol{X_i})$ and $\omega_i(0) = \frac{1}{(1-\pi_i)} = 1/P(T_i = 0|\boldsymbol{X_i})$, where $\omega_i(t)$ represents the weight of unit i at treatment level t, the inverse probability weighting estimator for the ATE is

$$\widehat{\text{ATE}} = \frac{\sum 1\,(T_i=1) Y_i \omega_i(1)}{\sum 1\,(T_i=1)\omega_i(1)} - \frac{\sum 1\,(T_i=0) Y_i \omega_i(0)}{\sum 1\,(T_i=0)\omega_i(0)}, \qquad (2.4)$$

where summations are taken over the entire sample.

CHAPTER 3. Generalizing the Propensity Score to the Continuous Treatment Setting

3.1. Introduction

It is not rare to find practice situations in which treatment regimens are not binary, and units might be exposed to different levels or doses of treatment. This is true, both in economic and in medical applications. In these situations, studying the impact of such treatment as though it were binary can mask some important features.
From a notational point of view, the framework of potential outcomes is easily adapted to situations where the treatment variable is continuous.

Suppose that T_i takes values within a real interval $\mathrm{T} = [t_{min}, t_{max}]$. The potential outcomes are now an uncountably infinite set $Y_i = \{Y_i(t): t \in \mathrm{T}\}$. The observed outcome can still be written as $Y_{i=}Y_i(t)$, and the causal effect for unit i of moving from treatment dose t to t^* is $Y_i(t^*) - Y_i(t)$.

The classic causal inference methods dealing with binary treatment, stratification, matching, weighting and covariance adjustment, do not have obvious extensions to continuous treatment. The multiple treatment problem has not received much attention, and there is considerably less evidence compared to the binary setting.

More than 10 years ago, several researchers proposed generalization of the propensity score methodology for the non-binary treatment regimens (Flores, Flores-Lagunes, Gonzalez, & Neumann, 2012; Hirano & Imbens, 2004; Imai & Van Dyk, 2004; Guido W Imbens, 2000). Novel applications appear also in Ertefaie and Stephens (2010) and Moodie and Stephens (2012).
Flores-Lagunes, Gonzales, and Neumann (2012), proposed two extensions to the method of Hirano and Imbens (2004) that aim to provide more a robust estimation through a more flexible response model.

In this work, the two main generalizations of the propensity score methods, the propensity function (PF) of Imai and van Dyk (2004), and the generalized propensity score (GPS) of Hirano and Imbens (2004), along with the Flores et al. (2012) extension, will be examined more closely and applied to a cohort study.

3.2. Generalized Propensity Score

One of the first studies dealing with continuous treatment is the work of Imbens (1999). He proposed an extension of the propensity score methodology allowing

for an estimation of average causal effects with multi-valued treatments. This works represents the starting point for the next analysis of Hirano and Imbens (2004), where the propensity score method is extended in a setting with continuous treatment to estimate the dose-response function (DRF), $\mu(t) = E[Y_i(t)]$.

The key assumption is the generalization of the unconfoundedness hypothesis for the binary treatment to the multivalued case:

$$Y(t) \perp T|X \qquad \text{for all } t \in T. \tag{3.1}$$

This is referred to as weak unconfoundedness by Hirano and Imbens (2004), since it doesn't require joint independence of all potential outcomes, $\{Y(t)\}_{t\in \mathrm{T}}$. However, it does require conditional independence for each value of the treatment.

Next step is the definition of the *Generalized Propensity Score* (GPS) as $R = r(T, X)$, where

$$r(t, x) = f_{T|X}(t|x)$$

is the conditional density of the treatment given the covariates. The generalized propensity score is calculated at each corresponding level of treatment, so that there are as many propensity scores as there are levels of treatment.

Together with the balancing property of the GPS, that assumes that within strata with the same value of $r(t, x)$ the probability that $T = t$ doesn't depend on the value of *X*, i.e. $X \perp \mathbb{1}\{T = t\}|\, r(t, X)$, the unconfoundedness assumption implies that assignment to treatment is unconfounded given the GPS. Thus, for every *t*,

$$f_T\big(t|r(t, X), Y(t)\big) = f_T\big(t|r(t, X)\big) \ . \tag{3.2}$$

This is the assumption of *weak ignorability of the treatment assignment given the propensity score*.

The GPS estimated can be used to eliminate any bias associated with differences in the covariates. The estimation of the parameter of interest $\mu(t) = E[Y(t)]$ is obtained with a two-step procedure that first estimates the conditional expectation of the outcome given the treatment level *T=t* and the GPS *(i)* , then averages the estimated conditional expectation in *(i) over the distribution of the pre-treatment variables*, to obtain the dose-response function at that particular level of treatment *(ii):*

(i) $\beta(t, r) = E[Y(t)|r\ (t, X) = r] = E[Y|T = t, r\ (t, X) = r]$
(ii) $\mu(t) = E[Y(t)] = E[\beta\big(t, r(t, X)\big)]$.

What has been proposed by the authors is a parametric procedure for which both the estimation and the inference problems are handled with a parametric approach,

while the basic framework is more general, thus nothing prevents the implementation with a more flexible approach, as stated by the authors.
In the first stage, it is suggested to use a normal distribution for the treatment given the covariates:

$$r(T,X) = T_i|X_i \sim N(\beta_0 + \beta_1' X_i, \sigma^2).$$

The estimated GPS is:

$$\widehat{R_i} = \frac{1}{\sqrt{2\pi\widehat{\sigma^2}}}\exp(-\frac{1}{2\widehat{\sigma^2}}\ (T_i - \hat{\beta}_0 - \hat{\beta}_1' X_i)^2$$

where $\beta_0, \beta_1,$ and σ^2 are estimated by maximum likelihood.

In the second stage the conditional expectation of Y_i given T_i and R_i is estimated using a quadratic approximation:

$$E[Y_i|T_i, R_i] = \alpha_0 + \alpha_1 T_i + \alpha_2 {T_i}^2 + \alpha_3 R_i + \alpha_4 {R_i}^2 + \alpha_5 T_i R_i \qquad (3.3)$$

The parameters of equation (3.3) are estimated by OLS using the estimated GPS $\widehat{R_i}$.

The last step estimates the average potential outcome at treatment level t as:

$$\widehat{E[Y(t)]} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\alpha}_0 + \hat{\alpha}_1 t + \hat{\alpha}_2 t^2 + \hat{\alpha}_3 \hat{r}(t, X_i) + \alpha_4 \hat{r}(t, X_i)^2 + \alpha_5 t\ \hat{r}(t, X_i)) \qquad (3.4)$$

To obtain an estimation of the entire dose-response function (DRF) this expected mean can be estimated for each level of treatment one is interested in.
Standard errors can be estimated using the bootstrap method on the entire process. Most of the time we are interested in estimating the relative dose-response function, that is $E\{Y(t) - Y(0)\}$, comparing the average outcome under each treatment level t with the control level $t=0$. However, the control group can be non-existent or not have a proper meaning, leaving a more interpretable result being expressed only by $E\{Y(t)\}$.

3.3. The Propensity Function

Another way to generalize the propensity score approach has been proposed by Imai and van Dyk (2004), through the Propensity Function (PF). They proposed a method capable of establishing causal estimates in observational studies, encompassing what was proposed by Hirano and Himbens in 2004. Relying on SUTVA and strong ignorability of treatment assignment, they defined the PF as the entire conditional probability density (or mass) function of the multivariate treatment T_i given the covariates X_i, i.e. $\pi_\varphi(.|X_i) = p_\varphi(.|X_i)$, parametrized by φ. This is analogous to the propensity score for the binary treatment, where $\pi_\varphi(.|X_i)$ is exactly determined by $p_\varphi(T = 1|X_i)$. Thus, Imai and van Dyk made an extra assumption of Uniquely Parametrized Propensity Function, in order to summarize the PF. This assumption states that for every value of X, there exists a unique finite-dimensional parameter, $\theta \in \Theta$, such that $\pi_\varphi(.|X_i)$ depends on X_i

only through $\theta_\varphi(X_i)$. In other words, the propensity function θ uniquely represents $\pi_\varphi\left(.\left|\theta_\varphi(X_i)\right.\right)$, that could be written as $\pi(.|\theta)$, for example, if we use a normal linear model for the treatment, $T_i \sim \mathcal{N}({X_i}'\beta, \sigma^2)$.
With $\varphi = (\beta, \sigma^2)$, then $\pi_\varphi(.|X_i)$ is uniquely represented by the scalar $\theta_i = {X_i}'\beta$.

Imai and van Dyk proposed different uses of the PF. One possibility is the sub-classification or stratification, that estimates the causal effects within these subclasses and obtain the overall treatment effect by a weighted average of the estimates across strata, using weights based on the number of observations in each subclass. The hypothesis under this method is that, if the PF is correctly specified, then the covariates are distributed evenly across groups defined by its quartiles. Within each subclass, the outcome can be modelled in a standard way as a function of the treatment.

A second approach proposed by Imai van Dyk is a smooth coefficient model, which allows the intercept and the slope to vary smoothly as a function of the PF:

$$\hat{E}(Y|T,\hat{\theta}) = f\left(\hat{\theta}\right) + g\left(\hat{\theta}\right) * T, \qquad (3.5)$$

where $f(\cdot)$ and $g(\cdot)$ are unknown but smooth continuous functions. This could be fit using cubic regression splines as the basis functions.
Computing the average causal effect in this last approach involves averaging $g\left(\hat{\theta}\right)$ across all units.

3.4. Extension to Hirano and Imbens GPS

Flores, Flores-Lagunes, Gonzales, and Neumann (2012) proposed two extensions to the method of Hirano and Imbens, aiming to overcome the bias resulting from misspecification of (2.3), by using a more flexible response model.
The first extensions generalizes (3.3) with $E\left(Y|T,\hat{R}\right) = \beta\left(T,\hat{R}\right)$, where $\beta\left(T,\hat{R}\right)$ is a flexible nonparametric model. The DRF is obtained through (3.4) with

$$\widehat{E\{Y(t)\}} = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}[t,\hat{r}(t,Xi)] \quad . \qquad (3.6)$$

The second method involves inverse weighting (IW) and nonparametric estimators based on a kernel function. The method starts from the assumption that the Horvitz-Thompson (1952) weighting can be applied to a continuous treatment by weighting units with the inverse of the GPS at any given level of treatment, and then smoothing over the levels with a kernel approach. The estimated DRF is

$$\hat{\mu}(t) = \frac{\sum_{i=1}^{N} K_h(T_i - t) Y_i w_i(t)}{\sum_{i=1}^{N} K_h(T_i - t) w_i(t)} \qquad (3.7)$$

where $K_h(T_i - t)$ is a kernel function (Gaussian kernel, triangular kernel density, or other shape), in which h is a bandwidth around t tending to 0 as $N \to \infty$. The

kernel function is used to give more influence to units closer to t than farther away.
In the variation proposed by Flores et al. (2012) each individual kernel weight is divided by its GPS at t. In particular let's consider $\widetilde{K}_h(T_i - t) = \frac{K_h(T_i - t)}{\hat{r}(t, \boldsymbol{X}_i)} = K_h(T_i - t)\widehat{w}_i(T_i = t)$, then an estimate of $\mu(t)$ is

$$\hat{\mu}(t) = \frac{\sum_{i=1}^{N} \widetilde{K}_h(T_i - t) Y_i}{\sum_{i=1}^{N} \widetilde{K}_h(T_i - t)} \tag{3.7}$$

This is an adaptation of the Nadaraya-Watson estimator (Nadaraya, 1964), that is a local constant regression weighted by the inverse of the GPS.

Another estimator proposed by Flores et al. (2012) is a local linear regression of Y on T with a weighted kernel function $\widetilde{K}_h(T_i - t)$:

$$\hat{\mu}(t)_{IW} = \frac{D_0(t)S_2(t) - D_1(t)S_1(t)}{S_0(t)S_2(t) - S_1^2(t)} \quad , \tag{3.8}$$

where $S_j(t) = \sum_{i=1}^{N} \widetilde{K}_h(T_i - t)\,(T_i - t)^j$ and $D_j(t) = \sum_{i=1}^{N} \widetilde{K}_h(T_i - t)\,(T_i - t)^j Y_i$.

Moodie and Stephens (2012) extended the HI method to the longitudinal setting to estimate the direct effect of a continuous treatment on a longitudinal response, where repeated measures of outcome and treatment may be present. They formulated an approach that deals with GPS for analysis, with repeated measures response data with interval dependent treatments. In their study they considered two time intervals and estimated the potential outcome in the first interval as a predictor for the second interval.

3.5. Extension to Imai and van Dyk

A modification of the Imai and van Dyk's method has been suggested by (Zhao, van Dyk, & Imai, 2013) in order to extend the propensity function estimator from the average effect to the dose-response curve estimation. In fact, the reason behind this method is the need to make the DRF more robust and flexible.

The new method considers a smooth function of the PF that is less rigid than subclassification of its values. To estimate the DRF the predictions are averaged over the empirical distribution of X_i for each value t of interest.

In details, the method first estimates the propensity function $\hat{\theta}_i = \hat{\theta}_\varphi(X_i)$ as a function of the treatment given the covariates.

Then it fits the model $\hat{E}\left(Y_i(T_i) \middle| T_i = t, \hat{\theta}_i\right) = f(\hat{\theta}_i, T_i)$, and estimates the smooth function $f(\cdot)$. The dose-response function is obtained by averaging over the empirical distribution of $\hat{\theta}$ at each value t of interest:

$$\hat{E}[Y(t)] = \frac{1}{N}\sum_{i=1}^{N} \hat{f}(\hat{\theta}_i, t). \quad (3.9)$$

The most important difference with the Hirano and Imbens's DRF is the use of the PF instead of the GPS in (3.9). The authors demonstrated that the θ is a much better predicted variable than the GPS, since the strong dependence of this last one from the treatment value t, together with the non-monotonicity of this dependence, make the model more complex and less robust in the estimates.

3.6. Importance Sampling

In the setting of the marginal structured models, Robins et al. (2000) applied the weighting approach to the continuous treatment. In order to obtain unbiased estimates of the effect, the authors proposed to use stabilized weight, under the assumption of normal distribution of the treatment, defined as $w_i(t) = g(T_i)/P(T_i = t|X_i)$, where $g(T_i)$ is the marginal density for T_i under normal distribution, and the denominator is the conditional density of the continuous variable T_i given X_i. To estimate $g(T_i)$ one can reasonably use the density of a normal variate with mean $\bar{T} = \frac{1}{N}\sum_{i=1}^{N} T_i$ and variance $\frac{\sum_{i=1}^{N}(T_i-\bar{T})^2}{(N-1)}$, and to estimate $P(T_i = t|X_i)$ by the normal density with mean $\hat{T}_i$ and variance $\frac{\sum_{i=1}^{N}(T_i-\hat{T})^2}{(N-m-1)}$, where $\hat{T}_i$ is the fitted value from the OLS regression of T_i on covariates. If no stabilization is performed in the continuous setting, then the unstabilized weights may have infinite variance and thus cannot be used.

The numerator of the stabilized weights instead is meant to adjust the fit to what it might have been, had the treatment been assigned independently from the covariates.

3.7.Covariate Balance and Common Support

One important feature of the analysis with the propensity score, is the check of the correct specifications of the propensity score model.
The quality of the PS depends on two major aspects: having the important and relevant covariates included in the model, and having a correct specification of the functional form of the covariates in the model (Guo S & Fraser, 2010). Correctly specified PS are able to successfully eliminate bias.

Although a unique test to assess the accuracy of model specification does not currently exist, two central properties of the PS can be examined to determine it: the covariate balance property and the common support region.

The idea of the balanced property is that the estimated PS and covariates should be balanced between treated and control units. Assessing balance involves comparing the distribution of the measured baseline covariates in the treated and control units, before and after conditioning on PS. This can be done by using standardized mean differences, statistical tests or graphical representations (Austin, 2011).

In the continuous treatment setting, there exist different methods for checking covariate balance. Hirano and Imbens (2004) test the balance property dividing the treatment values into K intervals according to the sample distribution (it can be quartiles), obtaining a series of binary treatments. Within each treatment interval the GPS is computed at a representative point of the treatment variable in the group, $\tilde{T}_k$, (it could be the mean, the median or another percentile) for each unit. The GPS value obtained at $\tilde{T}_k$ will be then defined as $\hat{r}(\tilde{T}_k, X)$. The covariate balance is checked for these intervals first by stratifying units on the scores $\hat{r}(\tilde{T}_k, X)$ in m intervals, defined by the quantiles of its distribution. Then within each m interval, two-sample t-tests are performed to compare the mean of each covariate between units that belong to the k treatment interval and units that are in the same group of GPS, but belong to another treatment interval. Finally, the within-strata differences in means and variances are combined to compute a single t-statistic for each covariate, by weighting with weights given by the number of observations in each GPS interval.

As pointed out by Imai (2008) and Austin (2011), the non-statistical significance in the tests doesn't imply lack of balance, since a small sample size can limit the ability to catch an imbalance in the covariate. This test-based diagnostic should be then taken with great care, being also the balance property referred to a particular sample and not easily and appropriately inferred to the total population (Imai et al., 2008).
Another method by Flores et al (2012) instead compares unrestricted (with all the covariates) and restricted (that sets the coefficient of all the covariates to 0) models by using likelihood ratio tests.

Imai and van Dyk (2004) suggest checking the balance by regressing each covariate on the treatment and comparing the distribution of t-statistics for each of the regression coefficients before and after conditioning on PF.

To guarantee a common support for units with different treatment levels, these ones are first binned, and then the GPS values are checked for each treatment level (Bia, Flores, Flores-Lagunes, & Mattei, 2014).

The procedure starts by first binning the treatment in K classes based on the treatment level, defined by the percentiles Q of its distribution. Then the GPS is computed at the median level of treatment in a specific percentile Q_k for each unit i, $\widehat{R_i^k} = \hat{r}(\tilde{T}_{ki}, X)$. The common support region with respect to a specific percentile is obtained by comparing the support of the distribution of $\widehat{R_i^k}$ for those units belonging to the percentile Q_k to that of units not belonging to that percentile. The common support region is defined by Bia et al (2014) as:

$$CS = \cap_{q=1}^{K} \left\{ i: \widehat{R_i^q} \in [max \left\{ \min_{\{j:\, Q_{j=q}\}} \hat{R}_j^q \, , \min_{\{j:\, Q_{j\neq q}\}} \hat{R}_j^q \right\}, \right.$$
$$\left. min \left\{ \max_{\{j:\, Q_{j=q}\}} \hat{R}_j^q \, , \max_{\{j:\, Q_{j\neq q}\}} \hat{R}_j^q \right\}] \right\}.$$

Units that fall in the overlapping region are kept and the others discarded. This process is repeated for each treatment class and the units not discarded are used in the analysis.

3.8.Conclusions

The continuous treatment problem is relatively poorly studied and is significantly more complicated than the binary one.
Hirano and Imbens (2004) proposed one of the most important work on continuous treatment, providing an estimate of a measure of effect at each level of the treatment. The focus of the work is to provide a dose-response curve (DRF) of the potential outcome given different levels of propensity score, that is not by definition an estimation of the average effect of the treatment.

Traditionally, the research on the potential outcomes aims at estimating an effect by comparing the treated and not-treated group. Instead in this work, a proper estimation of this kind of effect is not possible to obtain. Rubin (1974) stated that we need the information of the non-treated units in order to obtain some value of their potential outcomes, while in Hirano and Imben's work (2004) no comparison has been performed, since only information about treated units are used being in the continuous treatment framework.
One thing to consider about Hirano and Imbens' method is that their weak ignorability assumption of the treatment given the GPS doesn't guarantee that T will be uncorrelated with the potential outcomes, since to achieve conditional independence we would need to condition on the entire family of GPS, not only to

the values observed for the specific treatment level. That implies that not all the response models can be used. Stratification for example is not feasible, unless the classifying variable is low dimensional (Zhao et al., 2013). In fact, subclasses are defined by similar values for a single propensity score at a particular treatment level so that we can estimate the average potential outcome for that treatment level within the subclasses, and we do so separately for each treatment level, with different subclasses for each treatment level.

Imai and van Dyk (2004) proposed a method to estimate the average causal effect, that makes the most important difference with Hirano and Imbens's approach. What can be pointed out is that in this method the stratification on the PF implies a proportional structure of the levels of the treatment that might be not true. Other specifications, like the smooth approach, might be more suitable in some cases.

One interesting property of their method, however, is that since we can fit a response model conditional on T within each subclasses of PS, then we can all the same average these fitted models and estimate the DRF (Zhao et al., 2013).

CHAPTER 4. Cohort Study on Industrial Air Pollution Exposure from a Large Steel Plant and the Risk of Mortality and Morbidity in the Nearby Population in the Taranto Area, Apulia Region, Italy

4.1. Introduction

The city of Taranto is located in in the Apulia region in Italy and is one of the most industrialized areas in southern Italy and Europe. The area has been included among the 14 sites of national interest requiring remediation, and defined in the 1990s as an "area at high risk of environmental crisis" by the Italian government, for the massive presence, since the early 1960s, of industrial sites, including one of the largest steel plants in Europe, mineral deposits, oil refining, cement production, fuel storage, power production, mining industry, military plants and the harbour.

Several environmental monitoring studies and measurement campaigns of industrial emissions in the Taranto area showed a critical framework of environmental pollution and the relevant contribution of the steel factory on the levels of pollutants responsible of adverse health effects (Brand, Pulles, Gijlswijk, Fribourg-Blanc, & Courbet, 2004).

Epidemiological studies conducted in the area revealed high mortality risks in Taranto, for all causes, all cancers and in particular for lung, pleura, bladder, lympho-hematopoietic system and, respiratory disease and pneumonia (Graziano, Bilancia, Bisceglia, Pollice, & Assennato, 2009; Martinelli et al., 2009; Martuzzi, Mitis, Biggeri, Terracini, & Bertollini, 2002; Mitis, Martuzzi, Biggeri, Bertollini, & Terracini, 2005; M A Vigotti, Cavone, Bruni, Minerba, & Conversano, 2007). The most recent study SENTIERI lead by the National Institute of Health (ISS) analysed mortality of the population living near "sites of national interest for environmental remediation" (SIN). The study found excesses in mortality in the area of Taranto and the small municipality of Statte than expected for lung and pleura cancer, non-Hodgkin's lymphoma, cardiovascular and respiratory diseases, and digestive disease for both genders (Comba et al., 2012).

Most and less recent multicentre studies on short term effects of air pollution in Italy (MISA, EPIAIR1 and EPIAIR2), in terms of mortality/morbidity, found increases in risks in this area for all causes, cardiovascular and respiratory diseases (Alessandrini et al., 2013; Berti, Galassi, Faustini, & Forastiere, 2009; Biggeri, Bellini, & Terracini, 2004; Colais et al., 2009; Scarinzi et al., 2013; Stafoggia et al., 2009).

These findings, together with the massive industrial emissions and several high pollution episodes, have generated considerable echo and public concerns regarding the possible connection between adverse health outcomes and air pollution from the plant.

In 2011, the Court of Taranto requested a survey to evaluate emissions from the steel plants and an epidemiological study to investigate on the health status of the resident population in relation to the emissions of the plant (Forastiere, Biggeri, & Triassi, 2012; Sanna, Monguzzi, Santilli, & Felici, 2012). The results of these reports led the closure of many sectors of the steel plant ordered by the Taranto Court. The epidemiological study used the current cohort of residents from 1998 to 2010, and the analyses were performed by districts and socioeconomic position using Cox Proportional Hazard Models. Published results (Mataloni et al., 2012) showed greater effects on the health of the population residing in neighbourhoods closest to the industrial area, like Paolo VI, built in the late 1960s mainly as a residential area of the steel plant's workers, Tamburi, close to the mineral deposits, and Borgo. All these areas are the most polluted, according to the spatial distribution of pollutants in the area (Mangia, Gianicolo, Bruni, Vigotti, & Cervino, 2013). People living in these areas showed higher levels of mortality, especially for all malignant neoplasms, cancer of pancreas and lung, cardiac, respiratory, and digestive system diseases. Risks were higher in the Paolo VI district compared to the others. Low socio-economic position was associated with higher risks of mortality for all causes, cancers, and cardiovascular and respiratory diseases. On the same cohort of residents, followed-up until 2010, was based a more recent study (Maria Angela Vigotti, Mataloni, Bruni, Minniti, & Gianicolo, 2014), that explored the mortality risk in all the neighbourhoods of the Taranto area, including the industrial ones. The results of this study confirmed findings from Mataloni et al.(2012), showing higher risks in the neighbourhoods closest to this area and the highest increases in the Paolo VI district.

The present study used the same cohort of residents enrolled from 1998 to 2010 (Mataloni et al., 2012), followed up until 2014 and with mortality by cause assessed until 2013. The Mataloni et al. study, like many longitudinal European studies (Beelen et al., 2008; Beelen et al., 2014; Bentayeb et al., 2015; Giulia Cesaroni et al., 2013; Jerrett et al., 2009), included exposure estimated at individual level, but only at the study inception (baseline), as a marker for long term exposure.

Few longitudinal studies until now dealt with time-varying exposure assessment (Bentayeb et al., 2015; Lepeule, Laden, Dockery, & Schwartz, 2012; Tétreault et al., 2016; Wahida et al., 2016) for every year of follow-up, and among them only two studies accounted for all the residential history of participants (Bentayeb et al., 2015; Wahida et al., 2016). However, in both these studies, exposure assessment was at the zip code and census block level, respectively, and not at the residential address.

Purpose of the Study

In the present study, we incorporated into the exposure assessment the exposures associated with annual residential mobility patterns in the study period, in a time-dependent manner, considering also the past (before the baseline) residential history, which was available at the individual address for each year of permanence in the area.

The aim was to assess whether exposure to industrial emissions was related to increased mortality from all natural causes, and cardiovascular and respiratory conditions, as well as to increased cardiorespiratory morbidity in the nearby population. Mortality and incidence of lung cancer was also of interest. These conditions were of *a priori* interest because of the large literature on the health effects of fine particles, but few studies on populations living near industries. A large population-based cohort study has been conducted with a long follow-up of mortality, hospital admissions, and cancer incidence; exposure to industrial emission was estimated retrospectively for all the subjects.

A cohort study in environmental epidemiology evaluates the association between environmental exposures and health effects while controlling for potential measured confounders. In a residential cohort design, where all the residents in a specific area are selected instead of a population sample, the information at the individual level about all possible confounders is not available, since this could only come from a detailed survey with a direct questionnaire to every participant. One relevant unmeasured confounder is for example smoking habit, an important risk factor for most of the health outcomes of interest, and potentially heterogeneously distributed among those more or less exposed to air pollution.

In this study, a first causal inference exercise was performed to gauge the potential bias due to unmeasured confounders, strictly related to lifestyle. The relationship between exposure to industrial pollutants and the prevalence of smoking habit, alcohol consumption and obesity was explored following the indirect adjustment approach by Schneeweiss et al. (Schneeweiss, Glynn, Tsai, Avorn, & Solomon, 2005).

4.2.Methods

Study Design

The cohort of residents in the study area was enrolled using municipality data of Taranto and the two nearby towns of Massafra and Statte, located in the Apulia region. The cohort included all residents in the area as of January 1st 1998 and all subjects entering later for immigration or birth until December 31st 2010. The data source for the enrolment was the General Registry Office of the three cities. The municipality datasets underwent several quality controls, such as elimination of residents registered to AIRE (Register of Italians Resident Abroad) and double records. By using municipality data, vital status (alive, dead, migrant) of each

participant as of December 31st, 2013 was assessed. The municipalities databases contained for each person all changes of residence until the end of 2010. Each address was geocoded using ArcGIS software. A census block-level socioeconomic position (SEP) (Caranci et al., 2010) index was assigned to each participant based on his/her geocoded address. This index was composed by information recorded at the 2001 census: percentage of population with educational level equal to or less than primary school, percentage of the active population unemployed or looking for their first occupation, percentage of rented houses, percentage of single parent families, and population density (number of occupants per 100 m^2). For the present study, the index has been classified into 5 quintiles, representing high, middle-high, medium, middle-low and low SEP.

Information about cause of death (1998-2013) and hospitalization diagnoses (1998-2014) was retrieved from the Regional Health Databases, while cancer incidence was extracted from the Regional Cancer Registry. Events of mortality/morbidity, with corresponding dates and causes of death, principal diagnosis of the first hospital discharge (ICD-9 CM and ICD-10 revisions), and cancer diagnosis for the period 2006-2011 (ICDO-3T), were assigned to each person enrolled, upon record-linkage procedures between the cohort file and the health databases.

Air Pollution Exposure Assessment

Taranto is one of the most highly industrialized cities in Italy, where shipping, industries and urban activities co-exist. The main industrial activity is the integrated steelwork ("ILVA"). The areal source emissions are the hot emissions from the steelwork, the cold emissions generated by the fossil fuels processing products, the wind erosion from the large coal mining park, the handling on the conveyors and material transport.

Daily simulations were carried out to estimate three-dimensional concentrations of pollutants through a Lagrangian modelling system, developed by the Regional Environmental Protection Agency (ARPA) of the Apulia region for the year 2010. The dispersion model assessed the impact of the harbour, industrial and urban activities on air quality in the Taranto area. The modelling system included the SWIFT meteorological model, the SURFPRO turbulence pre-processor and the SPRAY Lagrangian particles dispersion model (Giua et al., 2014). SPRAY is a 3D model particularly suited to provide an accurate local distribution of the primary pollutants in the atmosphere in non-homogeneous and non-stationary conditions. The meteorology in the area was built with the SWIFT and SURFPRO codes on hourly basis, by using the products for the year 2007 supplied by the MINNI project (Zanini, 2009) as the input. Model results were validated using measured data in 9 fixed monitoring stations of the ARPA network.

In this study PM_{10} (Particulate Matter of less than 10 μm in aerodynamic diameter) and SO_2 (Sulphur dioxide) concentrations from industrial sources were

considered as exposure. The individual exposure to air pollutants was obtained in a multi-step procedure.

1) In the first step, the 2010 PM_{10} and SO_2 concentration maps were overlapped to the cohort geocoded addresses in order to assign to each residence that year's corresponding estimated industrial air pollutant concentration.
2) The ILVA steel plant provided productivity data per kiloton year (kton/a) of steel, coke, cast iron and others from 1965 to 2014. In addition, air pollutants emissions for every process of the industry (coke ovens, sintering plant, blast furnace, steel plant, mining parks and transport of materials) were supplied by the Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA) for the years 1990, 1995, 2000 and 2005. For the purposes of this study, only emissions from the steel production process were considered. Emissions for missing years were estimated by backward and forward interpolation of the emission series from 1965 to 2010, weighted with the productivity ratio of the current year to the previous/following one:
    - *1965-1990*: the emission in year *x* was obtained retrospectively, starting from the first known value in the 1990, weighting with the ratio of the productivity in the current year to the following one (4.1)

$$E(x) = E(x+1) * \frac{P(x)}{P(x+1)} \qquad (4.1)$$

      where x indicates the year, E is the estimated emission and P is the productivity
    - *1991-2010:* the emission in the year *x* was obtained through the interpolation of the emissions in the time intervals, always of a length of 5 years, between two observed knowns values (superior and inferior extreme of the interval, for example 2005 and 2010), weighting them with the relative productivity of the previous year (4.2)

$$E(x) = E(x+1) + \frac{E(sup)-E(inf)}{5} * \frac{P(x)}{P(x-1)} \qquad (4.2)$$

      where *E(sup)* and *E(inf)* represent the emissions at the two extremes of the time interval of 5 year length.
    - *2010-2014:* the emission in the year x was estimated starting form the last value provided in 2010, weighting prospectively with the relative productivity of the previous year (4.3)

$$E(x) = E(x-1) * \frac{P(x)}{P(x-1)} \qquad (4.3)$$

3) Once the annual emission series was complete for the entire period 1965-2014, an annual calibration factor (4.4) was computed as the ratio of

emissions on one year to the emissions in the year 2010, the same of the ARPA dispersion model

$$F_x = \frac{\hat{E}_x}{\hat{E}_{2010}} \qquad (4.4)$$

where x is the year and $\hat{E}$ is the estimated emission.
This factor was computed to modulate the exposure estimated from the dispersion model with the information acquired on the effective productivity and emissions from the steel plant.

4) The yearly calibration factor was then multiplied to the individual exposure from the dispersion model for every residence and every year. In this way we obtained an annual time-varying exposure from the industry based on the spatial pattern of the dispersion model and the temporal pattern estimated by the annual series of the productivity and the emissions of the steel plant.

Occupational Exposure

In Italy, retirement contribution history for workers in private sectors is currently available at INPS (National Social Insurance Agency) databases, where, for each year since 1974, the following information are recorded: amount of contributions paid, length of working period, worker task (blue-collar or white-collar workers) and the company where the activity was performed. Companies are moreover classified according to the branch of economic activity. Using the fiscal code it was possible to link the occupational history to each person enrolled in the cohort. Then people on duty during the period 1974-1997 were selected by branch of activity, so previous occupational history was categorized into five groups:

- Blue collar workers in steel factories;
- White collar workers in steel factories;
- Naval construction workers;
- Mechanical construction workers;
- Workers in other occupational branch or people without contribution payments (reference category).

Statistical Analyses

Each resident contributed to person-years at risk from the date of entry in the cohort (1998 or later within 2010) until date of exit for death, emigration or end of follow-up, whichever came first. The association between long-term exposure to air pollutants (defined as time-varying annual average) and mortality/morbidity was estimated using a survival analysis with multivariate Cox proportional hazard models. Hazard Ratios (HR) and corresponding 95% Confidence Intervals (95% CI) per 10 $\mu g/m^3$ increase of each pollutant were computed. The hypothesis of

proportional hazards was verified by using the command **stphtest** in Stata after estimating the Cox model. This test essentially asks whether or not the slope in the regression of time vs. Schoenfeld's residuals is flat (it should be if PH property holds).
Age was used as the time scale. Observation times were censored at the time of death for causes different from the one under study, emigration, loss to follow-up, or end of follow-up, whichever came first. *A priori* confounders included: gender, area-level socioeconomic position, occupation (recorded at baseline), calendar period in 3 classes (defined as time-dependent).
Air pollution exposure was modelled using alternative time-varying variables based on different time windows, with the aim of exploring effects attributable to different averaging periods: current exposure (lag 0, e.g. average exposure in the current year), and 5-year time-window lagged concentrations over the period considered (1-5, 6-10, 11-15, 16-20, 21-25, 26-30, 31-35, 35+ years). Subjects can belong to more than one time window depending on the length of their residential history of exposure.

Hazard Ratios were expressed for 10 $\mu g/m^3$ increment of pollutants for the average exposures. As an additional analysis, we relaxed the assumption of linearity of the concentration-response function by modelling exposure at lag 0 with a penalized spline with 2 degrees of freedom in the Cox proportional hazard models used for the main analysis.

Effects of air pollutants on mortality and cancer outcomes were analysed on the total cohort. Hospitalisation outcomes were analysed both in the total cohort and in the sub-cohort of children younger than 14 years.

Confounding by Smoking, Alcohol and Body Mass Index

In the present cohort, it was not possible to retrieve information on individual lifestyle factors, as we could not guarantee their random distribution across the PM_{10} exposure levels, thus leading to potential residual confounding and bias of the relationship between exposure and outcome.

The objective was to evaluate this potential bias derived from the lack of information on smoking habit, alcohol consumption and obesity in the cohort by using an ancillary data set of the PASSI (*Progressi delle Aziende Sanitarie per la Salute in Italia*) (Centro nazionale di epidemiologia) national surveillance survey about lifestyles and personal habits for the years 2008-2013, considered representative of the cohort of Taranto, Massafra and Statte. Out of a total of 1,755 subjects recruited in the survey between 2008 and 2013, 620 individuals were members of the cohort; they were geocoded and we assessed the exposure to industrial pollutants (PM_{10} and SO_2). The relationship between exposure to industrial pollutants (independent variables in quartiles) and the prevalence of smoking habit (smokers, ex-smokers vs. never smokers), alcohol consumption (Yes/No) and obesity (BMI>30 vs BMI<=30) was assessed by using a Poisson regression model (Prevalence Rate Ratios) adjusted for age, sex and socio-

economic position, and weighted for the age distribution of the original cohort (in classes 18-34, 35-49, 50-69).

4.3.Results

The area under study is represented in Figure 1, where it is possible to distinguish the administrative boundaries of the three municipalities of Taranto, Massafra and Statte. The area is characterized by the presence of the "ILVA" steel plant (red star), mineral deposits, oil refining, cement production, fuel storage, power production, mining industry, military plants and the harbour.

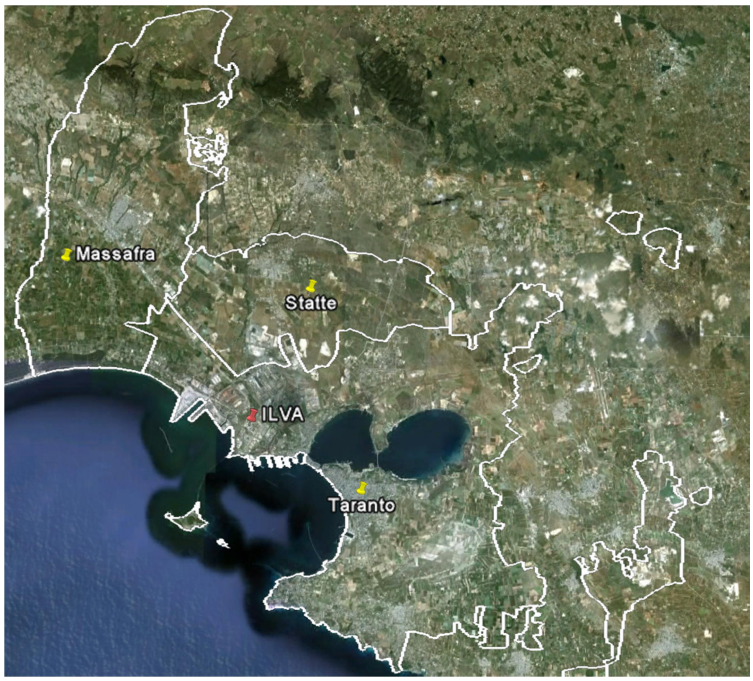

Figure 1: Study area, the municipalities of Taranto, Massafra and Statte, and the localization of the ILVA steel plant

A total of 321,356 individuals were enrolled in the cohort from 1998 to 2010 and followed-up until 2014 (51.1% females), of which 35,398 (11%) died at the end of follow-up (follow-up for mortality is until 2013) and for 27,260 (8.5%) the follow-up ended at the time subjects emigrated.
The main characteristics of the study cohort (age, socioeconomic position, length of residence at recruitment, occupation and vital status) are described in Table 1. Among the subjects recruited, 35% had low socioeconomic position, 34.6% had already been living in the area for more than 30 years, and 32.9% for less than 10 years, at the time of enrolment. There were subjects who had been employed at the iron and steel industry (13,556 subjects), in mechanical (17,035 subjects) and in naval constructions (1,238 subjects).

Cohort characteristics	Categories	N	%
Cohort		**321,356**	**100**
Gender	Males	157,031	48.9
	Females	164,325	51.1
Age class at enrolment (years)	0	34597	10.8
	1-9	29047	9.0
	10-19	36,224	11.3
	20-29	49,652	15.5
	30-39	45,674	14.2
	40-49	37,811	11.8
	50-59	34,213	10.7
	60-69	26,946	8.4
	70-79	18,502	5.8
	>=80	8,690	2.7
Area based socio-economic position	High	68,693	21.4
	Middle-High	39,095	12.2
	Medium	32,736	10.2
	Middle-Low	58,034	18.1
	Low	112,481	35.0
	Missing	10,317	3.2
Length of residence in the area at the enrolment (years)	0-10	105768	32.9
	11-19	46831	14.6
	21-30	50756	15.8
	31-40	111203	34.6
Occupational status			
	No	307,800	95.8
Employment at the Iron and Steel Industry[a]	Yes, laborer	9,633	3.0
	Yes, office worker	3,923	1.2
Employment in the Mechanical Construction Industry[a]	No	304,321	94.7
	Yes	17,035	5.3
Employment in the Naval Construction Industry[a]	No	320,118	99.6
	Yes	1,238	0.4
Vital Status			
	Alive	258,698	80.5
	Dead	35,398	11.0
	Lost to follow-up	27,260	8.5

[a]1974-1997

Table 1: Descriptive characteristics of the cohort (residents in Taranto, Massafra and Statte), study period 1998-2013.

Figure 2 and Figure 3 show the dispersion model concentration maps for PM_{10} and SO_2, and the cohort geocoded addresses in the study area (municipalities of Taranto, Massafra and Statte). The dispersion model displays the result of the Lagrangian particle model (simulated using data of 2010), illustrating higher levels of the two pollutants in the zone immediately surrounding the industrial area.

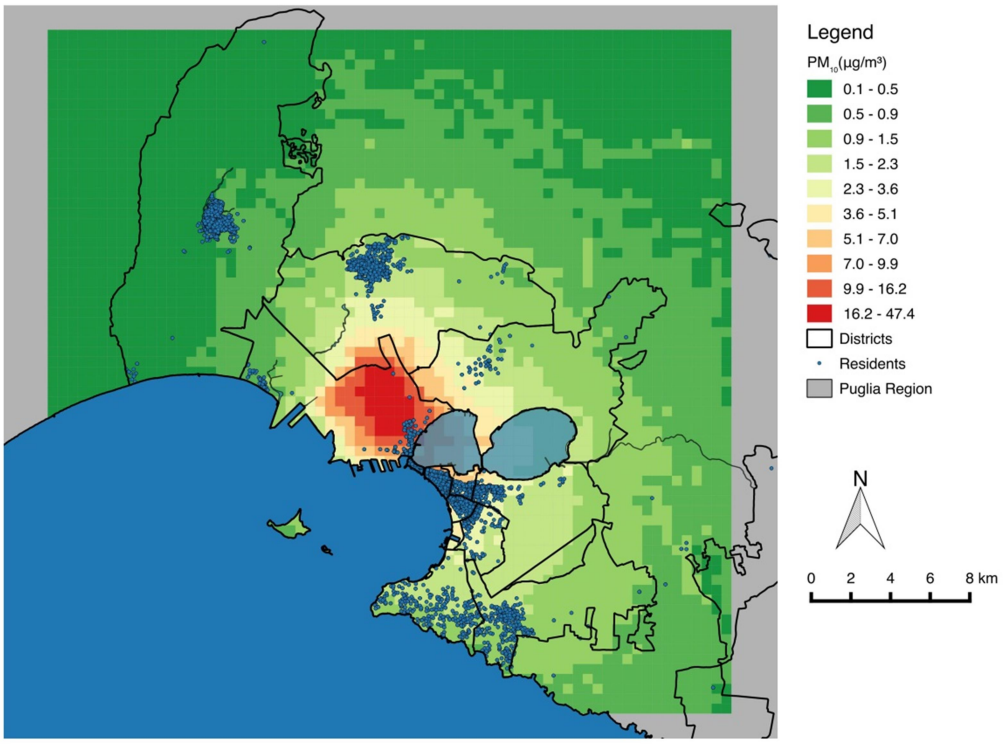

Figure 2: Dispersion model concentration map for PM_{10} (ARPA, year 2010), with geocoded addresses at baseline.

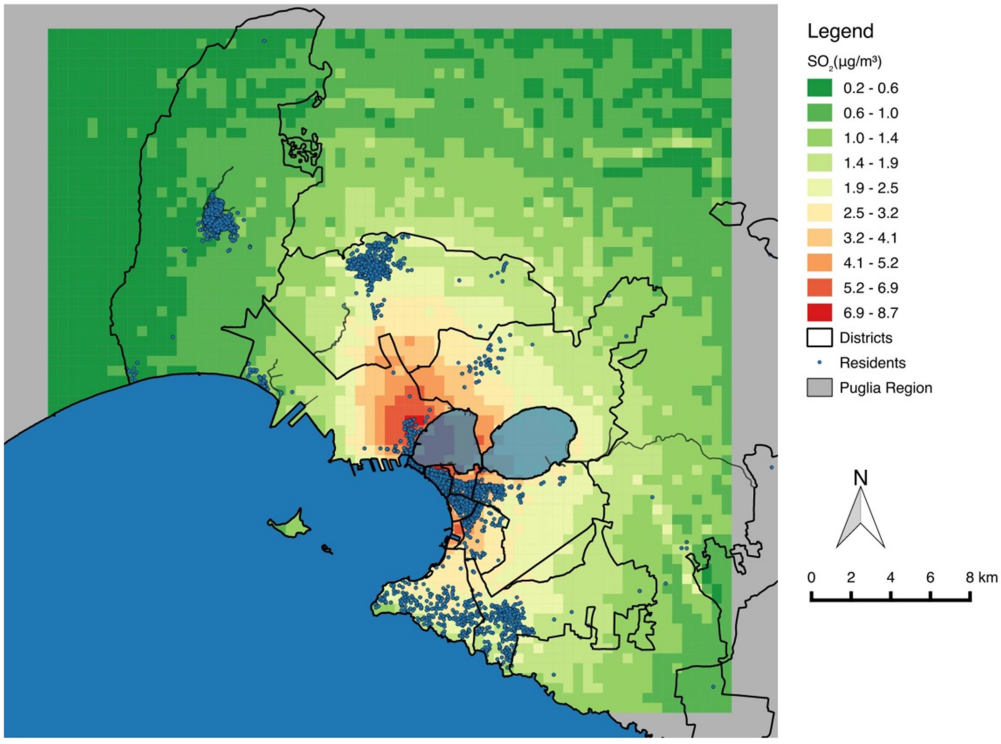

Figure 3: Dispersion model concentration map for SO_2 (ARPA, year 2010), with geocoded addresses at baseline.

The distribution of the socioeconomic position index (SEP) across the census tracts is represented in Figure 4, with values going from 1 (yellow colour, indicating high SEP) to 5 (dark brown, indicating low SEP). Some of the census tracts are not coloured, since the sample size was too small (less than 50 inhabitants) to calculate the index.

Productivity of the steel plant, emissions data and average exposures to PM_{10} and SO_2 at lag 0, resulting from the backward and forward extrapolation procedure, are represented in Figure 5. In the upper part of the figure, productivity and extrapolated emissions from the steel plant processes are plotted together for the period 1965-2013. Emissions follow the trend of productivity until the year 1995, when they decreased, while productivity started to increase until 2008. The average exposures to PM_{10} and SO_2 at lag 0 (in the bottom part of the figure) strictly follow emissions trends and behave similarly for the two pollutants. Both productivity and emissions decreased in the year 2009, possibly due to the economic crisis; consequently, a decrease in pollutant exposure has been observed.

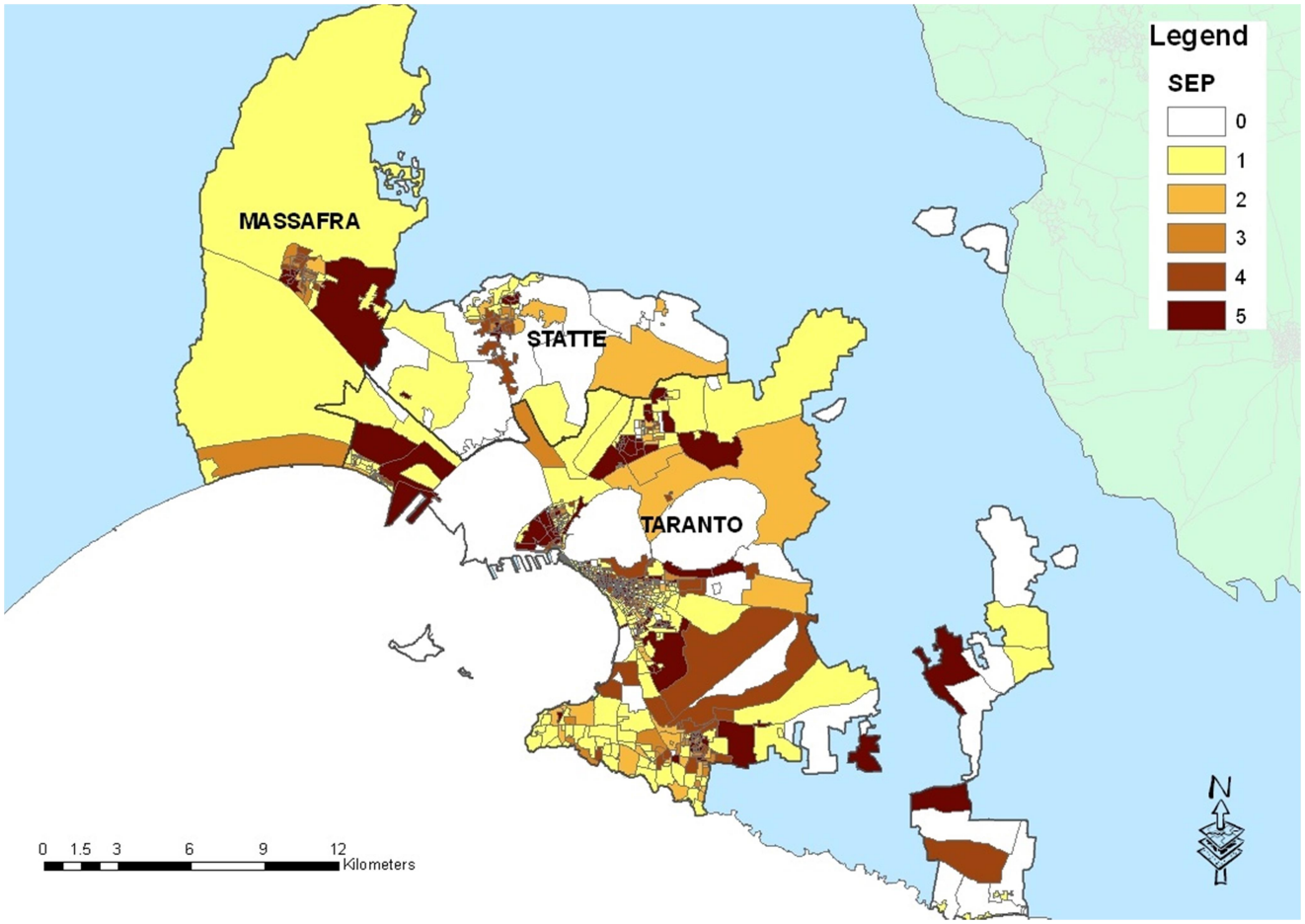

Figure 4: Socioeconomic position index (SEP) by census tract in the municipalities of Taranto, Massafra e Statte

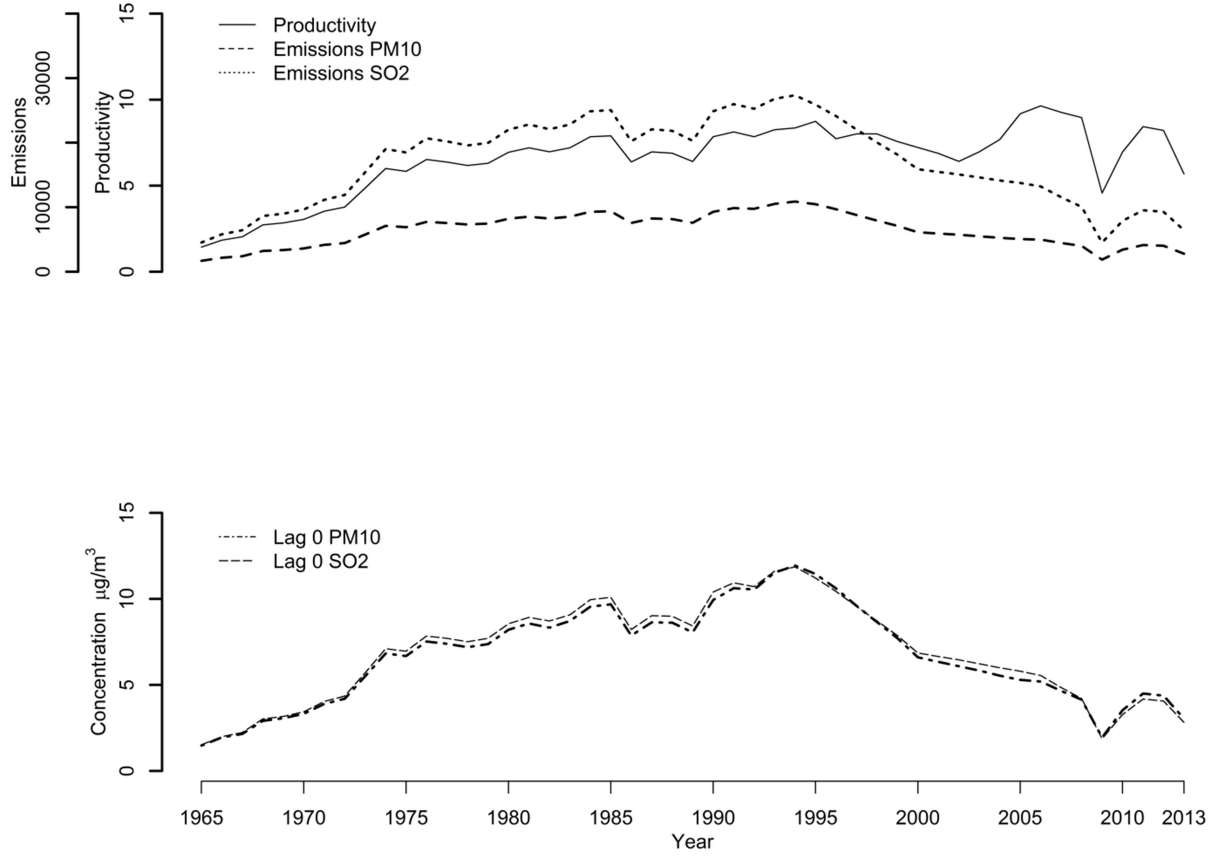

Figure 5: Temporal trend of the steel productivity, emissions and pollutants concentrations at lag 0, study period 1965-2014.

The time-varying exposure to the two pollutants in the past 35 years was attributed to each individual of the cohort. The annual average exposure at lag 0 at baseline (1998) was 9.03µg/m3 (±SD=9.53) for PM_{10} and 9.09µg/m3 (±SD=4.81) for SO_2 among 270,833 cohort members recruited at the start of the cohort (Table 2). There was a high correlation (r=0.7) between the two exposures under study.

Pollutant	Mean	SD	Min	Max	Percentiles				
					5°	25°	50°	75°	95°
PM_{10} (µg/m^3)	9.03	9.53	0	85.24	1.05	3.52	7.79	9.35	30.60
SO_2 (µg/m^3)	9.09	4.81	0	22.08	1.88	5.53	9.27	11.82	18.18

Table 2: Descriptive data of exposures to PM_{10} and SO_2 of industrial origin at lag 0 among 270,833 cohort members at the baseline in 1998.

Exposure to Industrial Pollution and Effects on Mortality

The association between air pollutants and mortality is shown in Table 3. For each 10µg/m^3 increment of PM_{10} and SO_2 exposure at lag 0, there was an increased risk of natural mortality (HR=1.04, CI 95% 1.02-1.06, HR=1.09, CI 95% 1.05-1.12, respectively), particularly mortality from heart disease (HR=1.05, CI 95% 1.02-1.09 and HR=1.11, CI 95% 1.04-1.18, respectively), and from acute myocardial infarction (HR=1.10, CI 95% 1.02-1.19, and HR=1.29, CI 95% 1.10-1.52 for PM_{10} and SO_2, respectively). Malignant neoplasms HR=1.08, CI 95% 1.02-1.15), and in particular lung cancer mortality (HR=1.17, CI 95% 1.03-1.34) showed positive associations with the average concentrations of SO_2. Moreover, a high risk of mortality for kidney diseases was found in relation to PM_{10} (HR=1.13, CI 95% 1.02-1.25 for 10µg/m^3 increase). Only weak associations were detected for respiratory disease mortality, and a negative association was found between both air pollutants and mortality from cerebrovascular diseases.

Causes of death (ICD-9CM)		PM_{10}		SO_2	
	N	HR*	95%CI	HR*	95%CI
Natural mortality (001-799)	33042	1.04	1.02-1.06	1.09	1.05-1.12
Malignant neoplasms (140-208)	10210	1.03	1.00-1.06	1.08	1.02-1.15
Trachea, bronchus, and lung (162)	2164	1.05	0.99-1.12	1.17	1.03-1.34
Bladder (188)	476	1.03	0.90-1.18	0.98	0.74-1.29
Kidney (189)	116	0.95	0.70-1.30	0.81	0.46-1.45
Lymphatic and hematopoietic tissue (200-208)	879	0.98	0.87-1.09	1.04	0.85-1.28
Diseases of the central nervous system (330-349)	1014	1.05	0.951.16	1.05	0.86-1.29
Diseases of the circulatory system (390-459)	12527	1.02	1.00-1.05	1.04	0.99-1.10
Heart diseases (390-429)	8857	1.05	1.02-1.09	1.11	1.04-1.18
Acute myocardial infarction (410-411)	1275	1.10	1.02-1.19	1.29	1.10-1.52
Cerebrovascular disease (430-438)	2903	0.90	0.85-0.96	0.80	0.72-0.89
Diseases of the respiratory system (460-519)	2741	1.02	0.97-1.08	1.02	0.91-1.14
Respiratory infections (460-466, 480-487)	751	0.90	0.80-1.02	0.85	0.69-1.04
COPD (490-492, 494, 496)	1618	1.03	0.95-1.10	1.04	0.90-1.21
Kidney disease (580-599)	707	1.13	1.02-1.25	1.16	0.93-1.45

*Hazard Ratio (HR) from a Cox model stratified for period of follow-up (3 categories) and sex, adjusted for age (temporal axis), socioeconomic position and occupational status

Table 3: Associations between annual average exposure to PM_{10} and SO_2 and mortality. Hazard ratio (HR) per 10 μg/m^3 increase in PM_{10} and SO_2 and 95% Confidence Intervals, CI, 1998-2013.

Linearity of the association between PM_{10} at lag 0 and mortality was confirmed by the linear splines represented in Figure 6 for all the causes studied. The exposure-response relationships between SO_2 and cause-specific mortality displayed some deviations from linearity, though with ample confidence bands at higher concentrations (Figure 7).

The latency of the effects on mortality was analysed estimating different independent models, in which one 5-year time window of exposure at the time was used, with the condition that subjects could belong to more than one time window depending on the length of their residential history of exposure. The effect estimates for these different time-windows of exposure were higher in the most proximal lags (up to 5 years), then decreased so to become almost null and then increased again for exposure occurring in the past, namely 26 years or more. (Figure 8).

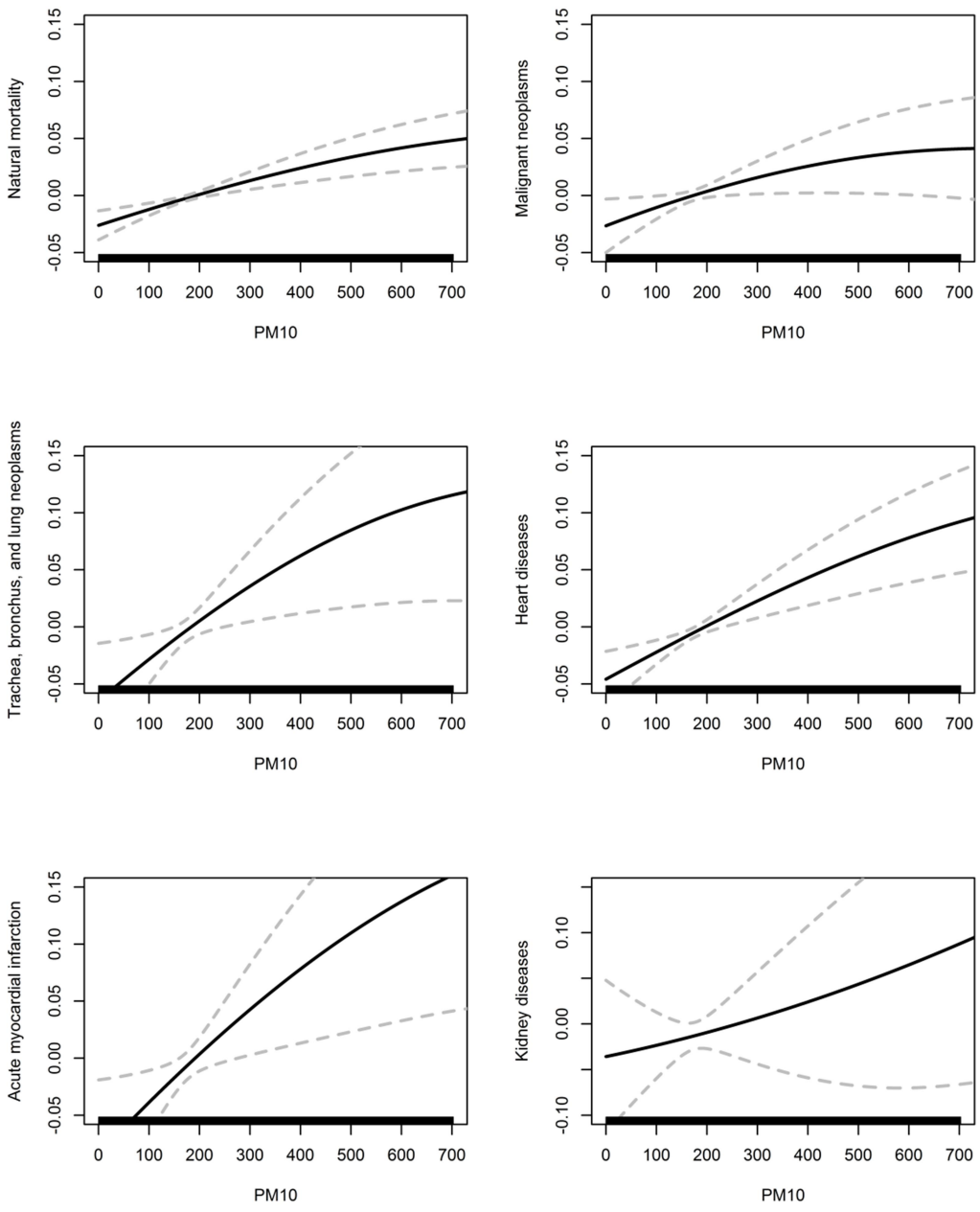

Figure 6: Penalized splines and confidence interval (95%CI) of the relationship between average PM_{10} exposure at lag 0 ($\mu g/m^3$) and mortality for natural mortality, mortality for malignant neoplasms, lung cancer, heart diseases, acute myocardial infarction and kidney diseases.

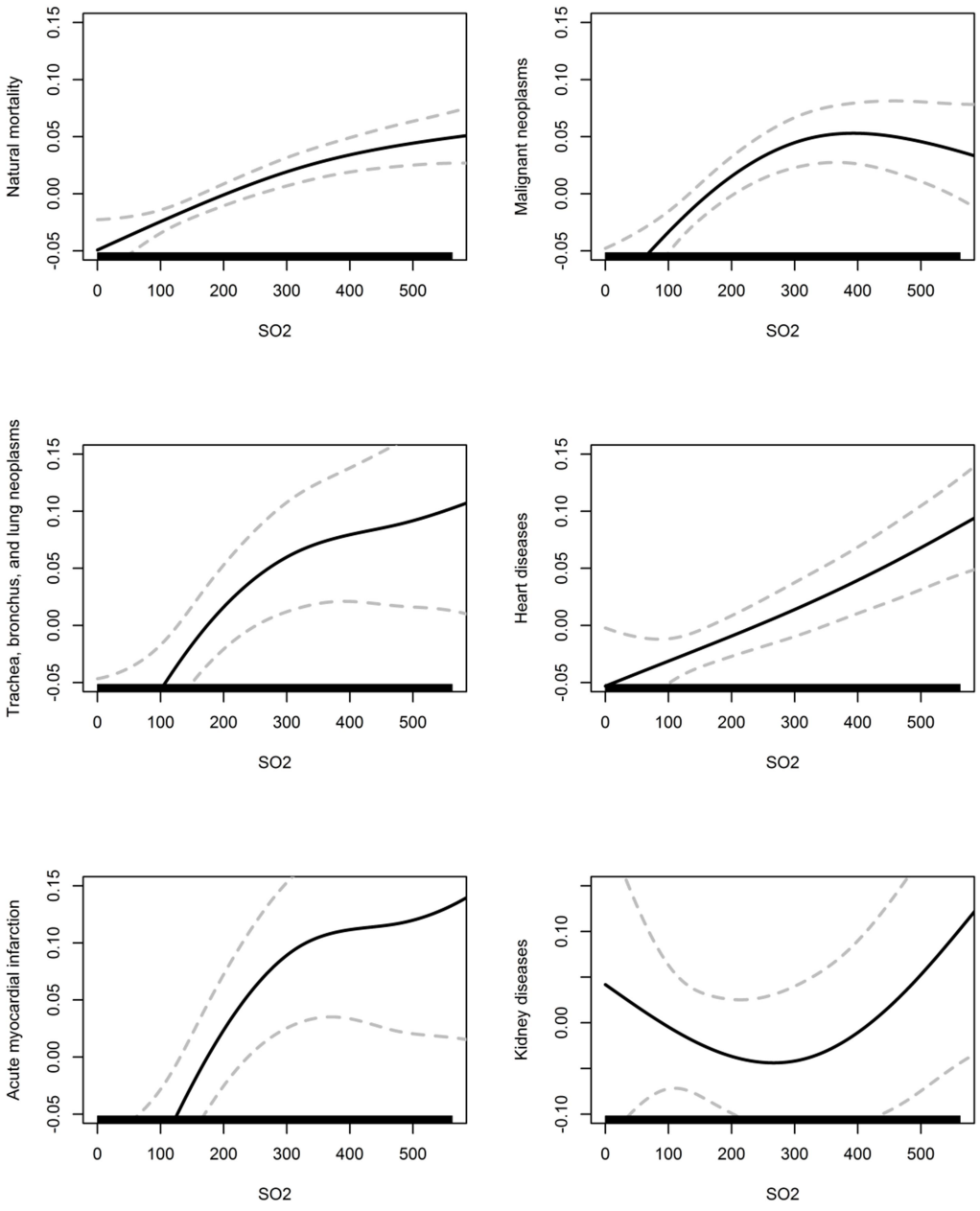

Figure 7: Penalized splines and confidence interval (95%CI) of the relationship between average SO_2 exposure at lag 0 ($\mu g/m^3$) and mortality for natural mortality, mortality for malignant neoplasms, lung cancer, heart diseases, acute myocardial infarction and kidney diseases.

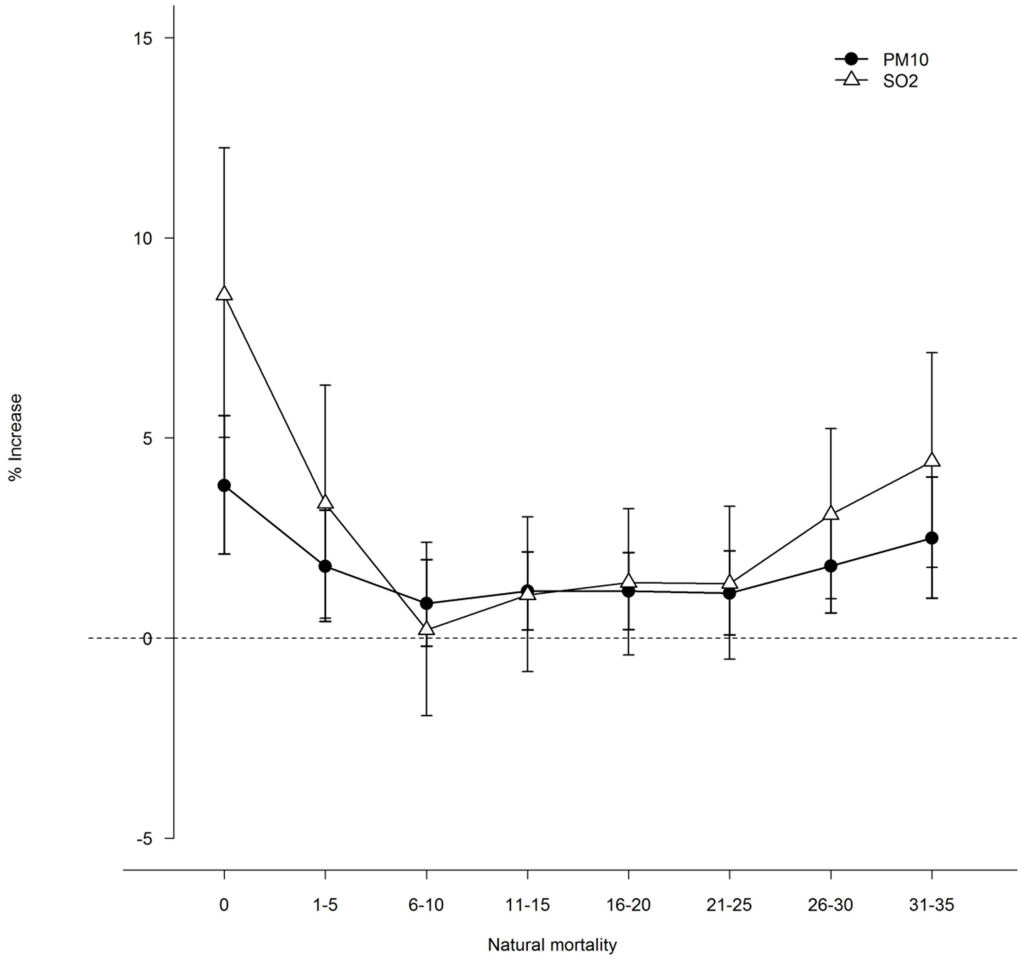

Figure 8: Distribution of the effects of PM_{10} and SO_2 in 5-year time windows on natural mortality. Results expressed as percent increase for 10µg/m^3 increment.

Exposure to Industrial Pollution and Effects on Hospitalizations

Table 4 shows the associations between average exposures to PM_{10} and SO_2 and hospital admissions. Both pollutants were positively associated at lag 0 with several conditions investigated. In particular, increased risks (ranging from 3% to 11% for 10µg/m^3 increment of PM_{10} and 6-35% for 10µg/m^3 increment of SO_2) were found for all natural causes, diseases of the central nervous system, heart diseases, respiratory infections, and kidney diseases. In addition, SO_2 was linked also to hospital admissions for acute myocardial infarction (HR=1.14, CI 95% 1.06-1.23), heart failure (HR=1.13, CI 95% 1.06-1.21).
There were associations with pediatric admissions for diseases of the respiratory system (HR=1.11 and HR=1.33, for 10µg/m^3 increases in PM_{10} and SO_2 at lag 0, respectively) and for respiratory infections (HR=1.15 for PM_{10} and HR=1.49 for SO_2). Also for hospital admissions, the shape of the relationship with pollutants was explored by estimating penalized splines. The Figure 9 shows the linear trend in the effects of PM_{10} and SO_2 on admissions for several conditions.

Diagnosis (ICD-9 CM)		PM_{10}		SO_2	
	N	HR	95%CI	HR	95%CI
All natural causes (001-799)[a]	193277	1.03	1.02-1.04	1.06	1.04-1.07
Diseases of the central nervous system (330-349)	8890	1.05	1.01-1.08	1.21	1.13-1.30
Diseases of the circulatory system (390-459)	49859	1.04	1.02-1.05	1.06	1.03-1.09
Heart diseases (390-429)	34316	1.05	1.04-1.07	1.10	1.07-1.14
Acute myocardial infarction (410-411)	7253	1.02	0.99-1.06	1.14	1.06-1.23
Heart failure (428)	8952	1.02	0.99-1.06	1.13	1.06-1.21
Cerebrovascular disease (430-438)	13236	1.01	0.98-1.04	0.92	0.87-0.97
Diseases of the respiratory system (460-519)	31091	1.07	1.05-1.08	1.15	1.12-1.19
Respiratory infections (460-466, 480-487)	13654	1.11	1.08-1.13	1.35	1.28-1.42
COPD (490-492, 494, 496)	7474	1.03	1.00-1.06	0.95	0.88-1.01
Asthma (493)	885	0.99	0.90-1.09	0.95	0.78-1.16
Kidney disease (580-599)	13184	1.08	1.05-1.11	1.09	1.04-1.15
Population 0-14 years[b]					
Diseases of the respiratory system (460-519)	9505	1.11	1.09-1.14	1.33	1.26-1.41
Respiratory infections (460-466, 480-487)	6746	1.15	1.11-1.18	1.49	1.39-1.59
Asthma (493)	272	0.77	0.60-0.98	0.55	0.37-0.81

[a]Hazard Ratio (HR) from a Cox model stratified for period of follow-up (3 categories) and sex, adjusted for age (temporal axis), socioeconomic position and occupational status

[b]Hazard Ratio (HR) from a Cox model stratified for period of follow-up (3 categories) and sex, adjusted for age (temporal axis), socioeconomic position

Table 4: Associations between average exposures to PM_{10} and SO_2 and hospital admissions. Hazard ratio (HR) per increase of 10 µg/m^3 in average exposure and 95% Confidence Intervals, 1998-2014.

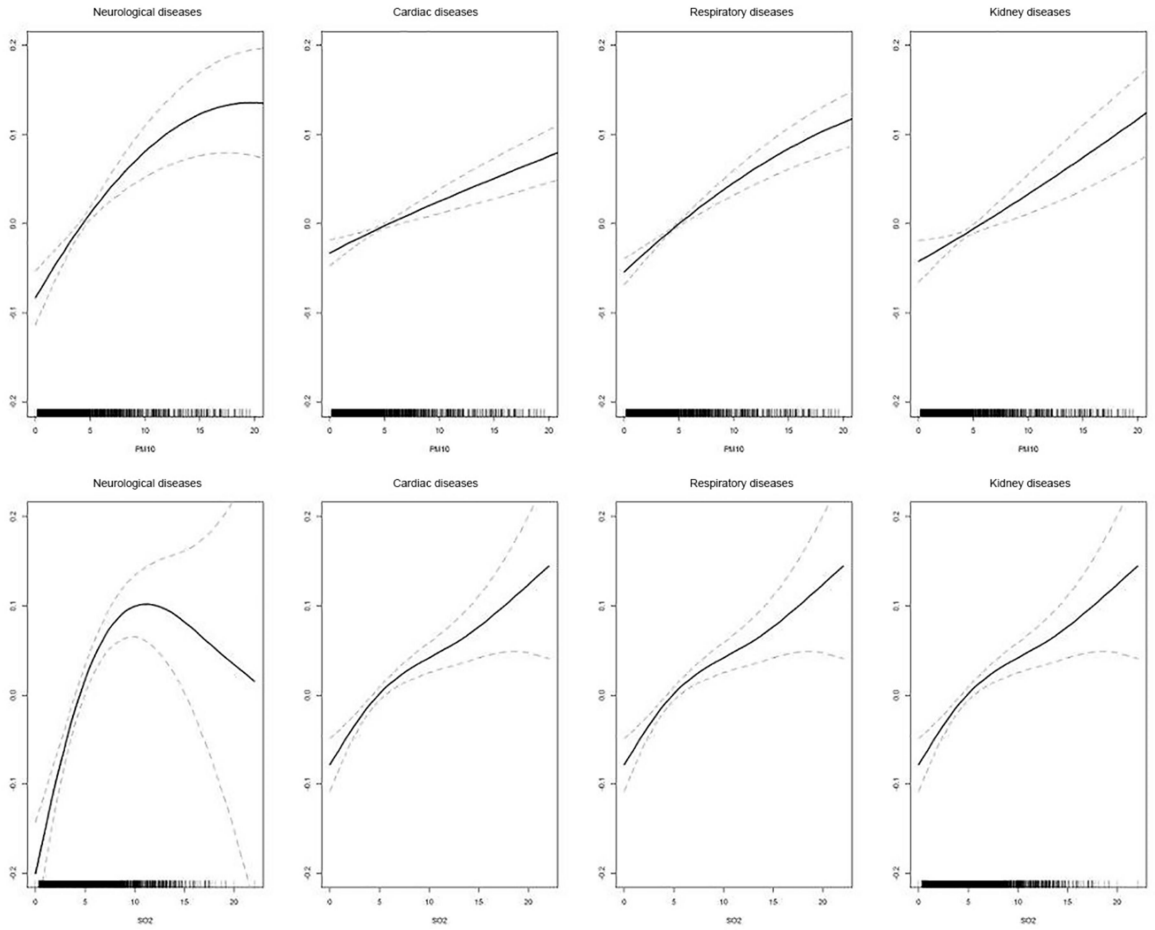

Figure 9: Penalized splines and confidence interval (95%CI) of the relationship between average PM_{10} and SO_2 exposure at lag 0, and hospital admissions for neurological disorders, heart, respiratory and kidney diseases.

Exposure to industrial pollution and effects on cancer incidence

Table 5 shows the association between pollutants from the industry and cancer incidence. An association was found with incidence of lung cancer (HR=1.29, 95%CI 1.14-1.45, and HR=1.42, 95%CI 1.10-1.84, for 10 μg/m^3 increases in PM_{10} and SO_2, respectively), kidney cancer (HR=1.32, 95%CI 1.01-1.73 for PM_{10} e HR=2.44, 95%CI 1.38-4.34 for SO_2). To note that exposure to PM_{10} was associated also with breast cancer among women (HR=1.27).

Site of cancer (ICD-O3T, ICD-O3M)		PM_{10}		SO_2	
	N	HR*	95%CI	HR*	95% IC
All site (ICDO3T C00-C809)	8999	1.14	1.09-1.19	1.05	0.97-1.14
Upper respiratory and digestive tract (ICDO3T C00-C14)	144	0.80	0.52-1.23	0.67	0.34-1.31
Oesophagus(ICDO3T C15)	27	0.30	0.06-1.48	0.20	0.04-1.08
Stomach (ICDO3T C16)	284	0.99	0.77-1.28	0.69	0.43-1.11
Colon-rectum and anus (ICDO3T C18-C21)	887	1.11	0.96-1.28	1.00	0.77-1.31
Liver(ICDO3T C22)	340	1.10	0.89-1.37	0.75	0.48-1.15
Gallbladder and biliary tract (ICDO3T C23-C24)	117	1.14	0.80-1.64	0.88	0.41-1.85
Pancreas (ICDO3T C25)	208	1.19	0.90-158	1.19	0.68-2.08
Larynx (ICDO3T C32)	91	1.39	0.99-1.96	1.39	0.62-3.13
Lung incl. trachea and bronchus (ICDO3T C33-C34)	943	1.29	1.14-1.45	1.42	1.10-1.84
Pleural cancer (ICDO3T C384)	89	0.96	0.61-1.52	1.15	0.50-2.64
Bone and cartilage (ICDO3T C40-C41)	22	0.59	0.16-2.22	0.53	0.09-2.96
Malignant melanoma of the skin (ICDO3T C44)	1944	1.15	1.04-1.26	1.08	0.90-1.30
Peripheral nerves, connective and soft tissue (ICDO3T C49)	40	1.22	0.66-2.27	0.62	0.17-2.26
Breast (ICDO3T C50)	1137	1.27	1.13-1.41	1.19	0.94-1.51
Prostate (ICDO3T C61)	653	1.09	0.92-1.29	1.06	0.77-1.45
Testis (ICDO3T C62)	42	1.08	0.58-2.01	0.96	0.30-3.11
Kidney (ICDO3T C64)	173	1.32	1.01-1.73	2.44	1.38-4.34
Renal pelvis and urinary organs (ICDO3T C65-C66, C68)	34	0.87	0.34-2.23	0.56	0.13-2.46
Bladder (ICDO3T C67)	415	1.07	0.88-1.32	0.91	0.61-1.35
Brain and central nervous system (ICDO3T C69-C72)	117	1.23	0.87-1.72	0.87	0.42-1.82
Thyroid (ICDO3T C73-C75)	365	0.97	0.75-1.25	0.76	0.49-1.17
Mesothelioma (ICDO3M 9050-9055)	72	0.96	0.57-1.60	0.93	0.36-2.37
Sarcoma Kaposi (ICDO3M 9140)	38	1.35	0.77-2.37	1.39	0.41-4.64
Hodgkin limphoma (ICDO3M 9650-9667)	52	0.98	0.54-1.78	1.56	0.54-4.50
Non-Hodgkin lymphoma (ICDO3M 9590-9596)	31	0.93	0.41-2.11	0.74	0.18-3.06
Multiple Myeloma (ICDO3M 9732)	98	0.91	0.56-1.46	0.76	0.34-1.69
Leukaemias (ICDO3T 9421, ICDO3M 9800-9948)	184	1.11	0.82-1.51	1.21	0.68-2.15

[a]Hazard Ratio (HR) from a Cox model stratified for period of follow-up (2 categories) and sex, adjusted for age (temporal axis), socioeconomic position and occupational status

Table 5: Associations between average exposures to PM_{10} and SO_2 from the industry and cancer incidence. Hazard ratio (HR) per increase of 10 µg/m^3 in average exposure and 95% Confidence Intervals, 2006-2011.

Residual Confounding

Table 6 shows the prevalence of the confounders analysed (smoking, alcohol, and obesity) by quartiles of the distribution of PM_{10} and SO_2 exposure in the ancillary data set used. There is no clear increasing trend of the prevalence of smoking, drinking alcohol and obesity with higher quartiles of exposure. No statistically significant association was found between exposures and individual smoking habits and obesity suggesting that they were not relevant confounders. Alcohol consumption was related to the 2^{nd} and 4^{th} quartile of PM_{10} only.

	PM_{10}				SO_2			
Risk factors	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**
Smoking								
Never smokers (%)	46.75	51.35	50.3	40.88	46.05	50.99	51.92	40.69
Smokers (%)	31.82	27.03	30.3	37.23	32.24	27.15	30.13	36.55
Ex-smokers (%)	21.43	21.62	19.39	21.9	21.71	21.85	17.95	22.76
Alcohol consumption (%)	52.6	66.9	52.12	62.77	54.61	62.91	55.77	60
Obesity (BMI>=30kg/m^2) (%)	16.88	12.16	10.3	16.06	15.79	13.25	11.54	14.48
Adjusted models[a] -RR(95%CI)								
Smokers [b]	1.00	0.87 (0.62-1.22)	0.97 (0.72-1.32)	1.24 (0.92-1.68)	1.00	0.88 (0.63-1.24)	0.94 (0.69-1.27)	1.24 (0.92-1.66)
Ex-smokers [b]	1.00	0.97 (0.64-1.46)	0.85 (0.57-1.27)	0.98 (0.65-1.47)	1.00	0.90 (0.60-1.37)	0.74 (0.49-1.12)	0.95 (0.64-1.41)
Alcohol consumption [c]	1.00	1.32 (1.09-1.59)	1.02 (0.82-1.26)	1.25 (1.02-1.52)	1.00	1.20 (0.99-1.46)	1.06 (0.86-1.30)	1.15 (0.95-1.40)
Obesity [d]	1.00	0.66 (0.37-1.17)	0.59 (0.33-1.04)	0.87 (0.51-1.50)	1.00	0.76 (0.43-1.34)	0.69 (0.39-1.23)	0.82 (0.47-1.45)

[a] Separated Poisson models adjusted for age, gender and socioeconomic position index.
[b] Reference category "non-smokers"; [c] reference category "No alcohol consumption"; [d] reference category "Not obese (BMI<30kg/m^2)"
Abbreviations: Qx= quartile; BMI=Body Mass Index; RR= Relative Risks; CI=Confidence Interval

Table 6: Association between smoking, alcohol consumption, obesity and quartiles of PM_{10} and SO_2 exposures.

4.4.Discussion

Results of this study showed higher risks of mortality for all causes, lung cancer, heart disease, acute myocardial infarction and kidney disease for people exposed to PM_{10} and SO_2 from the ILVA industry.

The study also showed higher risks in the first years of exposure as well as the distant past of more than 25 years back.. Pollutants were more strongly associated to hospital admissions than to mortality. Exposure to PM_{10} was associated with an increase in risk of natural mortality of 4%, while SO_2 with an increase of 9% for all natural diseases. The analysis of cancer incidence highlighted a positive relationship between the two pollutants and cancer of the lung, of the breast for women and of the kidney for both genders.

In the previous studies of Mataloni et al. (2012) and Vigotti et al. (2007) in the same area, highest risks were observed in the neighbourhoods closest to the industrial area (Paolo VI and Tamburi) where air pollution levels were greater than in the other districts. The risks remained high even when they corrected for the socioeconomic position (Mataloni et al., 2012). The relationships found in the present study are coherent with previous studies having the neighbourhood as unit of the study, whereas we conducted the investigation on an individual level. The SO_2 alone showed risks higher than PM_{10} and similar to those observed in Paolo VI district in the study of Mataloni et al. In particular, a recent study of Mangia et al.(2013) revealed that SO_2 identified very well the most polluted area close to the steel factory, exhibiting higher mean values and positive correlations with wind speed, when the monitoring station is downwind from the industrial site. The monitoring station located in Paolo VI district recorded the highest SO_2 mean concentration values, compared to the other neighbourhoods.

One important limitation of most longitudinal studies (Beelen et al., 2008; Bentayeb et al., 2015; Giulia Cesaroni et al., 2013; Jerrett et al., 2009; Mataloni et al., 2012) is taking individual levels of exposure contrasts at study inception (baseline) as representative for long term exposures. This approach is prone to exposure misclassification in case of widespread mobility patterns or changing spatial distribution of exposure over time. Few longitudinal studies have dealt with time-varying exposure assessment (Bentayeb et al., 2015; Lepeule et al., 2012; Tétreault et al., 2016; Wahida et al., 2016) for every year of follow-up, and among them only two studies attempted backward reconstruction of individual residential histories (Bentayeb et al., 2015; Wahida et al., 2016). The study conducted in Paris took into account residential mobility by weighting the individual cumulative exposure at census block level by the probability of moving from a census block to another, on a specific year, tabled in a matrix of all possible movements (Wahida et al., 2016). However, in the two studies, aggregation of exposure assessment was at the zip code and census block level, respectively. In both studies (Bentayeb et al., 2015; Wahida et al., 2016), an average cumulative exposure measure was estimated, which in one case (Bentayeb et al., 2015) lead to similar or weaker results compared to the mean annual exposure.

One of the strengths of this study was the availability of complete residential history as well high quality mortality and hospital admissions data, which for the latter arrived until the end of follow up. Record linkage procedures attributed almost 98% of the causes of death to individuals, and rigorous protocols were adopted to select appropriate hospital admissions.

We used a method that is quite novel to long-term effect studies; that is, taking the entire residential history from the time people became residents in the area, including the time preceding follow-up, considering all address changes and emigration/immigration in order to account for past exposure. The level of exposure was assessed on a very fine scale, through a well-reasoned procedure, where dispersion estimates were weighted with production and emissions data of the steel plant. In Table 1, we saw that more than 34% of subjects were residing in the area for more than 30 years, showing that the population is stable, with a percentage of 81-85% of cohort subjects who never changed address during follow-up.

However, individual exposures came from models and were not measured directly through the follow-up. This assumes that subjects would remain at home all day. Since a large proportion of subjects works and spends time in the workplace, this may have introduced a misclassification bias in our study.

In the analyses conducted, we confirmed a linear relationship linking PM_{10} time-dependent concentration and all causes mortality and morbidity, while SO_2 showed some parabolic associations. The analysis of the latency on mortality showed that risks of natural mortality were strongest in the first year of exposure and after more than 25 years of permanence in the area, while Bentayeb et al. (2015) showed greatest effects were found after one year of exposure. However the high correlation (r=0.7) amongst lags cannot allow us to think in terms of independent effects.

A more accurate lag- distributed model should be applied, like Lepeule et al. (2012) did in six Harvard cities from 1974 to 2009, to inspect the behaviour and trend of time of exposure on the onset of diseases. It is in fact also of interest to explore whether the concentration-response curve has changed over time, as particle composition and anthropogenic activities have changed over the 48-year period considered in our study.

Gasparrini's (2014) method for modelling exposure–lag–response associations with distributed lag non-linear models in longitudinal studies was implemented at the first instance, but with difficulties of application to this cohort. In fact, his method, working with matrices of data, is built on complete cases, or in situations in which missing data could be replaced by zero without changing the variable's meaning. In this analysis, there was much missing data, as very few observations had effectively all the data for the 35-year time period investigated. It was not feasible in this scenario to consider a missing exposure value as zero exposure. After evaluating the different characteristics in the age structure of this sample of residents, present in the area for 35 years, compared to the whole cohort, it was decided to estimate independent models for each 5-year time window, as described above.

A limitation of the study was the unavailability of important individual risk factors, such as smoking, physical activity and diet. Estimates on a sub-group of people for whom information on smoking, alcohol consumption and obesity was available indicated a negligible association to PM_{10} and SO_2 exposure. Prevalence of smoking in the sample of the cohort members was 31.5%, higher than the prevalence of 27-28% in the Apulia region and 28% in Italy globally; however, the higher prevalence was not related to air pollution levels from the industry. It is therefore extremely unlikely that smoking plays a confounding role in this cohort. In addition, for the other individual risk factors, it should be noted that we considered area-based socioeconomic indicators, thus making confounding less likely.

CHAPTER 5. Did Design Applied to the Cohort of Residents in the Taranto Area, Apulia Region, Italy

5.1. Introduction

In this part of the study, a "differences in differences" approach was used to investigate the relationship between exposure to air pollution from a specific industrial source and mortality in the nearby population. The method has been applied in the past in the econometric literature, and it can be seen as a "before/after study with a control group". The essence of the design is that differences in exposures across time are related to differences in rates of diseases in the same populations. The roles of potential individual and behavioral factors are thus cancelled out, as the comparisons are occurring within populations. Of course, variability of exposure across time is essential to appreciate differences in disease occurrence. In environmental epidemiology a recent study in the USA (Wang et al., 2016) used a variant of this method to evaluate the causal effects of long-term $PM_{2.5}$ exposure on mortality in New Jersey. The authors assumed the potential outcome (the aggregated number of death) in a census tract as a function of spatial confounders that vary among census tracts, but not over time, as well as temporal confounders that vary over time but not spatially, confounders that vary over time and space, and an offset term for the logarithm of the population in the census tract. Thus, by estimating the differences between years, they removed the confounding by variables varying by census tract but not time, and by estimating the differences between census tracts, they removed the confounding by variables that vary over time but not census tracts. They built the DID by estimating a Poisson regression model for the logarithm of the aggregated number of death by year and census tract in which they controlled for indicator variables for tracts and years.

The present study presents a variant of the DID design proposed by Wang et al. (2016) with an application to the cohort of residents in the Taranto area, south of Italy. The description of the cohort building and its characteristics is provided in Chapter 4.

The study has been described in a paper for publication, annexed to this monography in Annex B. Here I will provide a short summary.

5.2. Summary of the Study

5.2.1. Methods

The study cohort included all subjects residing in the study area at 01/01/1998, plus those entering the three municipalities until 31/12/2010 (see Chapter 4 for

more details on the cohort building). The cohort was followed up until December 2014. For the aims of this part of the study, only the latest period, 2008-2014, was retained, as it matched the availability of environmental data and was short enough to reduce potential residual confounding from temporal covariates.

Exposure Assessment

The exposure assessment constituted one of the most complex parts of the study. In fact, the objective was to estimate individual exposure to PM_{10} from industrial origin for each year of the study, combining daily data from four monitoring stations of the ARPA network for daily PM_{10} and NO_2 concentrations, with industrial concentrations coming from the Lagrangian dispersion model developed by ARPA for the year 2010, described in Chapter 4.
One station was located in an urban area of Taranto ("Via Alto Adige"); another one was in a suburban setting away from the industrial area ("Via Ugo Foscolo" in the Talsano district), in order to measure background concentrations. The remaining two stations were located near the industrial area: "Via Machiavelli" and "Via Archimede", in the Tamburi district at the border of the industrial area. These data were used to describe temporal variability in exposure over the study area.
Annual measurements from the monitoring stations and the data of the 2010 dispersion model map were combined in order to estimate industrial PM_{10} exposures for each year of the 2008-2014 period according to a complex mathematical methodology described in the Supplemental material of the paper.

Mortality Data

Causes of death were ascertained using record linkage procedures with the mortality registry of the Local Health Authority in Taranto and the following causes of death were analyzed: natural causes (ICD-9CM 001-799, ICD-10 A00-R99), circulatory system diseases (ICD-9CM 390-459, ICD-10 I00-I99), heart diseases (ICD-9CM 390-429, ICD-10 I00-I51), and respiratory diseases (ICD-9CM 460-519, ICD-10 J00-J99).

Statistical Analyses

For each year (7: 2008-2014), area-unit (11: 9 Taranto districts + Massafra and Statte) and age class (4: 0-34, 35-64, 65-74, >75 years) aggregated counts of cause-specific deaths were computed. Furthermore, using cohort information and individual residential history, the person-years were determined and used as denominators (or "offset") in the multivariate Poisson regression analysis described below:

$$\ln\left[E\left(Y_{q,t,e}\right)\right] = \beta_0 + \beta_1 I_q + \beta_2 T + \beta_3 I_e + \beta_4 I_q * T + \beta_5 I_e * T + \beta_6 PM_{10\,q,t,e} + \ln\left(P_{q,t,e}\right)$$

This final model, used and described in detail in the paper, can be interpreted as a variant of the difference-in-differences method; it is similar to the method proposed by Wang et al.(2016). The idea underlying the model is that a causal effect of PM_{10} is obtained by removing confounding from spatio-temporal covariates by design. This is achieved with the introduction of age-specific and district-specific linear trends in the regression model. The drawback of the model, on the other hand, is that only fluctuations of PM_{10} around linear trends are contrasted to concurrent fluctuations in mortality rates, with consequent lack of statistical power.

Results were presented as percent increase risk of death, and 95% confidence intervals, relative to 1μg/m^3 variation of industrial PM_{10}.

Effect modification by age in two classes (<65 years, 65+ years old) was also tested, and the p-value of the relative effect modification (p-REM) was evaluated (Biggeri et al., 2004).

The follow-up and the statistical analysis were performed using SAS 9.0, Stata 13, and geographical data were analyzed using Arcgis and Qgis.

5.2.2. Results

First results on the productivity (kton/year) of the ILVA plant in the period under study by type (steel, coke, sinter and pig iron) and the annual average PM_{10} (measured overall concentrations) assessed in the four monitoring stations, showed a non- negligible agreement between those measures. (Figure 3 and 4 of the paper)

During the study period of 2008-2014 37,736 deaths (11.7%) were observed from a cohort counting 262,375 people in the year 2008. We calculated 1,726,353 person-years of follow-up, differently distributed across the area districts (Table 1 in Annex B), and 15,303 natural deaths, 71.8% of which constituted subjects 75+ years old, and 0.8% from those below 35 years of age.

For descriptive purposes, the study area was divided into three sub-areas only: "Tamburi, Lido azzurro", "Isola, Borgo" and "Other" (which includes all the districts and municipalities different from the previous two). For each of the three area units the absolute change in the estimated industrial PM_{10} between the yearly value and the mean area was calculated. Thus, in Figure 5 and 6 of Annex B it is possible to observe that the annual trends of the industrial PM_{10} resemble the trends in mortality rates (per 1,000 person-years) in the "Tamburi, Lido azzurro" and "Isola, Borgo" districts, i.e. the districts most influenced by industrial emissions, whereas the patterns of exposure and mortality in the other areas deviate.

The results from the specified DID model (Table 2 in Annex B) showed a 1.86% (95% Confidence Interval (CI): -0.06, 3.83%) increase of natural mortality of relative to 1 μg/m^3 variation of industrial PM_{10}. In particular, we found a 2.37%

(95% CI: 0.31, 4.47%) increase in natural mortality (REM p-value= 0.22) among subjects 65+ years old (Figure 5 in Annex B). Furthermore, we found a 8.74% (95% CI: 1.50, 16.51%) increase in respiratory mortality. For the same outcome, there was no evidence of effect modification by age (REM p-value= 0.96). Effect estimates for circulatory and cardiac diseases were positive, but affected by larger standard errors.

5.2.3. Conclusions

The study supports the plausibility of a causal effect of industrial PM_{10} on mortality in the study area, especially in the elderly population.

This study has, to our judgement, several strengths: careful reconstruction of the individual residential histories, good quality geocoding, high completeness in the assignment of causes of death. However, it should be noted that the cohort lacks data on individual risk factors (such as smoking and other lifestyle characteristics, individual estimates of exposure to meteorological parameters, etc.). However, such factors might have confounded the estimates under investigation only under the assumption that they varied differently across age groups and districts, and that such differences were not adequately captured by linear trends. We believe it to be unlikely. Further support to this is provided by a recent re-analysis of the cohort study in Chapter 4 and Annex A ("*Studio di coorte sugli effetti delle esposizioni ambientali ed occupazionali sulla morbosità e mortalità della popolazione residente a Taranto*", (2016)) which applied indirect adjustment methods on data collected from the PASSI Surveillance System. That study showed that the associations between industrial pollutants and mortality/morbidity were not biased by unmeasured individual confounders such as smoking and body-mass index.

CHAPTER 6. Generalized Propensity Score in the Continuous Treatment Setting: an Application to the Cohort of Residents in the Taranto Area, Apulia Region, Italy

6.1. Introduction

In this chapter I will present the part of the study concerning the application of the main generalized propensity score methods (Hirano & Imbens, 2004; Imai & Van Dyk, 2004; Robins & Finkelstein, 2000) on the cohort of residents in the Taranto area.

The aim was to calculate through different methods the effect of exposure to industrial PM_{10} on mortality amongst residents living around the large steel plant located in the Taranto area, starting from what has been estimated in the main cohort study using the classical Cox regression model described in Chapter 4.

Recently, Zigler and Dominici (2012) discussed the priority role of more rigorous studies providing evidence of causality in the body of evidence for air pollution regulation. They provided good practice for causality in air pollution studies not strictly related to causal inference methods, but most importantly to the design of the study and its ability to render observational data as coming from a randomized experiment. A set of questions was then elaborated to check whether a study on air pollution provides evidence of causality.
In a second work (C. Zigler et al., 2016), the authors provided new statistical methods (mediation analysis and principal stratification) and perspectives for drawing causal inference on the long term effects air quality regulations. They encouraged the deployment of potential-outcomes methods for direct-accountability assessment, sometimes framed as intervention studies that analyse a large change in air pollution, and in general for air pollution epidemiology.

Many studies have been published on the use of causal methods based on propensity score methodology in observational studies and time-to-event analysis. However, there is few evidence on the use of the propensity score approach in the continuous treatment setting (Comfort et al., 2017; Jiang & Foster, 2013; Moodie & Stephens, 2012), and even less has been published in the time-to-event analysis framework (Koch et al., 2010; Podolanczuk et al., 2017).

In environmental epidemiology few studies implemented propensity score methodologies, prevalently PS matching, to obtain causal estimates of air pollution on health (Baccini et al., 2017; Capuno et al., 2018; Kirby et al., 2016; Mueller et al., 2011; Neupane et al., 2015; Rosa et al., 2017; Sætterstrom et al.,

2012; Wylie et al., 2014; Wylie, Coull, et al., 2015; Wylie, Singh, et al., 2015). Among them only three studies so far directly referred to air pollution as for particulate matter (Baccini et al., 2017; Rosa et al., 2017; Schwartz et al., 2015).
Only two studies focused on the generalized propensity score approach to obtain causal estimates of the effect of air pollution on health (Baccini et al., 2017; Schwartz et al., 2015).
The first one by Baccini et al. (2017) aimed at assessing the short-term impact (two days) of high daily levels of PM_{10} on mortality in Milan, Italy. The authors applied the propensity score approach for the prediction of the dichotomized variable for PM_{10}, defined as the daily concentration above and under the threshold $40\mu g/m^3$ (annual limit). Days were then matched using the estimated propensity score, and the impact was obtained by comparing mortality between matched days.
The authors did not evaluate a concentration-response curve; instead they focused on a binary exposure measure with the aim of investigating a "safe threshold" below which the pollutant doesn't affect health, with public health implications.

Generalized propensity scores for continuous exposure to particles with diameter less than 2.5μm have instead been estimated by Schwartz et al. (2015), in order to obtain the average causal association with daily mortality in a time-series study in Boston. They used a method similar to the one proposed by Imai and van Dyk (2004), in which they first modelled the $PM_{2.5}$ in a linear regression as a flexible function of time, temperature, trend and co-pollutants. They then took the predicted values of this model to get the propensity score. By first trimming on the upper and lower 5%, they then subclassified in deciles and used the resulting dummy variables in a quasi-Poisson model together with the two-day mean $PM_{2.5.}$

A topic widely explored by researchers regarding air pollution is the shape of the concentration-response. This issue is critical for public health assessment and it may have consequences if such thresholds exist.
Many studies have examined the shape of the concentration–response curve for long-term as well as short-term exposure (Crouse et al., 2012; M. J. Daniels, Dominici, Samet, & Zeger, 2000; Schwartz et al., 2001; Schwartz, Coull, Laden, & Ryan, 2008; Schwartz & Zanobetti, 2000; Shi et al., 2016). It was demonstrated that the association between particulate matter and mortality exists at any concentration, even below the maximum air quality thresholds currently in place in the USA and the EU.
Here, the Hirano and Imbens (2004) approach for the calculation of the DRF that computes the observed causal outcome for each level of the pollutant will be presented.

The current study aimed at advancing the knowledge of the application of generalized propensity score methods with continuous treatment (or exposure), in the domain of the air pollution observational studies, where it is more frequent to find residual confounding of the association, due to individual factors related to both the outcome and the exposure and not directly measured at the individual level.

Furthermore, in longitudinal studies on the association of air pollution on health, the main measure of individual confounders and risk factors is usually the socioeconomic index. This index comes from several census block characteristics that are synthetized through a principal component factor analysis to obtain one single score.
In this study, where the main exposure is the PM_{10} concentration at the subject's residence address at baseline, it was proposed to use a more complete list of individual confounders and risk factors as predictors not of the outcome, but of the exposure, i.e. the propensity score.
We can therefore identify these major objectives:

1. The first one was to better explain the socioeconomic inequalities in the population living around the steel plant in the Taranto area, compared to the index used in the main study, through a complete list of determinants available at the census block level and collected every 10 years by the National Statistics Institute survey;
2. To include as possible relevant variables, health indicators at district level that may serve as surrogate of individual behaviours like smoking habit;
3. To predict the generalized propensity score of the continuous exposure to PM_{10} and use it to obtain the average treatment effect over all the values and to estimate an exposure-response curve.

6.2.Methods

6.2.1. Mortality Data

The data used for this study were those of the main cohort described in Chapter 4, composed by residents enrolled in the area of Taranto, Massafra and Statte, at January 1st 1998 and all subjects later entering for immigration or birth until December 31st 2010, followed up until 2014.

For each resident, health data on mortality, with corresponding dates and causes of death (ICD-9CM revision), have been assigned through record-linkage procedures with the Regional Health Databases.

6.2.2. Air Pollution Data

PM_{10} data concentrations came from a Lagrangian modelling system developed by the Regional Environmental Protection Agency (ARPA) of the Apulia region for the year 2010. Differently from the main cohort study described in Chapter 4, the level of exposure has been attributed to each subject at its first residence address, through a geocoding procedure(the distribution over the area is shown in Figure 2 in Chapter 4).

6.2.3. Socioeconomic Position Data

A census block-level socioeconomic position (SEP) index was assigned originally to each participant on the basis of his/her geocoded address and was composed by

information recorded at the National Statistics Institute 2001 census: percentage of population with educational level equal to or less than primary school, percentage of the active population unemployed or seeking their first occupation, percentage of rented houses, percentage of single parent families, population density (number of occupants per 100 m^2). A composite index was computed through factor analysis and classified into 5 quintiles, representing high, middle-high, medium, middle-low and low SEP. (Caranci et al., 2010)

In order to obtain a better description of the population and a good prediction of the propensity score it was decided to consider a more exhaustive list of variables retrieved from the 2001 census, which is provided in Annex C.
Not all the census variables were selected, but only those identified as importantly related to both the outcome and the exposure.
Summary informative quantities from those variables, such as the proportion over the population or the ratio, were calculated, standardized with their province mean, and attributed to each subject according to his census block at baseline:

- Percentage of male and female residents.
- Percentage of married, separated or divorced residents.
- Percentage of single-person households, percentage of five-person or more households.
- Percentage of graduates, percentage of diploma or no-diploma residents.
- Percentage of foreign residents.
- Percentage of workers in agriculture, industry, commerce or public administration, service.
- Employment Rate: defined as the ratio of employed 15 years old and over people and the total of people aged 15 years and over who are in the labor force, expressed per 1000.
- Unemployment Rate defined as the ratio of unemployed 15 years old and over people and the total of people aged 15 years and over who are in the labor force, expressed per 1000.
- Percentage of entrepreneur, self-employed residents, percentage of employed residents, percentage of not in the labor force residents, percentage of housewives.
- Percentage of commuters: defined as the proportion of residents moving daily outside the town to work.
- Percentage of private houses, percentage of houses with potable water, percentage of houses with heating, percentage of habitable buildings.
- Population density (number of occupants per 100 m^2).

6.2.4. Statistical Methods

Each resident contributed to person-years at risk from the date of entry in the cohort (1998 or later within 2010) until date of exit for death, emigration or end of follow-up, whichever came first.
The association between long-term exposure to air pollutants (defined as time-varying annual average) and mortality was first estimated using a standard survival analysis with multivariate Cox proportional hazard models, as described in Chapter 4. In that model, age was used as the time scale. *A priori* confounders

included: gender, area-level SEP, occupation (recorded at baseline and defined as blue collar/white collars workers in steel factories; naval construction workers; mechanical construction workers; workers in other occupational branches or people without contribution payments (reference category)); calendar period, in 3 classes (defined as time-dependent).

The generalization of the propensity score was applied according to two of the main methods: the Propensity Function (PF) of Imai van Dyk (2004) and the Generalized Propensity Score (GPS) at each exposure level by Hirano and Imbens (2004). Extensions to HI method have been also considered (Bia et al., 2014).
The first method (Imai & Van Dyk, 2004) allows estimation of the average treatment effect (ATE), while the second one estimates the concentration-response curve.
A third method by Robins (2000) on the Inverse Probability of Treatment Weighting (IPTW) has also been applied following the guidelines of Austin (2011) for propensity score in time-to-event observational studies.

In the PF and GPS methods the generalized propensity score has been estimated through a linear regression of the exposure to PM_{10} against census indicators list, age, gender and occupation. The last three regressors were already present in the standard Cox model used for the main analysis and are now predictors of the propensity score. The SEP index has been instead replaced by the extensive list of census variables to predict the exposure.

The propensity score analysis was implemented in Cox Proportional Hazard models stratifying for calendar period in 3 classes and gender, with age as the temporal axis.

Mortality from natural causes (ICD-9CM: 001-799), cardiovascular (ICD-9CM: 390-459), cardiac (ICD-9CM: 390-429) and respiratory (ICD-9CM: 460-519) causes was studied.

A complete cases analysis was performed, since the information on socioeconomic factors from the 2001 census was missing for some census tracts.
This choice is questionable, since it can lead to biased causal inference, unless the data are missing at random (D'Agostino Jr & Rubin, 2000). In our case the size of data did not allow multiple imputation techniques, since they require huge computational time and resources.
The common support condition was verified through the method defined by Bia et al. (2014). Observations out of the common support were discarded.

Imai van Dyk Propensity Function

Following the proposal by Imai and van Dyk (2004) the PF was calculated as the linear prediction of the regression model of PM_{10} on the selected variables and then subclassified. The PF was first subclassified in 5 groups according to the

quintiles of its distribution (Austin, 2011), and then also in 10 groups defined by the deciles, following the suggestion of Imai van Dyk to increase the number of classes to obtain a reduction in bias. Hazard ratios (HR) for each stratum were computed and averaged to obtain the ATE, weighting for the number of observations in each class.
Another analysis on the response model was performed to explore how estimate would change following the suggestion of Imai van Dyk (2004), which is to adjust in each stratum at least for the PF score.
Covariate balance of the PF was evaluated by regressing the covariates as dependent variables on the PM_{10} exposure, before and after adjusting for the PF.
Hazard Ratios (HR) and corresponding 95% Confidence Intervals (95% CI) per 10 $\mu g/m^3$ increases of the pollutant were computed.

Hirano and Imbens Generalized Propensity Score and DRF

Hirano and Imbens (2004) concentration-response curve has been estimated through the R package **causaldrf** (Galagate, Schafer, & Galagate, 2015) that allows the specification of different models for the relationship between the outcome and the exposure and the GPS. As a first step, the GPS is calculated as the conditional density of the exposure to PM_{10} given the covariates, for each level of exposure.
Then three estimates of the DRF were derived for fixed values of exposure:

1) The HI model, with a quadratic parametrization of the relationship between outcome and exposure and the GPS, as well as an interaction term (2.3)

$$E[Y_i|T_i, R_i] = \alpha_0 + \alpha_1 T_i + \alpha_2 {T_i}^2 + \alpha_3 R_i + \alpha_4 {R_i}^2 + \alpha_5 T_i R_i;$$

2) A more flexible function for the GPS in a generalized additive model (GAM), through the package **mgcv** (Wood & Wood, 2007) inside **causaldrf**, that regress Y on the exposure and cubic regression spline terms for the GPS fit;
3) Flores et al. (2012) extension of HI method, that is an additive spline estimator, where additive spline bases values are created for both the exposure and the estimated GPS. Then the outcome is regressed on the exposure, the GPS, exposure basis, and GPS bases. Three knots were chosen for both splines.

These methods for calculating the DRF were presented in R and Stata packages only for outcomes Y assumed to be generated from a particular distribution in the exponential family, a large range of probability distributions that includes the normal, binomial, Poisson and gamma distributions, among others. Survival analysis instead involves the modelling of time-to-event data; in this context, death or failure is considered an "event" in the survival analysis literature, and the Cox proportional hazard model is usually adopted to estimate through partial-likelihood the multiplicative relationship between the covariates and the hazard ratio.

However, there is a relationship between proportional hazards models and Poisson regression models, which is sometimes used to fit approximate proportional hazards models in softwares for Poisson regression. (Berry, 1983; Laird & Olivier, 1981). McCullagh and Nelder's (1987) book on generalized linear models has a chapter on converting proportional hazards models to generalized linear models, where rates are modelled instead of survival time, based on counting processes and martingales.
The Andersen–Gill (1993) counting process method for analysis of hazard functions is very similar in this context to Poisson regression methods for the analysis of rates.

In practice in this study the cohort was split by year of residence in the area throughout the follow-up period. The event of interest (in this case, death) was put as the dependent variable of a Poisson GLM on the set of pre-defined regressors. The parameters were then estimated by using iteratively reweighted least squares with person-years as an "offset".
Before estimating the DRF the equivalence between the Cox model and the Poisson model was verified.

Covariate balance using the HI methods and extensions was verified in an earlier step, by using an automatic procedure developed by Bia and Mattei (2008) in the Stata package **drf** and **gpscore** (Bia et al., 2014), considering intervals of exposure and GPS based on quartiles and quintiles of the distribution, respectively. A correspondent command in R wasn't found; hence, for this part of the analysis it was necessary to export the database in the Stata software to perform this check.

Robins' Importance Sampling

The last part of the analysis implemented the importance of sampling procedure of Robins (2000), estimating Cox PH models weighted for the inverse of the weights defined by the author for the continuous case, under the assumption of normal distribution of the exposure, as $w_i(t) = g(T_i)/P(T_i = t|\boldsymbol{X}_i)$, where $g(T_i)$ is the marginal density for T_i under normal distribution, and the denominator is the conditional density of the continuous variable for exposure T_i given the set of covariates $\boldsymbol{X}_i$.
To verify the stability of weights and the presence of outliers affecting the average causal effect, a trimming on the 5% and 95% of their distribution was also performed, in addition to the Robins' approach in a further analysis.

Additional Analysis on Propensity Score Model Specification

Two additional analyses were performed to verify the robustness of the propensity score model specification.

The first analysis considered also the SEP index in the prediction model, since some of the information used to compute the index was different from the ones considered in the specification model.

Two additional factors predicting the exposure were included in the propensity model in a later analysis in order to verify its correct specification and improve the prediction. These factors were the Age-Standardized Hospitalization Rates (direct method per 100,000) for chronic obstructive pulmonary disease (COPD) and lung cancer, amongst males and females. The rates were computed for each neighbourhood in the period 1998-2010 and attributed to each individual at the residence neighbourhood level. These rates are surrogates of the smoking prevalence in the specific neighbourhood.

6.2.5. Results

The cohort from the complete cases analysis counted 309,069 subjects in total, of which 36,384 (11.8%) died, in the period from 1998 to 2014.
Among the deceased, 34,947 (96%) died of natural causes, 12,932 (35.5%) died of cardiovascular diseases, 9,203 (25.3%) of cardiac diseases, and 2,799 (7.7%) of respiratory diseases (results not shown).
Among the list of parameters collected during the 2001 census by the National Statistics Institute survey, only some, listed in Table 7 (Part 1) and Table 8 (Part 2), were considered as possible predictors of the propensity score in the regression model. Total number of people for each category, as well as their percentages or ratios, was reported and referred to the population of Taranto, Massafra and Statte municipalities resident in 2001, together with correspondent mean and standard error in the cohort under study. The cohort of residents under study showed the same mean values of the selected parameters.
These quantities, computed for each census tract and standardized by the province mean, were then attributed to each subject, according to the census tract of its first residence, in the hypothesis of a stability over time.

		N	%*	Mean % in the Cohort	SD in the Cohort
Demographics	**Population**	**247,541**	**100**	**-**	**-**
	Males	119,158	48.14	48.09	5.55
	Females	128,383	51.86		
	Married	121,811	49.21	48.39	6.86
	Separated or divorced	5,403	2.18	2.27	1.80
	Foreigners	823	0.33	0.37	0.82
Households	**Number of households**	**85,580**	**100**		
	Single-person households	14,905	10.35	10.94	8.97
	5-person or more houselholds	8,855	17.42	17.28	9.29
Education	**5-year old and over population**	**235,912**	**100**		
	Bachelor	17,177	7.28	7.11	7.92
	Diploma	59,870	22.49	24.27	11.99
	No diploma	28,425	12.04	12.71	6.70

* Percentages may not sum up to 100, because some categories are missing or subjects can belong to more than one category

Table 7: Distribution of indicators and main indexes of the ISTAT census in the municipalities of Taranto, Massafra and Statte, 2001 (Part 1).

		N	%*	Mean % in the Cohort	SD in the Cohort
Occupation	**Population**	**247,541**	**100**		
	Population 15 years old and over in the labour force	87,124	35.20	35.32	8.52
Domain	Workers in agriculture	3,965	4.55	4.54	7.18
	Workers in industry	18,299	21.00	20.88	8.34
	Workers in commerce	11,478	13.17	13.00	5.31
	Workers in public administration/service	22,758	26.12	25.26	13.14
Type of employment	Entrepreneur-self-employed	4,597	5.28	4.82	4.97
	Employed	55,173	63.33	62.78	11.76
	Population not in the labour force	121,492	49.08	49.08	9.09
	Housewives/Househusbands	50,709	20.49	20.79	6.31
	Employment Rate (x1000)		784.51	773.72	127.70
	Unemployment Rate (x1000)		119.01	123.25	74.23
	Commuters	12,173	4.92	4.45	5.25
Living conditions	**Number of dwellings**	**100,912**	**100**		
	Private houses	58,504	58	66	22
	Houses with potable water	98,112	97	98	9
	Houses with heating	74,128	86.71	86.07	15.90
	Number of buildings	**26,518**	**100**		
	Habitable buildings	22,563	85.09	88.12	19.01
	Population density (number of occupants per 100 m^2)		3.02	3.31	3.40

* Percentages may not sum up to 100, because some categories are missing or subjects can belong to more than one category

Table 8: Distribution of indicators and main indexes of the ISTAT census in the municipalities of Taranto, Massafra and Statte, 2001 (Part 2).

Upon preliminary correlation analyses, some of these variables were found to be highly correlated and consequently collinear.

The final propensity score multivariate model for the PM_{10} exposure gave coefficients and standard error as reported in Table 9.
Age, previously only considered in the temporal axis in the previous standard Cox model, (described in Chapter 4), was also used in the propensity score model.
The same goes for the gender which was only considered to stratify the risk at baseline. It was decided instead to only include occupational status in the PS prediction.
All the covariates were significantly associated with exposure, except the percentage of families with 5 or more members.
Once calculated the propensity score in both methods, the covariate balance was checked through the HI and the IvD methodologies. The HI performed a series of *t*-tests that the conditional mean of the pre-treatment variables given the generalized propensity score is not different between units who belong to a particular treatment interval and units who belong to all other treatment intervals for each covariate, and satisfied the balance property at a level lower than 0.01.
Also the IvD procedure obtained not significant p-values for the t-tests of each covariate regressed on the treatment after conditioning on PF.

	Coefficient	Standard Error	95% CI	
% males	0.387	0.011	0.36	0.41
% females	-0.048	0.011	-0.07	-0.03
% separated or divorced	-0.712	0.014	-0.74	-0.68
% foreigners	-1.252	0.025	-1.30	-1.20
% single-person households	0.367	0.011	0.35	0.39
% 5-person or more households	0.003	0.008	-0.01	0.02
% bachelor	1.312	0.010	1.29	1.33
% diploma	-0.970	0.009	-0.99	-0.95
% no diploma	0.051	0.009	0.03	0.07
Unemployment Rate	0.188	0.008	0.17	0.20
% workers in agriculture	-1.261	0.014	-1.29	-1.23
% workers in industry	-0.179	0.010	-0.20	-0.16
% workers in commerce	0.231	0.008	0.22	0.25
% workers in services	-0.732	0.013	-0.76	-0.71
% self-employed workers	0.358	0.013	0.33	0.38
% employed workers	0.701	0.011	0.68	0.72
% not in labor force	-0.646	0.009	-0.66	-0.63
% housewives/househusbands	2.202	0.010	2.18	2.22
% commuters	-1.052	0.007	-1.07	-1.04
% private houses	0.948	0.005	0.94	0.96
% houses with drinkable water	0.574	0.008	0.56	0.59
% houses with heating	-1.180	0.006	-1.19	-1.17
% habitable buildings	-0.056	0.006	-0.07	-0.04
Population density	0.588	0.005	0.58	0.60
Age	0.000	0.000	0.00	0.00
Gender	0.021	0.007	0.01	0.03
Employment in the Mechanical Construction	0.137	0.015	0.11	0.17
Employment in the Naval Construction Industry[a]	0.396	0.053	0.29	0.50
Employment at the Iron and Steel Industry-	0.059	0.020	0.02	0.10
Employment at the Iron and Steel Industry- Office	-0.399	0.030	-0.46	-0.34

[a]1974-1997

Table 9: Factors associated with PM_{10} in the propensity score prediction model, coefficients, standard error and Confidence Intervals (CI)

Estimates of the causal effect of exposure on mortality were first calculated for natural mortality, to check the validity of the model on the largest sample of data. Cox standard model, Cox model stratified with Imai van Dyk (IvD) Propensity function in five subclasses, and Cox model with IPTW with trimmed weights according to Robins (RISt) were estimated. (Table 10)

The Hazard Ratio of PM_{10} on natural mortality estimated by the standard Cox model was 1.04, with confidence interval (CI) of 1.01-1.06.

Estimates from standard Cox and IvD were the same in the size of the effect and its standard error (HR=1.04, CI 95% 1.01-1.06), while the RISt model obtained a

higher estimate of the effect (HR=1.09, CI 95% 1.06-1.11) for 10μg/m^3 increase of PM_{10} from industrial origin.

When the same Cox model with IvD stratification was applied stratifying the PF in 10 subclasses instead of 5 or adjusting within the strata for the PF score (Table 11), the effect estimate had approximately the same size (from HR=1.04 to HR=1.03) but with larger confidence intervals. An explanation is that in the case of stratification for more subclasses the model could in fact have estimated stratum effects among few observations, losing power. In the case of adjustment for the PF the model could have instead produced an overfitting within the strata.

Method	Model	HR	95%CI	
Standard	Cox model with SEP[a]	1.04	1.01	1.06
Imai and van Dyk (IvD)	Cox model stratified for PF[b]	1.04	1.01	1.06
Robins Importance Sampling Trimmed (RISt)	Cox model with trimmed IPTW[c]	1.09	1.06	1.11

[a]Cox model stratified by follow-up period (three classes) and gender, age on the temporal axis, socioeconomic position and occupational status.
[b]Cox model stratified by follow-up period (three classes) and gender, stratified for PF in 5 subclasses
[c]Cox model stratified by follow-up period (three classes) and gender, with weights based on GPS and trimmed at 5% and 95%
Abbreviations: HR=Hazard Ratio; CI=Confidence Intervals; IvD= Imai van Dyk; RISt= Robins Importance Sampling Trimmed; IPTW=Inverse Probability Treatment Weighting; PF= Imai van Dyk's Propensity Function; SEP= socioeconomic position

Table 10: Associations between annual average exposure to industrial PM_{10} and natural mortality (Number of deceased=34,947). Hazard ratio (HR) per 10 μg/m^3 increase in PM_{10} and 95% Confidence Intervals, CI, 1998-2014.

Method	Model	HR	95%CI	
IvD	Cox IvD stratified for PF[a]	1.04	1.01	1.06
	adjusted for PF[a]	1.03	1.00	1.06
	with subclassification in 10 classes[b]	1.03	1.00	1.06

[a]Cox model stratified by follow-up period (three classes) and gender, stratified for PF in 5 subclasses
[b]Cox model stratified by follow-up period (three classes) and gender, stratified for PF in 10 subclasses

Abbreviations: HR=Hazard Ratio; CI=Confidence Intervals; IvD= Imai van Dyk.

Table 11: Propensity Function model by Imai van Dyk for the ATE of exposure to industrial PM_{10} on natural mortality (Number of deceased=34,947) in different outcome response models. Hazard ratio (HR) per 10 μg/m^3 increase in PM_{10} and 95% Confidence Intervals, CI, 1998-2014.

To verify the robustness of the propensity score model it was decided to try different models for its estimation.
The causal models chosen, in which different specifications of the PS were tested, were those defined in Table 10:

- IvD: Cox model stratified by PF in 5 subclasses, follow-up period (three classes) and gender
- Robins: Cox model stratified by follow-up period (three classes) and gender, with weights defined by Robins et al. (2000), trimmed at 5% and 95%.

Results are shown in Table 12. First the socioeconomic position index (SEP) was included in the PS estimation model. The estimates of the causal effect in the IvD model was reduced and lost some power when adding the SEP in the propensity score, probably for an overestimation of the PS score (HR=1.03, CI 95% 1.00-1.06 with SEP vs HR=1.04, CI 95% 1.01-1.06 without the SEP). In the RISt model, the inclusion of SEP determined as well a slight reduction of the effect of PM_{10} (HR=1.08, CI 95% 1.06-1.11 with SEP vs HR=1.09, CI 95% 1.06-1.11 without SEP).

Then the age-standardized hospitalization rates for Chronic Obstructive Pulmonary Disease (COPD) and Lung cancer, estimated in the period 1998-2010 by gender and district (Table 13), were included in the propensity model as proxies for the smoking habits distribution, since this information was not collected in the census and in the cohort.
The two rates were added one at a time, then together and, as a final model, the two rates together with the SEP index. Again, the IvD and RISt models were estimated with the different propensity scores obtained.

The inclusion of the age-standardized rates for COPD alone did not change the final effects, whereas the inclusion of both COPD and lung cancer SHRs in the prediction model increased the causal effect of PM_{10} on natural mortality in both IvD and RISt models (HR=1.05, CI 95% 1.02-1.08 and HR=1.10, CI 95% 1.07-1.13 in IvD and RISt, respectively).
When including the two SHRs together with the SEP index in the PS and then in the IvD model, the result did not change compared to the PS specification with only the two SHRs (HR=1.05, 95%CI 1.02-1.08), while in the RISt the effect decreased (HR=1.09, 95%CI 1.06-1.12).

Method	Model	HR	95%CI	
Standard	Cox model with SEP[a]	1.04	1.01	1.06
IvD	Cox model stratified for PF[b]	1.04	1.01	1.06
	+SEP in the estimation of PF	1.03	1.00	1.06
	+SHR COPD in the estimation of PF	1.04	1.01	1.08
	+SHR Lung cancer in the estimation of PF	1.03	1.00	1.06
	+SHR_{COPD}+SHR_{LC} in the estimation of PF	1.05	1.02	1.08
	+SEP+SHRCOPD+SHRLC in the estimation of PF	1.05	1.02	1.08
RISt	Cox model with trimmed IPTW[c]	1.09	1.06	1.11
	+SEP in the weights prediction	1.08	1.06	1.11
	+SHR COPD in the weights prediction	1.09	1.06	1.12
	+SHR Lung cancer in the weights prediction	1.09	1.06	1.12
	+SHR_{COPD}+SHR_{LC} in the weights prediction	1.10	1.07	1.13
	+SEP+SHRCOPD+SHRLC in the weights prediction	1.09	1.06	1.12

[a]Cox model stratified by follow-up period (three classes) and gender, age on the temporal axis, socioeconomic position and occupational status.
[b]Cox model stratified by follow-up period (three classes) and gender, stratified for PF in 5 subclasses
[c]Cox model stratified by follow-up period (three classes) and gender, with weights based on GPS and trimmed at 5% and 95%
Abbreviations: HR=Hazard Ratio; CI=Confidence Intervals; IvD= Imai van Dyk; RISt= Robins Importance Sampling Trimmed; IPTW=Inverse Probability Treatment Weighting; PF= Imai van Dyk's Propensity Function; SEP= socioeconomic position; COPD=Chronic Obstructive Polmunary Diseases; LC=Lung Cancer

Table 12: Associations between annual average exposure to PM_{10} and natural mortality (Number of deceased=34,947) in different propensity specification models. Hazard ratio (HR) per 10 µg/m^3 increase in PM_{10} and 95% Confidence Intervals, CI, 1998-2014.

District	SHR COPD		SHR Lung Cancer	
	Male	Female	Male	Female
Borgo	214,3	103,4	87,2	9,7
Tamburi, Isola, Porta Napoli, Lido Azzurro	367,1	133,9	117,4	10,0
Italia Montegranaro	190,8	93,3	77,0	12,9
San Vito, Lama, Carelli	189,8	110,3	63,5	18,9
Paolo VI	374,4	156,5	131,0	10,3
Salinella	232,0	93,6	89,0	14,3
Solito Corvisea	215,3	102,3	80,3	6,2
Talsano	246,5	98,8	99,0	7,6
Tre Carrare, Battisti	241,1	109,5	89,9	10,1
Massafra	275,3	92,9	48,2	7,6
Statte	288,1	104,0	77,6	5,4

[a] Direct standarization (Italian population at 1991)
Abbreviations: SHR=Standardized Hospitalization Rates

Table 13: Chronic Obstructive Pulmonary Disease (COPD) and Lung Cancer Age-Standardized Hospitalization Rates[a] (SHR) per 100,000, by gender and district in the cohort of residents in Taranto, Massafra and Statte, 1998-2010.

The last analysis explored the effects of PM_{10} exposure on specific causes of death: cardiovascular, cardiac and respiratory causes through standard Cox, IvD and RISt models (Table 14). The Cox models for the three methods are defined in Table 10, from which the HR values for natural mortality were taken.
From Table 13 it was observed that for the standard Cox model, most of the effect on natural mortality was driven by the effect on cardiac mortality (HR=1.08, 95%CI 1.03-1.13). Cardiovascular and respiratory mortality had positive, but not statistically significant estimates (HR= 1.03, 95%CI 0.99-1.07 and HR=1.02, 95%CI 0.94-1.11 for cardiovascular and respiratory mortality, respectively).
The IvD model confirmed the results obtained for all causes from the standard Cox model, with a slight decrease in the effect for cardiac mortality (HR=1.07, 95%CI 1.02-1.12).
Causal estimates of PM_{10} on cause-specific mortality in the RISt model were higher than the ones observed with the other two methods, as was observed for natural mortality. In the case of cardiovascular mortality the effect also reached statistical significance (HR=1.09, 95%CI 1.04-1.13).

		Standard Cox[a]			IvD[b]			RISt[c]		
Cause of death (ICD-9CM)	**N**	**HR**	**95%CI**		**HR**	**95%CI**		**HR**	**95%CI**	
Natural mortality (001-799)	34,947	1.04	1.01	1.06	1.04	1.01	1.06	1.09	1.06	1.11
Cardiovascular mortality (390-459)	12,932	1.03	0.99	1.07	1.03	0.99	1.08	1.09	1.04	1.13
Cardiac mortality (390-429)	9,203	1.08	1.03	1.13	1.07	1.02	1.12	1.15	1.10	1.20
Respiratory mortality (460-519)	2,799	1.02	0.94	1.11	1.04	0.94	1.14	1.09	1.00	1.19

[a]Cox model stratified by follow-up period (three classes) and gender, age on the temporal axis, socioeconomic position and occupational status.
[b]Cox model stratified by follow-up period (three classes) and gender, stratified by PF in 5 subclasses
[c]Cox model stratified by follow-up period (three classes) and gender, with weights based on GPS and trimmed at 5% and 95%
Abbreviations: HR=Hazard Ratio; CI=Confidence Intervals; IvD= Imai van Dyk; RIS= Robins Importance Sampling; RIST= Robins Importance Sampling Trimmed

Table 14: Associations between annual average exposure to industrial PM_{10} and mortality for natural, cardiovascular, cardiac and respiratory mortality in standard Cox, IvD and RISt models. Hazard ratio (HR) per 10 µg/m^3 increase in PM_{10} and 95% Confidence Intervals, CI, 1998-2014.

Figure 10 shows the causal exposure-response curves estimated through Hirano and Imbens' GPS method for natural, cardiovascular, cardiac and respiratory mortality. Three estimation models were used:

1) Black line: a quadratic parametrization of the relationship between outcome and exposure and the GPS, as well as an interaction term
2) Red dashed curve: additive spline estimator (Flores et al. (2012) extension of HI method))
3) Blue dotted curve: flexible function for the GPS in a generalized additive model

The three curves in our data show approximately the same behaviour on the response outcome. With HI exposure-response function at increasing exposures to $PM_{10,}$ the estimated mortality rate increases rapidly up to 10μm and then slightly decreases for higher levels of PM_{10}. The same behaviour is observed with the additive splines that give more flexibility to the model. However, the GAM estimate increases rapidly for lower levels of PM_{10} and then continues growing for higher levels of the pollutant. The cause-specific curves are almost identical in the behaviour, but not in the magnitude of the rate for the cause-specific mortality.

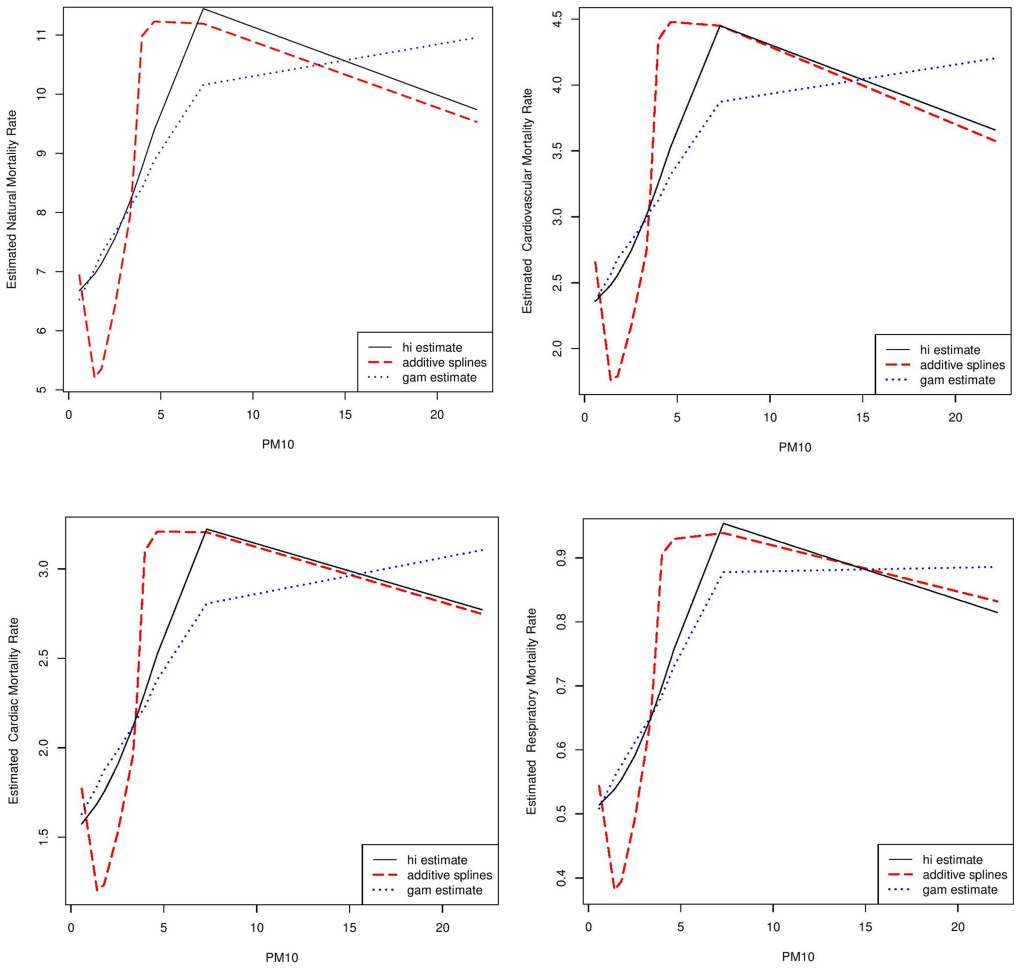

Figure 10: Exposure-response curves for the relationship between PM_{10} and mortality rates from all-natural, cardiovascular, cardiac, and respiratory causes, from Hirano and Imbens' method and its extensions, 1998-2014.

6.2.6. Discussion

This study revealed interesting results in the estimated effect of exposure to PM_{10} on mortality. In the standard Cox model it was observed a clear positive effect of PM_{10} on mortality from natural causes and cause-specific. The Imai van Dyk Propensity Function model, applied to the data to obtain causal marginal estimates, showed similar results, while the Inverse Probability of Treatment Weighting method by Robins (2000) gave higher estimates of the effect.
It was beyond the purpose of the study to establish which statistical method could lead to the least biased estimates. However, some literature exists discussing the pros and cons of the main propensity scores' methodologies (covariate adjustment, stratification, matching, and inverse probability of treatment weighting) in time-to-event studies that could help interpreting the results.

It's the case of some recent studies by Austin (Austin, 2013, 2014; Austin, Grootendorst, & Anderson, 2007) that have compared the performance of different propensity scores methods in survival analysis and found that both stratification on propensity score and covariate adjustment resulted in biased estimation of marginal hazard ratios.
Among the commonly used propensity scores methods, only IPTW based on propensity score has been considered allowing for the true estimation of marginal hazard ratios with negligible bias when estimating the ATE (Austin & Stuart, 2015c; Gayat, Resche-Rigon, Mary, & Porcher, 2012).
IPTW, based on PS, in a series of simulations by Austin Stuart (2015c), produced estimates of the marginal hazard ratio with negligible bias with both a weak and a strong treatment specification model.
In another study by Austin (2014), the use of PS methods for stratification and covariate adjustment in Cox proportional hazard models was discouraged, because these methods result in conditional estimates of the treatment effect, rather than marginal effect. Moreover, both stratification on the PS and covariate adjustment using the PS can result in biased estimations of the conditional hazard ratio that would be obtained by adjusting for all prognostically important covariates (Austin, 2013). Therefore Austin (Austin, 2009a, 2014) recommended that researchers use either IPTW using PS or PS matching in studies estimating the effect of treatment on survival outcomes that demonstrated to eliminate the systematic differences between treated and untreated subjects to approximately the same degree.
On the other hand, Rubin (2004) suggested that matching and stratification may be preferable to IPTW and covariate adjustment in general, since the latter two directly use the estimated propensity score and may thus be more adversely affected by misestimation or instability in the estimated propensity scores. Zhao et al. (2013) obtained a very unstable fit characterized by very large standard errors with IPTW for the estimation of the DRF.

Diagnostics on observational studies using PS are crucial in determining whether conditioning on the estimated PS has removed all the observed systematic differences between treated and control subjects.

Consequently, another point of discussion is that while balance diagnostics have been described for covariate adjustment using the propensity score, stratification and matching using PS, these diagnostics are less transparent and less described in the context of IPTW using PS (Austin, 2008, 2009c; Austin & Stuart, 2015a). Even less is present for the continuous setting on balancing for IPTW. However the quantity computed for the denominator in the formula for inverse probability weights (Robins & Finkelstein, 2000) is based on the Hirano and Imbens GPS (Hirano & Imbens, 2004).
In our study IPTW estimates consistently showed higher estimates, even if they have been trimmed to exclude extreme values. No direct covariate balance diagnostic has been done, since no reference methodology is present related to the continuous setting, however the Bia and Mattei (2014) procedure for the GPS balancing has been implemented before creating the weights. It is believable that this procedure could detect the presence of unbalance among the observations.

The results obtained with Imai van Dyk method were not robust to the change in the final response model. In fact, the HR estimated for PM_{10} in the model stratified for 10 subclasses of the PF was smaller (and not significant) than what was observed in the model stratified for 5 subclasses. The same has been observed when adjusting for the PF in each stratum of the PF function.
Since no evidence or comparison is present in literature about this response model specification, no indication can be given on which of these methods is better to apply. In our case, the subclassification in 5 classes, following indication from Austin 2013, resulted in a more conservative approach that prevented the loss of power in the estimates.

Imai and van Dyk verified through simulations that bias and MSE reduction through subclassification on the PF is relatively robust to model misspecification (Imai & Van Dyk, 2004). However, care should be taken when selecting both the PF model and the response model. Model diagnostics in our study showed covariate balance for all the different specification of the model tested (Table 12).
The response model produced different hazard ratio estimates according to the different model specifications. In fact, for both the IvD and RISt methods applied, the introduction of the SEP index in the propensity score specification model slightly reduced the intensity of the effects, while instead the introduction of both the age-standardized hospitalization rates for COPD and lung cancer increased the estimates by 1%. What happened for the SEP is probably due to an overfitting of the model, that already contained all the covariates from which the SEP index was built. For the second result, since the two SHRs were used in the model as proxies for smoking, this result could be interpreted as a better explication of the differences among the units that accounts now also for the previously unmeasured smoking habit. The changes in the final effect estimates were very slight but attention should be given to the model selection.

In fact, there is lack of consensus in literature as to which variables to include in the propensity score model. Rosenbaum (Paul R Rosenbaum, 2010, p. 356) suggested that one should address the issue of variable selection for PS by asking “which covariates do you wish to balance by matching on the propensity score”.

What Austin (2014; 2007) recommended was to identify the potentially prognostically important covariates basing on "subject matter expertise and a review of the existing literature, rather than on formal statistical hypothesis testing in the study sample". Causal diagrams can help identify those covariates in the spirit of basing the decision in the design phase of the study and not in the analysis phase, through statistical testing. (Rubin, 2007)

Exposure-response curves estimated through Hirano and Imbens, Flores and with additive splines produced one of the first evidence in the causal inference setting in air pollution studies.
Several studies investigated the shape of the exposure-response curve for particulate matter, all confirming a linear relationship with mortality (M. Daniels, Dominici, Zeger, & Samet, 2004; Evangelia Samoli et al., 2005; Schwartz & Zanobetti, 2000; Stafoggia et al., 2013). In this study three response models were specified according to different GPS relationship with exposure and outcome. The curves show a linear sharp increase of mortality rates starting from very low levels of PM_{10}, and a subsequent slight decrease in the response. No bootstrap standard errors have been calculated for these curves, so it is possible that a flat zone could be resulting where we currently see a decrease in the curve.
It has been demonstrated that the performance of Hirano and Imbens's method is poor and slightly differs from the unadjusted regression, while the additive spline curve by Flores can be an improvement of the estimate, especially for higher order treatments, though it can introduce a cyclic artefact to the fit (Zhao et al., 2013), as we also observed in our data.
It is very difficult to accurately represent the response model, and results are strongly dependent on parametric assumptions. One way to improve it, like we did, is to use more flexible response model (Flores et al., 2012; Zhao et al., 2013), to overcome the extrapolation risks of fitting the response model at all values of the GPS, including at unobserved combinations of exposures and covariates.
To obtain an estimation of the DRF with the Cox model, it was necessary to reconduct the data to a Poisson process. Even though the equivalence between the two models was verified, some concerns could still be present about the holding of the Rubin Causal Model (RCM) assumption in this setting, where units are split for their time of follow-up, but keeping the same exposure at baseline fixed.
More research is needed to adapt the current knowledge and methods of GPS to the Cox proportional hazard models.

One of the assumptions made in this study was the unconfoundedness assumption (Paul R Rosenbaum & Rubin, 1983). This condition is untestable from the data. In this study several census tract level socioeconomic indicators and SHRs were considered and included in the PS specification, to account for all individual level factors that could confound the association between exposure and outcome.
However, it is always possible that some unmeasured confounders could still bias the results, due to the fact that the information is not taken at the individual level. It has been argued that census block socioeconomic data as a surrogate for individual level socioeconomic position indicators may actually underestimate the

severity of the individual level SEP (Krieger et al., 2002; Subramanian, Chen, Rehkopf, Waterman, & Krieger, 2006).
One other assumption that is at the basis of the PS methods is the SUTVA (Stable Unit Treatment Value Assumption) that requires independence of the potential outcomes in one unit from the specific treatment assigned to other units. In the case of this study, and in general in air pollution studies, this is hard to maintain.
In fact, the exposure in one address for one person in a specific point in time could affect or be affected by the exposure level of a nearby address. It means that we can't guarantee that in the Taranto area the exposure levels estimated in the Borgo district are not affected by the exposure levels observed in the Tamburi district, where the ILVA plant is situated.

In long term-term effect studies, time also plays a role. In this part of the study we decided to consider only the exposure to PM_{10} at baseline, while in the main study in Chapter 4 all the exposure history has been reconstructed, and effects on health were estimated in time-varying models. It is legitimate to think that the past exposure of one person could have an effect in the present, even if the person is currently exposed to a less harmful level of pollution. As a consequence, the correlation in time of the treatment/exposure needs to be taken into account properly. Cumulative exposures could probably be used to overcome the SUTVA assumption. The time varying continuous exposure was beyond the objectives of this research and it was not treated, but it could certainly be a development of the study. No evidence is available up till now on time-varying continuous treatment for propensity score analysis.

Results from the models estimated through GPS and PF are coherent for all the causes of mortality and also coherent with what found in the main study (Chapter 4).
We believe that the potential bias due to possible differences in the exposed and not-exposed subjects have been taken into account satisfactorily by using the census tract socioeconomic information at baseline, both in the main and in the generalized propensity score approaches.
In fact the indirect adjustment methodology, applied in Chapter 4, already detected a negligible or absent correlation between lifestyle factors (smoking, drinking alcohol, and obesity), not directly measured in the cohort, and the exposure.

The exercise about the Generalized Propensity score methods proposed in this thesis highlighted how further research is needed to fill the existing gaps in the methodologies, and consequently in the literature, regarding the use of the GPS for continuous exposures in survival analysis.

6.3.Conclusions

This thesis has explored the use of causal inference methods in public health, where more often observational studies are conducted rather than randomized

experiments, due to the unethical, impractical and length implications for a timely decision making (Glass, Goodman, Hernán, & Samet, 2013).
The most important consequence deriving from the use of observational studies is that the interventions under consideration may be vaguely defined. A way to overcome this concern is to design the observational studies in order to mimic a randomized experiment with a well-defined intervention/exposure.
The knowledge of the intrinsic assignment mechanism is crucial to make the compared groups conditionally exchangeable, by controlling for example for variables that influence both assignment and outcomes, when they are measures, or by correctly designing natural experimental studies (Craig et al., 2017).

Three methods were presented in this thesis to estimate the association of industrial air pollution with health effects on a at-risk population in an observational longitudinal study.
The first method was the standard regression adjustment with multivariable Cox models, which controls for observed differences between exposure and control groups. This approach takes into account factors that are associated with both the exposure and the outcome, and assume that all such factors have been measured accurately, so that there are no unmeasured confounders (Craig et al., 2012; Craig et al., 2017).

The second approach used was the DID that compares the change in an outcome among people who are newly exposed to an intervention with change among those who remain unexposed. The strength of this method is that it controls for both observed and unobserved fixed characteristics of the groups and it is less prone to be affected by bias due to omitted variables or measurement error. One main limitation instead is the fact that the method relies on the assumption of no group-specific trends that may influence outcomes, i.e. the *parallel trend assumption*.
This strong assumption can't be tested and can be easily violated by differential changes in the composition of the two compared groups.

The third method was a propensity score based approach, with a generalization to the continuous exposures. In the absence of a randomization of the exposure, the PS based method tries to recreate the allocation mechanism, by using regression models for the conditional probability of being assigned a specific exposure given a set of covariates. The strength of this method is that it adjusts for the differences in the characteristics of the exposure in the exposed and unexposed groups, and doing so, minimizes the effects of confounding and allows for balanced comparisons. The main limitation of this method is that it strictly depends on the specification model for the propensity score given the observed and available variables, with the main assumption that no other unmeasured confounder plays a role in the mechanism of exposure allocation and no measurement error occurs (Craig et al., 2012; Craig et al., 2017)

In practice, none of these approaches provide a comprehensive unique solution to the central problem of selective exposure to the intervention (Rubin, 2008).
In fact, methods like matching, regression, and propensity score, controlling for observed factors, can be used to reduce bias but they are vulnerable to selection on

unobservables. Methods dealing with unobservable factors, like difference-in-differences, instrumental variables, and regression discontinuity design, rely on strong and untestable assumptions.

To provide more strength to the inference the use of conjoint methods as well as the replication of the study by using different designs, and falsification tests are indicated. Consistent findings from studies using different design, in fact, make less likely that the same common biases are present, and consistent findings increase the confidence in the generalisability of the causal inferences.
It is where the current study is positioned, showing that the three approaches all lead to the same conclusions, despite the different study designs and study periods involved.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies, 72*(1), 1-19.

Afendulis, C. C., He, Y., Zaslavsky, A. M., & Chernew, M. E. (2011). The impact of Medicare Part D on hospitalization rates. *Health services research, 46*(4), 1022-1038.

Alessandrini, E. R., Faustini, A., Chiusolo, M., Stafoggia, M., Gandini, M., Demaria, M., . . . Canova, C. (2013). Inquinamento atmosferico e mortalità in venticinque città italiane: risultati del progetto EpiAir2. *Epidemiologia e Prevenzione, 37*(4-5), 220-229.

Ambiente, C. S. (2016). Studio di coorte sugli effetti delle esposizioni ambientali ed occupazionali sulla morbosità e mortalità della popolazione residente a Taranto.

Analitis, A., Katsouyanni, K., Dimakopoulou, K., Samoli, E., Nikoloulopoulos, A. K., Petasakis, Y., . . . Cambra, K. (2006). Short-term effects of ambient particles on cardiovascular and respiratory mortality. *Epidemiology, 17*(2), 230-233.

Ancona, C., Badaloni, C., Mataloni, F., Bolignano, A., Bucci, S., Cesaroni, G., . . . Forastiere, F. (2015). Mortality and morbidity in a population exposed to multiple sources of air pollution: A retrospective cohort study using air dispersion models. *Environmental Research, 137*, 467-474.

Andersen, P. K., Borgan, O., Gill, R. D., & Keiding, N. (1993). *Statistical models based on counting processes*: Springer Science & Business Media.

Andjelkovich, D. A., Mathew, R. M., Richardson, R. B., & Levine, R. J. (1990). Mortality of iron foundry workers: I. Overall findings. *Journal of Occupational and Environmental Medicine, 32*(6), 529-540.

Ashenfelter, O. C., & Card, D. (1985). Using the longitudinal structure of earnings to estimate the effect of training programs: National Bureau of Economic Research Cambridge, Mass., USA.

Austin, P. C. (2008). Goodness-of-fit diagnostics for the propensity score model when estimating treatment effects using covariate adjustment with the propensity score. *Pharmacoepidemiology and drug safety, 17*(12), 1202-1217.

Austin, P. C. (2009a). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in medicine, 28*(25), 3083-3107.

Austin, P. C. (2009c). The relative ability of different propensity score methods to balance measured covariates between treated and untreated subjects in observational studies. *Medical Decision Making, 29*(6), 661-677.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research, 46*(3), 399-424.

Austin, P. C. (2013). The performance of different propensity score methods for estimating marginal hazard ratios. *Statistics in medicine, 32*(16), 2837-2849.

Austin, P. C. (2014). The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. *Statistics in medicine, 33*(7), 1242-1258.

Austin, P. C., Grootendorst, P., & Anderson, G. M. (2007). A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a Monte Carlo study. *Statistics in medicine, 26*(4), 734-753.

Austin, P. C., & Stuart, E. A. (2015a). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in medicine, 34*(28), 3661-3679.

Austin, P. C., & Stuart, E. A. (2015c). The performance of inverse probability of treatment weighting and full matching on the propensity score in the presence of model misspecification when estimating the effect of treatment on survival outcomes. *Statistical methods in medical research*, 0962280215584401.

Baccini, M., Mattei, A., Mealli, F., Bertazzi, P. A., & Carugno, M. (2017). Assessing the short term impact of air pollution on mortality: a matching approach. *Environmental Health, 16*(1), 7.

Bae, S., Kim, H. C., Ye, B., Choi, W. J., Hong, Y. S., & Ha, M. (2017). Causal inference in environmental epidemiology. *Environ Health Toxicol, 32*, e2017015. doi: 10.5620/eht.e2017015

Beelen, R., Hoek, G., van Den Brandt, P. A., Goldbohm, R. A., Fischer, P., Schouten, L. J., . . . Brunekreef, B. (2008). Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study). *Environmental health perspectives, 116*(2), 196.

Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z. J., Weinmayr, G., Hoffmann, B., . . . Nieuwenhuijsen, M. (2014). Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *The Lancet, 383*(9919), 785-795.

Bentayeb, M., Wagner, V., Stempfelet, M., Zins, M., Goldberg, M., Pascal, M., . . . Eilstein, D. (2015). Association between long-term exposure to air pollution and mortality in France: A 25-year follow-up study. *Environment international, 85*, 5-14.

Berry, G. (1983). The analysis of mortality by the subject-years method. *Biometrics*, 173-184.

Berti, G., Galassi, C., Faustini, A., & Forastiere, F. (2009). EPIDEMIOLOGIA & PREVENZIONE. *Rivista dell'Associazione italiana di epidemiologia ANNO, 33*(6), 1.

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

Bia, M., & Mattei, A. (2008). A Stata package for the estimation of the dose–response function through adjustment for the generalized propensity score. *The Stata Journal, 8*(3), 354-373.

Biggeri, A., Bellini, P., & Terracini, B. (2004). Meta-analysis of the Italian studies on short-term effects of air pollution--MISA 1996-2002.

Blundell, R., & Dias, M. C. (2009). Alternative approaches to evaluation in empirical microeconomics. *Journal of Human Resources, 44*(3), 565-640.

Bourgkard, E., Wild, P., Courcot, B., Diss, M., Ettlinger, J., Goutet, P., . . . Rigal, C. (2008). Lung cancer mortality and iron oxide exposure in a French steel-producing factory. *Occupational and Environmental Medicine*.

Brand, R., Pulles, T., Gijlswijk, R. v., Fribourg-Blanc, B., & Courbet, C. (2004). EPER European Pollutant Emission Register. *European Commission*.

Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American journal of epidemiology, 163*(12), 1149-1156.

Caliendo, M., & Kopeinig, S. (2008). SOME PRACTICAL GUIDANCE FOR THE IMPLEMENTATION OF PROPENSITY SCORE MATCHING. *Journal of Economic Surveys, 22*(1), 31-72. doi: 10.1111/j.1467-6419.2007.00527.x

Cancer, I. A. f. R. o. (2014). IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. *Outdoor Air Pollution. Lyon, France, 109*.

Capuno, J. J., Tan, C. A. R., & Javier, X. (2018). Cooking and coughing: Estimating the effects of clean fuel for cooking on the respiratory health of children in the Philippines. *Global Public Health, 13*(1), 20-34. doi: 10.1080/17441692.2016.1202297

Caranci, N., Biggeri, A., Grisotto, L., Pacelli, B., Spadea, T., & Costa, G. (2010). The Italian deprivation index at census block level: definition, description and association with general mortality. *Epidemiologia e Prevenzione, 34*(4), 167-176.

Card, D., & Krueger, A. B. (1994). Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania: National Bureau of Economic Research.

Centro nazionale di epidemiologia, s. e. p. d. s. d. I. s. d. s. La sorveglianza PASSI. *Il portale dell'epidemiologia per la sanità pubblica*.

Cesaroni, G., Agabiti, N., Rosati, R., Forastiere, F., & Perucci, C. (2006). An index of socioeconomic position based on 2001 Census, Rome. *Epidemiologia e Prevenzione, 30*(6), 352-357.

Cesaroni, G., Badaloni, C., Gariazzo, C., Stafoggia, M., Sozzi, R., Davoli, M., & Forastiere, F. (2013). Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. *Environmental health perspectives, 121*(3), 324.

Cochran, W. G. (1968). The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*, 295-313.

Cochran, W. G., & Chambers, S. P. (1965). The Planning of Observational Studies of Human Populations. *Journal of the Royal Statistical Society. Series A (General), 128*(2), 234-266. doi: 10.2307/2344179

Colais, P., Serinelli, M., Faustini, A., Stafoggia, M., Randi, G., Tessari, R., . . . Vigotti, M. A. (2009). Air pollution and urgent hospital admissions in nine Italian cities. Results of the EpiAir Project. *Epidemiologia e Prevenzione, 33*(6 Suppl 1), 77-94.

Cole, S. R., & Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American journal of epidemiology, 168*(6), 656-664.

Comba, P., Pirastu, R., Conti, S., De Santis, M., Iavarone, I., Marsili, G., . . . Minerba, S. (2012). Environment and health in Taranto, southern Italy: epidemiological studies and public health recommendations. *Epidemiologia e Prevenzione, 36*(6), 305-320.

Comfort, A., Leegwater, A., Nakhimovsky, S., Kansembe, H., Hamainza, B., Bwalya, B., . . . Olsho, L. (2017). Exploring the use of routinely-available, retrospective data to study the association between malaria control scale-up and micro-economic outcomes in Zambia. *Malaria journal, 16*(1), 15.

Correia, A. W., Pope III, C. A., Dockery, D. W., Wang, Y., Ezzati, M., & Dominici, F. (2013). The effect of air pollution control on life expectancy in the United States: an analysis of 545 US counties for the period 2000 to 2007. *Epidemiology (Cambridge, Mass.), 24*(1), 23.

Cox, D. R. (1958). Planning of experiments.

Craig, P., Cooper, C., Gunnell, D., Haw, S., Lawson, K., Macintyre, S., . . . Sutton, M. (2012). Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health*, jech-2011-200375.

Craig, P., Katikireddi, S. V., Leyland, A., & Popham, F. (2017). Natural experiments: an overview of methods, approaches, and contributions to Public Health intervention research. *Annual review of public health, 38*, 39-56.

Crouse, D. L., Peters, P. A., van Donkelaar, A., Goldberg, M. S., Villeneuve, P. J., Brion, O., . . . Pope III, C. A. (2012). Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study. *Environmental health perspectives, 120*(5), 708.

D'Agostino Jr, R. B., & Rubin, D. B. (2000). Estimating and using propensity scores with partially missing data. *Journal of the American Statistical Association, 95*(451), 749-759.

Daniels, M., Dominici, F., Zeger, S., & Samet, J. (2004). The National Morbidity, Mortality, and Air Pollution Study. Part III: PM10 concentration-response curves and thresholds for the 20 largest US cities. *Research report (Health Effects Institute)*(94 Pt 3), 1-21; discussion 23-30.

Daniels, M. J., Dominici, F., Samet, J. M., & Zeger, S. L. (2000). Estimating particulate matter-mortality dose-response curves and threshold levels: an analysis of daily time-series for the 20 largest US cities. *American journal of epidemiology, 152*(5), 397-406.

Di Gilio, A., Ventrella, G., Giungato, P., Tutino, M., Giua, R., Assennato, G., & de Gennaro, G. (2017). An intensive monitoring campaign of PAHs for assessing the impact of a steel plant. *Chemosphere, 168*, 171-182.

Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., . . . Speizer, F. E. (1993). An association between air pollution and mortality in six US cities. *New England journal of medicine, 329*(24), 1753-1759.

Enderlein, G. (1987). McCullagh, P., JA Nelder: Generalized linear models. Chapman and Hall London–New York 1983, 261 S.,£ 16,–: Wiley Online Library.

Ertefaie, A., & Stephens, D. A. (2010). Comparing approaches to causal inference for longitudinal data: inverse probability weighting versus propensity scores. *The international journal of biostatistics, 6*(2).

Filleul, L., Rondeau, V., Vandentorren, S., Le Moual, N., Cantagrel, A., Annesi-Maesano, I., . . . Paris, C. (2005). Twenty five year mortality and air pollution: results from the French PAARC survey. *Occupational and Environmental Medicine, 62*(7), 453-460.

Finkelstein, M. M., & Wilk, N. (1990). Investigation of a lung cancer cluster in the melt shop of an Ontario steel producer. *American journal of industrial medicine, 17*(4), 483-491.

Fletcher, A. C., & Ades, A. (1984). Lung cancer mortality in a cohort of English foundry workers. *Scandinavian journal of work, environment & health*, 7-16.

Flores, C. A., Flores-Lagunes, A., Gonzalez, A., & Neumann, T. C. (2012). Estimating the effects of length of exposure to instruction in a training program: the case of job corps. *Review of Economics and Statistics, 94*(1), 153-171.

Forastiere, F., Biggeri, A., & Triassi, M. (2012). Perizia Epidemiologica. *Environment Ministry*.

Galagate, D., Schafer, J., & Galagate, M. D. (2015). Package 'causaldrf'.

Gasparrini, A. (2014). Modeling exposure–lag–response associations with distributed lag non-linear models. *Statistics in medicine, 33*(5), 881-899.

Gayat, E., Resche-Rigon, M., Mary, J. Y., & Porcher, R. (2012). Propensity score applied to survival data analysis through proportional hazards models: a Monte Carlo study. *Pharmaceutical statistics, 11*(3), 222-229.

Gillespie, I. A., Floege, J., Gioni, I., Drüeke, T. B., Francisco, A. L., Anker, S. D., . . . Froissart, M. (2015). Propensity score matching and persistence correction to reduce bias in comparative effectiveness: the effect of cinacalcet use on all-cause mortality. *Pharmacoepidemiology and drug safety, 24*(7), 738-747.

Giua, R., Morabito, A., Tanzarella, A., Spagnolo, S., Pastore, T., Bevere, M., . . . Brusasca, G. (2014). Application of a Lagrangian particle model to the source apportionment for primary macropollutants in Taranto area (South Italy). *International Journal of Environment and Pollution, 55*(1-4), 31-40.

Glass, T. A., Goodman, S. N., Hernán, M. A., & Samet, J. M. (2013). Causal inference in public health. *Annual review of public health, 34*, 61-75.

Graziano, G., Bilancia, M., Bisceglia, L., Pollice, A., & Assennato, G. (2009). Statistical analysis of the incidence of some cancers in the province of Taranto 1999-2001. *Epidemiologia e Prevenzione, 33*(1-2), 37-44.

Greenland, S., Pearl, J., & Robins, J. M. (1999). Causal diagrams for epidemiologic research. *Epidemiology, 10*(1), 37-48.

Gu, X. S., & Rosenbaum, P. R. (1993). Comparison of multivariate matching methods: Structures, distances, and algorithms. *Journal of Computational and Graphical Statistics, 2*(4), 405-420.

Guo S, & Fraser, M. (2010). Propensity score analysis: Statistical methods and analysis: Thousand Oaks, CA: Sage.

Guo, S., & Fraser, M. W. (2014). *Propensity score analysis* (Vol. 12): Sage.

Hamra, G. B., Guha, N., Cohen, A., Laden, F., Raaschou-Nielsen, O., Samet, J. M., . . . Yorifuji, T. (2014). Outdoor particulate matter exposure and lung cancer: a systematic review and meta-analysis. *Environmental health perspectives, 122*(9), 906.

Hernán, M. A., Alonso, A., Logan, R., Grodstein, F., Michels, K. B., Stampfer, M. J., . . . Robins, J. M. (2008). Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology (Cambridge, Mass.), 19*(6), 766.

Hill, A. B. (1965). THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION? *Proc R Soc Med, 58*, 295-300.

Hirano, K., & Imbens, G. W. (2004). The propensity score with continuous treatments. *Applied Bayesian modeling and causal inference from incomplete-data perspectives, 226164*, 73-84.

Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association, 81*(396), 945-960. doi: 10.1080/01621459.1986.10478354

Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association, 47*(260), 663-685.

Hoshuyama, T., Pan, G., Tanaka, C., Feng, Y., Liu, T., Liu, L., . . . Takahashi, K. (2006). Mortality of iron-steel workers in Anshanl China: A retrospective cohort study. *International journal of occupational and environmental health, 12*(3), 193-202.

IARC. (1987). monographs on the Evaluation of Carcinogenic Risks to Humans. Overall Evaluation of Carcinogenicity: An Updating of IARC Monographs. *1-42, Supplement7*, 224-225.

Imai, K., King, G., & Stuart, E. A. (2008). Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the royal statistical society: series A (statistics in society), 171*(2), 481-502.

Imai, K., & Van Dyk, D. A. (2004). Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association, 99*(467), 854-866.

Imbens, G. W. (1999). The Role of the Propensity Score in Estimating Dose-Response Functions. *National Bureau of Economic Research Technical Working Paper Series, No. 237*. doi: 10.3386/t0237

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika, 87*(3), 706-710.

Jacob, B. J., Sutradhar, R., Moineddin, R., Baxter, N. N., & Urbach, D. R. (2013). Methodological approaches to population based research of screening procedures in the presence of selection bias and exposure measurement error: colonoscopy and colorectal cancer outcomes in Ontario. *BMC medical research methodology, 13*(1), 59.

Jerrett, M., Finkelstein, M. M., Brook, J. R., Arain, M. A., Kanaroglou, P., Stieb, D. M., . . . Chapman, K. R. (2009). A cohort study of traffic-related air pollution and mortality in Toronto, Ontario, Canada. *Environmental health perspectives, 117*(5), 772.

Jiang, M., & Foster, E. M. (2013). Duration of breastfeeding and childhood obesity: a generalized propensity score approach. *Health services research, 48*(2pt1), 628-651.

Kirby, M. A., Nagel, C., Rosa, G. A., Thomas, E. A., & Clasen, T. F. (2016). Assessing use, exposure and health impacts of an advanced water filter and advanced cookstove distribution program in Rural Rwanda. *American Journal of Tropical Medicine and Hygiene, 95*(5), 582-583. doi: 10.4269/ajtmh.abstract2016

Kloog, I., Ridgway, B., Koutrakis, P., Coull, B. A., & Schwartz, J. D. (2013). Long-and short-term exposure to PM2. 5 and mortality: using novel exposure models. *Epidemiology (Cambridge, Mass.), 24*(4), 555.

Koch, C. G., Li, L., Kaplan, G. A., Wachterman, J., Shishehbor, M. H., Sabik, J., & Blackstone, E. H. (2010). Socioeconomic position, not race, is linked to death after cardiac surgery. *Circulation: Cardiovascular Quality and Outcomes*, CIRCOUTCOMES. 109.880377.

Krall, J. R., Anderson, G. B., Dominici, F., Bell, M. L., & Peng, R. D. (2013). Short-term exposure to particulate matter constituents and mortality in a national study of US urban communities. *Environmental health perspectives, 121*(10), 1148.

Krieger, N., Chen, J. T., Waterman, P. D., Soobader, M.-J., Subramanian, S., & Carson, R. (2002). Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter? the Public Health Disparities Geocoding Project. *American journal of epidemiology, 156*(5), 471-482.

Laden, F., Schwartz, J., Speizer, F. E., & Dockery, D. W. (2006). Reduction in fine particulate air pollution and mortality: extended follow-up of the Harvard Six Cities study. *American Journal of Respiratory And Critical Care Medicine, 173*(6), 667-672.

Laird, N., & Olivier, D. (1981). Covariance analysis of censored survival data using log-linear analysis techniques. *Journal of the American Statistical Association, 76*(374), 231-240.

Lee, B. K., Lessler, J., & Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in medicine, 29*(3), 337-346.

Lepeule, J., Laden, F., Dockery, D., & Schwartz, J. (2012). Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009. *Environmental health perspectives, 120*(7), 965.

Little, R. J., & Rubin, D. B. (2000). Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu Rev Public Health, 21*, 121-145. doi: 10.1146/annurev.publhealth.21.1.121

Liu, L., Kauri, L. M., Mahmud, M., Weichenthal, S., Cakmak, S., Shutt, R., . . . Kumarathasan, P. (2014). Exposure to air pollution near a steel plant and effects on cardiovascular physiology: a randomized crossover study. *International journal of hygiene and environmental health, 217*(2), 279-286.

Lu, B., & Marcus, S. (2012). Evaluating long-term effects of a psychiatric treatment using instrumental variable and matching approaches. *Health Services and Outcomes Research Methodology, 12*(4), 288-301.

Lu, B., Zanutto, E., Hornik, R., & Rosenbaum, P. R. (2001). Matching with doses in an observational study of a media campaign against drug abuse. *Journal of the American Statistical Association, 96*(456), 1245-1253.

Lu, M. (1999). The productivity of mental health care: an instrumental variable approach. *The journal of mental health policy and economics, 2*(2), 59-71.

Madouasse, A., Browne, W., Huxley, J., Toni, F., & Green, M. (2012). A semi-parametric model for lactation curves: Development and application. *Preventive veterinary medicine, 105*(1), 38-48.

Mangia, C., Gianicolo, E. A., Bruni, A., Vigotti, M. A., & Cervino, M. (2013). Spatial variability of air pollutants in the city of Taranto, Italy and its potential impact on exposure assessment. *Environmental monitoring and assessment, 185*(2), 1719-1735.

Martinelli, D., Mincuzzi, A., Minerba, S., Tafuri, S., Conversano, M., Caputi, G., . . . Prato, R. (2009). Malignant cancer mortality in Province of Taranto (Italy). Geographic analysis in an area of high environmental risk. *Journal of preventive medicine and hygiene, 50*(3).

Martuzzi, M., Mitis, F., Biggeri, A., Terracini, B., & Bertollini, R. (2002). Environment and health status of the population in areas with high risk of environmental crisis in Italy. *Epidemiologia e Prevenzione, 26*(6 Suppl), suppl 1-53.

Mataloni, F., Stafoggia, M., Alessandrini, E., Triassi, M., Biggeri, A., & Forastiere, F. (2012). Studio di coorte sulla mortalità e morbosità nell'area di Taranto. *Epidemiol Prev, 36*(5), 237-252.

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological methods, 9*(4), 403.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of business & economic statistics, 13*(2), 151-161.

Mitis, F., Martuzzi, M., Biggeri, A., Bertollini, R., & Terracini, B. (2005). Industrial activities in sites at high environmental risk and their impact on the health of the population. *International journal of occupational and environmental health, 11*(1), 88-95.

Moodie, E. E., & Stephens, D. A. (2012). Estimation of dose–response functions for longitudinal data using the generalised propensity score. *Statistical methods in medical research, 21*(2), 149-166.

Mueller, V., Pfaff, A., Peabody, J., Liu, Y., & Smith, K. R. (2011). Demonstrating bias and improved inference for stoves' health benefits. *International Journal of Epidemiology, 40*(6), 1643-1651. doi: 10.1093/ije/dyr150

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications, 9*(1), 141-142.

Neupane, M., Basnyat, B., Fischer, R., Froeschl, G., Wolbers, M., & Rehfuess, E. A. (2015). Sustained use of biogas fuel and blood pressure among women in rural Nepal. *Environmental Research, 136*, 343-351. doi: 10.1016/j.envres.2014.10.031

Park, R. M., Ahn, Y. S., Stayner, L. T., Kang, S. K., & Jang, J. K. (2005). Mortality of iron and steel workers in Korea. *American journal of industrial medicine, 48*(3), 194-204.

Peng, R. D., Samoli, E., Pham, L., Dominici, F., Touloumi, G., Ramsay, T., . . . Cohen, A. (2013). Acute effects of ambient ozone on mortality in Europe and North America: results from the APHENA study. *Air Quality, Atmosphere & Health, 6*(2), 445-453.

Podolanczuk, A. J., Oelsner, E. C., Barr, R. G., Bernstein, E. J., Hoffman, E. A., Easthausen, I. J., . . . Raghu, G. (2017). High Attenuation Areas on Chest CT and Clinical Respiratory Outcomes in Community-dwelling Adults. *American Journal of Respiratory And Critical Care Medicine*(ja).

Pope 3rd, C. (1989). Respiratory disease associated with community air pollution and a steel mill, Utah Valley. *American Journal of Public Health, 79*(5), 623-628.

Pope 3rd, C., Hill, R. W., & Villegas, G. M. (1999). Particulate air pollution and daily mortality on Utah's Wasatch Front. *Environmental health perspectives, 107*(7), 567.

Pope, C. A., Burnett, R. T., Thurston, G. D., Thun, M. J., Calle, E. E., Krewski, D., & Godleski, J. J. (2004). Cardiovascular mortality and long-term exposure to particulate air pollution. *Circulation, 109*(1), 71-77.

Pope III, C. A., Ezzati, M., & Dockery, D. W. (2009). Fine-particulate air pollution and life expectancy in the United States. *N Engl J Med, 2009*(360), 376-386.

Pope III, C. A., Schwartz, J., & Ransom, M. R. (1992). Daily mortality and PM10 pollution in Utah Valley. *Archives of Environmental Health: An International Journal, 47*(3), 211-217.

Robins, J. M., & Finkelstein, D. M. (2000). Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics, 56*(3), 779-788.

Rodríguez, V., Tardón, A., Kogevinas, M., Prieto, C. S., Cueto, A., García, M., . . . Zaplana, J. (2000). Lung cancer risk in iron and steel foundry workers: a nested case control study in Asturias, Spain. *American journal of industrial medicine, 38*(6), 644-650.

Rosa, M. J., Pajak, A., Just, A. C., Sheffield, P. E., Kloog, I., Schwartz, J., . . . Wright, R. J. (2017). Prenatal exposure to PM2.5 and birth weight: A pooled analysis from three North American longitudinal pregnancy cohort studies. *Environ Int, 107*, 173-180. doi: 10.1016/j.envint.2017.07.012

Rosenbaum, P. R. (1987). Model-based direct adjustment. *Journal of the American Statistical Association, 82*(398), 387-394.

Rosenbaum, P. R. (1996). 6 Observational studies and nonrandomized experiments *Handbook of Statistics* (Vol. 13, pp. 181-197): Elsevier.

Rosenbaum, P. R. (2010). Design sensitivity and efficiency in observational studies. *Journal of the American Statistical Association, 105*(490), 692-702.

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

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association, 79*(387), 516-524.

Rubin, D. B. (1973a). Matching to remove bias in observational studies. *Biometrics*, 159-183.

Rubin, D. B. (1973c). The use of matched sampling and regression adjustment to remove bias in observational studies. *Biometrics*, 185-203.
Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology, 66*(5), 688.
Rubin, D. B. (1976). Multivariate Matching Methods That are Equal Percent Bias Reducing, I: Some Examples. *Biometrics, 32*(1), 109-120. doi: 10.2307/2529342
Rubin, D. B. (1978). Bayesian Inference for Causal Effects: The Role of Randomization. *The Annals of Statistics, 6*(1), 34-58.
Rubin, D. B. (1979). Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies. *Journal of the American Statistical Association, 74*(366), 318-328. doi: 10.2307/2286330
Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of internal medicine, 127*(8_Part_2), 757-763.
Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology, 2*(3), 169-188.
Rubin, D. B. (2004). On principles for modeling propensity scores in medical research. *Pharmacoepidemiology and drug safety, 13*(12), 855-857.
Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association, 100*(469), 322-331.
Rubin, D. B. (2007). The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Statistics in medicine, 26*(1), 20-36.
Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *The Annals of Applied Statistics*, 808-840.
Rubin, D. B., & Thomas, N. (1996). Matching using estimated propensity scores: relating theory to practice. *Biometrics*, 249-264.
Rubin, D. B., & Thomas, N. (2000). Combining propensity score matching with additional adjustments for prognostic covariates. *Journal of the American Statistical Association, 95*(450), 573-585.
Sætterstrom, B., Kruse, M., BrØnnum-Hansen, H., BØnlØkke, J. H., Meulengracht Flachs, E., & SØrensen, J. (2012). A method to assess the potential effects of air pollution mitigation on healthcare costs. *Journal of Environmental and Public Health, 2012*. doi: 10.1155/2012/935825
Samoli, E., Aga, E., Touloumi, G., Nisiotis, K., Forsberg, B., Lefranc, A., . . . Niciu, E. (2006). Short-term effects of nitrogen dioxide on mortality: an analysis within the APHEA project. *European Respiratory Journal, 27*(6), 1129-1138.
Samoli, E., Analitis, A., Touloumi, G., Schwartz, J., Anderson, H. R., Sunyer, J., . . . Pekkanen, J. (2005). Estimating the exposure–response relationships between particulate matter and mortality within the APHEA multicity project. *Environmental health perspectives, 113*(1), 88.
Samoli, E., Stafoggia, M., Rodopoulou, S., Ostro, B., Declercq, C., Alessandrini, E., . . . Le Tertre, A. (2013). Associations between fine and coarse particles and mortality in Mediterranean cities: results from the MED-PARTICLES project. *Environmental health perspectives, 121*(8), 932.

Sanna, M., Monguzzi, R., Santilli, N., & Felici, R. (2012). Perizia chimica della Procura di Taranto. *Environment Ministry*.

Scarinzi, C., Alessandrini, E. R., Chiusolo, M., Galassi, C., Baldini, M., Serinelli, M., . . . De, A. T. (2013). Air pollution and urgent hospital admissions in 25 Italian cities: results from the EpiAir2 project. *Epidemiologia e Prevenzione, 37*(4-5), 230-241.

Schneeweiss, S., Glynn, R. J., Tsai, E. H., Avorn, J., & Solomon, D. H. (2005). Adjusting for unmeasured confounders in pharmacoepidemiologic claims data using external information: the example of COX2 inhibitors and myocardial infarction. *Epidemiology, 16*(1), 17-24.

Schwartz, J. (2004). The effects of particulate air pollution on daily deaths: a multi-city case crossover analysis. *Occupational and Environmental Medicine, 61*(12), 956-961.

Schwartz, J., Austin, E., Bind, M.-A., Zanobetti, A., & Koutrakis, P. (2015). Estimating causal associations of fine particles with daily deaths in Boston. *American journal of epidemiology, 182*(7), 644-650.

Schwartz, J., Ballester, F., Saez, M., Pérez-Hoyos, S., Bellido, J., Cambra, K., . . . Sunyer, J. (2001). The concentration-response relation between air pollution and daily deaths. *Environmental health perspectives, 109*(10), 1001.

Schwartz, J., Bind, M.-A., & Koutrakis, P. (2017). Estimating causal effects of local air pollution on daily deaths: effect of low levels. *Environmental health perspectives, 125*(1), 23.

Schwartz, J., Coull, B., Laden, F., & Ryan, L. (2008). The effect of dose and timing of dose on the association between airborne particles and survival. *Environmental health perspectives, 116*(1), 64.

Schwartz, J., & Zanobetti, A. (2000). Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. *Epidemiology, 11*(6), 666-672.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-experimental Designs for Generalized Causal Inference*: Houghton Mifflin.

Shadish, W. R., & Steiner, P. M. (2010). A primer on propensity score analysis. *Newborn and Infant Nursing Reviews, 10*(1), 19-26.

Shi, L., Zanobetti, A., Kloog, I., Coull, B. A., Koutrakis, P., Melly, S. J., & Schwartz, J. D. (2016). Low-concentration PM2. 5 and mortality: Estimating acute and chronic effects in a population-based study. *Environmental health perspectives, 124*(1), 46.

Shimizu, M., Yi, S., Tuot, S., Suong, S., Sron, S., Shibanuma, A., & Jimba, M. (2016). The impact of a livelihood program on depressive symptoms among people living with HIV in Cambodia. *Global health action, 9*(1), 31999.

Smith, J. A., & Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of econometrics, 125*(1), 305-353.

Sorahan, T., Faux, A. M., & Cooke, M. A. (1994). Mortality among a cohort of United Kingdom steel foundry workers with special reference to cancers of the stomach and lung, 1946-90. *Occupational and Environmental Medicine, 51*(5), 316-322.

Splawa-Neyman, J., Dabrowska, D. M., & Speed, T. P. (1990). On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science, 5*(4), 465-472.

Stafoggia, M., Faustini, A., Rognoni, M., Tessari, R., Cadum, E., Pacelli, B., . . . Vigotti, M. A. (2009). Inquinamento atmosferico e mortalità in dieci città italiane. Risultati del Progetto EpiAir. *Epidemiol Prev, 33*(6 Suppl 1), 65-76.

Stafoggia, M., Samoli, E., Alessandrini, E., Cadum, E., Ostro, B., Berti, G., . . . Pascal, M. (2013). Short-term associations between fine and coarse particulate matter and hospitalizations in Southern Europe: results from the MED-PARTICLES project. *Environmental health perspectives, 121*(9), 1026.

Stampf, S., Graf, E., Schmoor, C., & Schumacher, M. (2010). Estimators and confidence intervals for the marginal odds ratio using logistic regression and propensity score stratification. *Statistics in medicine, 29*(7-8), 760-769.

Stölzel, M., Breitner, S., Cyrys, J., Pitz, M., Wölke, G., Kreyling, W., . . . Peters, A. (2007). Daily mortality and particulate matter in different size classes in Erfurt, Germany. *Journal of Exposure Science and Environmental Epidemiology, 17*(5), 458-467.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics, 25*(1), 1.

Subramanian, S., Chen, J. T., Rehkopf, D. H., Waterman, P. D., & Krieger, N. (2006). Comparing individual-and area-based socioeconomic measures for the surveillance of health disparities: a multilevel analysis of Massachusetts births, 1989–1991. *American journal of epidemiology, 164*(9), 823-834.

Tétreault, L.-F., Doucet, M., Gamache, P., Fournier, M., Brand, A., Kosatsky, T., & Smargiassi, A. (2016). Childhood exposure to ambient air pollutants and the onset of asthma: An administrative cohort study in Quebec. *Environmental health perspectives, 124*(8), 1276.

Thoemmes, F. J., & Kim, E. S. (2011). A systematic review of propensity score methods in the social sciences. *Multivariate behavioral research, 46*(1), 90-118.

Vigotti, M. A., Cavone, D., Bruni, A., Minerba, S., & Conversano, M. (2007). Mortality analysis in a site with point source emissions: the Taranto case (in Italian). *Rapporti ISTISAN, Impatto sulla salute dei siti inquinati:metodi e strumenti per la ricerca e le valutazioni.* (07/50), 155-165.

Vigotti, M. A., Mataloni, F., Bruni, A., Minniti, C., & Gianicolo, E. A. (2014). Mortality analysis by neighbourhood in a city with high levels of industrial air pollution. *International journal of public health, 59*(4), 645-653.

Wahida, K.-T., Padilla, C. M., Denis, Z.-N., Olivier, B., Géraldine, L. N., Philippe, Q., & Séverine, D. (2016). A Conceptual Framework for the Assessment of Cumulative Exposure to Air Pollution at a Fine Spatial Scale. *International journal of environmental research and public health, 13*(3), 319.

Wang, Y., Kloog, I., Coull, B. A., Kosheleva, A., Zanobetti, A., & Schwartz, J. D. (2016). Estimating causal effects of long-term PM2. 5 exposure on mortality in New Jersey. *Environmental health perspectives, 124*(8), 1182.

WHO. (2013). Final technical report.

Wood, S., & Wood, M. S. (2007). The mgcv package. *www. r-project. org*.

Wylie, B. J., Coull, B. A., Hamer, D. H., Singh, M. P., Jack, D., Yeboah-Antwi, K., . . . Macleod, W. B. (2014). Impact of biomass fuels on pregnancy outcomes in central East India. *Environmental Health: A Global Access Science Source, 13*(1). doi: 10.1186/1476-069X-13-1

Wylie, B. J., Coull, B. A., Hamer, D. H., Singh, M. P., Jack, D., Yeboah-Antwi, K., . . . MacLeod, W. B. (2015). Impact of biomass fuels on pregnancy outcomes in central East India. *Environmental Health: A Global Access Science Source, 13*(1). doi: 10.1186/1476-069X-13-1

Wylie, B. J., Singh, M. P., Coull, B. A., Quinn, A., Yeboah-Antwi, K., Sabin, L., . . . Macleod, W. B. (2015). Association between wood cooking fuel and maternal hypertension at delivery in central East India. *Hypertension in Pregnancy, 34*(3), 355-368. doi: 10.3109/10641955.2015.1046604

Zanini, G. (2009). Il sistema MINNI, modello integrato nazionale per la valutazione degli effetti dell'inquinamento atmosferico e dell'efficacia delle politiche di riduzione delle emissioni di inquinanti atmosferici. *Epidemiologia e Prevenzione, 33*(6), 35-42.

Zanobetti, A., Dominici, F., Wang, Y., & Schwartz, J. D. (2014). A national case-crossover analysis of the short-term effect of PM 2.5 on hospitalizations and mortality in subjects with diabetes and neurological disorders. *Environmental Health, 13*(1), 38.

Zanobetti, A., & Schwartz, J. (2009). The effect of fine and coarse particulate air pollution on mortality: a national analysis. *Environmental health perspectives, 117*(6), 898.

Zeger, S. L., Dominici, F., McDermott, A., & Samet, J. M. (2008). Mortality in the Medicare population and chronic exposure to fine particulate air pollution in urban centers (2000–2005). *Environmental health perspectives, 116*(12), 1614.

Zhao, S., van Dyk, D. A., & Imai, K. (2013). Propensity-score based methods for causal inference in observational studies with fixed non-binary treatments: Mimeo (http://imai. princeton. edu/research/files/gpscore. pdf).

Zhu, Y., Coffman, D. L., & Ghosh, D. (2015). A boosting algorithm for estimating generalized propensity scores with continuous treatments. *Journal of causal inference, 3*(1), 25-40.

Zigler, C., Kim, C., Choirat, C., Hansen, J., Wang, Y., Hund, L., . . . Dominici, F. (2016). Causal Inference Methods for Estimating Long-Term Health Effects of Air Quality Regulations.

Zigler, C. M., Dominici, F., & Wang, Y. (2012). Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes. *Biostatistics, 13*(2), 289-302. doi: 10.1093/biostatistics/kxr052

Annex A- Paper on the Main Taranto Cohort Study

Title: A cohort study on air pollution from a large steel plant and mortality and morbidity in the population of Taranto, Italy

Authors: Ester Rita Alessandrini, Massimo Stafoggia, Carla Ancona, Simona Leogrande, Francesca Mataloni, Lucia Bisceglia, Roberto Giua, Antonia Mincuzzi, Sante Minerba, Angela Morabito, Stefano Spagnolo, Giorgio Assennato and Francesco Forastiere, on behalf of the CSA Puglia Study Group

Affiliations

Department of Epidemiology, Lazio Regional Health Service, Rome – ASL Roma 1, Italy (Ester Rita Alessandrini, Massimo Stafoggia, Carla Ancona,Francesca Mataloni, Francesco Forastiere)

ARPA Puglia, Bari, Italy (Roberto Giua, Angela Morabito, Stefano Spagnolo, Giorgio Assennato)

Local Health Service Taranto, Italy (Simona Leogrande, Antonella Mincuzzi, Aldo Minerba)

AReS Puglia, Bari, Italy (Lucia Bisceglia, CSA Puglia Study Group)

Contact information

Ester Alessandrini

Address: via Cristoforo Colombo 112, 00154 – Rome

Email: ester.alessandrini@gmail.com

Shot running title: Air pollution and mortality/morbidity near a large steel plant.

Conflict of Interests. Dr. Forastiere has served as expert for the judge of the Taranto Court in a case involving the ILVA plant; there are no other competing financial interests. The other authors declare they have no actual or potential competing financial interests.

Abstract

Introduction

The association between long-term exposure to air pollution and mortality or morbidity has been widely studied, however the health effects of industrial emissions are less clear, and the relevant time-windows of exposure need to be established. We conducted a cohort study to examine the association between residential exposure to air pollution from a large steel plant located in Taranto (South Italy) and cause-specific mortality and morbidity in the population living in the area.

Methods

The cohort included all subjects (321,356) residents in the area in 1998-2010, followed until the end of 2013. Exposure to PM_{10} and SO_2 originating from the steel plant at each residential address of the study participants (on 2010) was assessed using a Lagrangian dispersion model. A backward and forward extrapolation of the annual exposures during 1965-2013 was applied based on steel production and emission data. The time-dependent annual average exposure was used in Cox proportional-hazard models to investigate the effect of industrial air pollutants on cause-specific mortality, hospital admissions, and cancer incidence. Adjustment was made on individual and contextual variables. The latency of the effects was explored on different 5-year time windows.

Results

A total of 33,042 subjects died from natural causes by the end of the follow-up. Exposures to both industrial PM_{10} and SO_2 at lag 0 (same year) were associated with natural mortality: Hazard Ratios (HR) 1.04 (95% Confidence Interval (CI):1.02-1.06) and 1.09 (95%CI:1.05-1.12) for 10μg/m^3 increments, respectively. Larger associations were observed for mortality from heart diseases (HR 1.05 for PM_{10} and 1.11 for SO_2) and acute myocardial infarction (HR 1.10 for PM_{10} and 1.29 for SO_2). The models based on 5-year exposure windows showed that the effects of both pollutants on natural mortality were driven by the recent exposure (previous 5 years) and the exposures in the distant past (30-35 years lag). Results for hospital admissions and lung cancer incidence were in line with the mortality findings.

Conclusions: Estimated exposures to industrial PM_{10} and SO_2 in this industrial area were associated with several health outcomes. Both recent and distant exposures were responsible for the negative effects on natural mortality.

MAIN TEXT

The health effects of air pollution have been widely studied and there is a large body of evidence indicating that chronic exposure to particulate matter is associated with cardiorespiratory mortality and morbidity (WHO, 2013). The International Agency for Research on Cancer (IARC) has concluded that there is sufficient evidence of the carcinogenicity of atmospheric particulate matter (PM_{10} and $PM_{2.5}$) in particular on lung cancer (Cancer, 2014). The association between air pollution and, cardiovascular and respiratory diseases has been observed in several studies (Hamra et al., 2014; Pope et al., 2004), and has been well documented by the WHO(WHO, 2013). The available literature on the sources of air pollution indicates that traffic as well as industry (combustion sources and coal-fired power plants) could be the sources mostly responsible for the adverse health effects as compared to other sources like crustal soil or sea spray (WHO, 2013). However, the literature on specific sources is limited and few studies exist on the health effect of steel industry.

The steel industry, given the specific treatment of coal at high temperature, emits several pollutants including fine particles with varying composition and containing several toxic elements, including polycyclic aromatic hydrocarbons (PAH) and transition metals (Di Gilio et al., 2017). Dioxins and others organic pollutants can originate from the combustion. Workers employed in the steel industry are exposed to several toxic substances including PAH, dioxins, and asbestos. The epidemiologic literature on the potential health effects of emissions from a steel plant is limited to the studies conducted in the Utah Valley during the eighties/nineties (Pope 3rd, 1989; Pope 3rd, Hill, & Villegas, 1999; Pope III,

Schwartz, & Ransom, 1992) with increase in mortality, respiratory hospital admissions and lung cancer. More recently, a study conducted in Canada showed that outdoor air pollution in the vicinity of a steel plant can influence cardiovascular physiology, in particular pulse rate and Flow-mediated vasodilation (FMD) (Liu et al., 2014). IARC has already classified occupational exposure in the steel industry as group 1 carcinogen (IARC, 1987) and studies of large cohorts of steel workers indicate excesses of cancers of the lung, pleura, bladder, and stomach (Andjelkovich, Mathew, Richardson, & Levine, 1990; Bourgkard et al., 2008; Finkelstein & Wilk, 1990; Fletcher & Ades, 1984; Hoshuyama et al., 2006; Park, Ahn, Stayner, Kang, & Jang, 2005; Rodríguez et al., 2000; Sorahan, Faux, & Cooke, 1994).

We have conducted a large population based cohort study with a long follow-up of mortality, hospital admissions, and cancer incidence near a large steel plant. Exposure to industrial emissions was estimated retrospectively for all the subjects. The aim of the study was to assess whether long-term exposure to industrial emissions is related to increased mortality from all-natural causes, cardiovascular and respiratory hospital admissions as well as cancer incidence. These conditions were of a priori interest because of the large literature on the health effects of fine particles but few studies nearby industries.

Materials and methods

Setting

The city of Taranto is located in the Apulia region in Southern Italy and is one of the most industrialized areas in Europe. The area has been included among the 14 Italian sites of national interest requiring remediation, and defined in the 1990s as an "area at high risk of environmental crisis" by the Italian government, because of the presence of one of the largest steel plants in Europe. Opened on November 27th, 1964, the ILVA steel plant has a nominal steel production capacity of 8 million tons/year; in addition to finished products for the domestic and foreign markets, it also supplies a wide range of Italian industries, including the automotive sector, home appliances, energy, construction, shipbuilding and packaging. The productivity of the plant changed during the period of its activity with an increase up to the eighties, a decline following the economic crisis (2009), a subsequent increase in the years 2010-2012, and a decline in 2013-2014. The production trend, and therefore the variation in emissions, had an effect on pollution levels in the neighboring districts.

Several environmental monitoring studies and measurement campaigns of industrial emissions in the Taranto area showed a relevant contribution of the steel factory on the levels of measured pollutants in the area. (Brand et al., 2004)

Study design

All residents living in Taranto (and Massafra and Statte, two nearby communities) on January 1, 1998, as well as those later entering for immigration or birth until December 31st 2010, were enrolled in the study cohort. For each subject, personal data, address, date of immigration (and possible emigration), and vital status were available from the General Registry Office of the three cities. The municipalities'

databases contained for each person all changes of residence since the fifties until the end of 2010. Each address was geocoded using ArcGIS software. The follow-up for vital status at December 31st 2013 was carried out through the municipal register and record linkage with the regional mortality database, which includes all deaths of the resident population.

A census block-level socioeconomic position (SEP) index was assigned to each participant on the basis of his/her geocoded address. SEP is a small area index (high, middle-high, medium, middle-low and low) based on the census-block (around 500 inhabitants) and built up on information recorded at the 2001 census (percentage of population with educational level equal to or less than primary school, the percentage of the active population unemployed or looking for their first job, percentage of rented houses, percentage of single parent families, population density) (G Cesaroni, Agabiti, Rosati, Forastiere, & Perucci, 2006).
For each subject, information about cause-specific mortality (1998-2013) and hospitalization (1998-2014), and cancer incidence (2006-2011) were retrieved from the Regional Health Databases and the Regional Cancer Registry. For hospitalization, only the principal diagnosis was considered.

Air pollution exposure assessment

The emission scenario of the ILVA plant is characterized by a high number of sources, both conveyed and fugitive. Fugitive emissions are the hot emissions from the steelwork, the cold emissions generated by the fossil fuels processing products, wind erosion from the large uncovered coal mining park, the handling

on the conveyors and the transportation of material. Daily simulations for the year 2010 were carried out to estimate three-dimensional concentrations of pollutants using a Lagrangian modelling system and considering both conveyed and fugitive emissions. The modelling system included the SWIFT meteorological model, the SURFPRO turbulence pre-processor and the SPRAY Lagrangian particles dispersion model (Giua et al., 2014). The SPRAY model simulates the transport, dispersion and deposition of pollutants using orography, meteorological data, turbulence and hourly spatial distribution (horizontal and vertical) of emissions, based on the characteristics of the single source and on mass fluxes (g/h). The model follows the path of fictitious particles in the atmospheric turbulent flow, and it can take into account complex situations, such as the presence of obstacles, breeze cycles, strong meteorological non-homogeneities and non-stationary, wind calm conditions. The meteorology in the area was built with the SWIFT and SURFPRO codes on hourly basis, by using the products for the year 2007 supplied by the MINNI project (Zanini, 2009) as the input. Model results were validated using measured data in 9 fixed monitoring stations of the ARPA network. Particulate Matter (PM_{10}) and Sulfur dioxide (SO_2) were considered as surrogate measures of contaminants emitted by the industrial plants. Annual average emissions were estimated for each 500m x 500m cell. The annual individual exposure to each air pollutant was obtained in a multi-step procedure that is described in the Supplemental material. The main measures of individual exposures used in the analysis were the time-varying exposures to industrial PM_{10} and SO_2 from 1965 till end of follow-up (additional details of the methodology in the Supplemental material).

Occupational status

Each cohort participant was linked to the National Pension Fund database, which includes information on individual jobs with employment companies since 1974. All people employed during the period 1974-1997 were selected by branch of activity, so previous occupational history was categorized in five groups:

- Blue collar workers in steel factories;
- White collar workers in steel factories;
- Naval construction workers;
- Mechanical construction workers;
- Workers in other occupational branches or people without contribution payments (reference category).

Statistical analyses

Each resident contributed to person-years at risk from the date of entry in the cohort (1998 or later up to 2010) until date of exit for death, emigration or end of follow-up. The association between long-term exposure to air pollutants (defined as time-varying annual average) and mortality/morbidity was estimated using a survival analysis with multivariate time-dependent Cox proportional hazard models.

Age was used as the time scale. Observation times were censored at the time of death for causes different from the one under study, emigration, loss to follow-up, or end of follow-up, whichever came first. A priori confounders included: gender,

area-level socioeconomic position occupation (recorded at baseline); calendar period (in 3 classes, time-dependant variable). Air pollution exposure was modelled using alternative time-varying variables based on different time windows, with the aim of exploring effects attributable to different averaging periods: current exposure (lag 0, e.g. average exposure in the current year), and 5-year time-window lagged concentrations over the period considered (1-5, 6-10, 11-15, 16-20, 21-25, 26-30, 31-35, 35+ years).

Hazard Ratios (HR) and corresponding 95% Confidence Intervals (95% CI) per 10 $\mu g/m^3$ increases of each pollutant were computed.

As an additional analysis, we relaxed the assumption of linearity of the concentration-response function by modelling exposure at lag 0 with a penalized spline with 2 degrees of freedom in the Cox proportional hazard models used for the main analysis.

Effects of air pollutants on mortality and cancer outcomes were analysed on the total cohort. Hospitalization outcomes were analysed both in the total cohort and in the sub-cohort of children younger than 14 years.

Confounding by smoking, alcohol and body mass index

We did not have information on individual lifestyle factors. If these variables were heterogeneously distributed across classes of exposure, a lack of adjustment could lead to residual confounding and bias in the relationship between exposure and outcome. We attempted to overcome the lack of information on smoking

habit, alcohol consumption and obesity in the cohort by using an ancillary data set of the PASSI (*Progressi delle Aziende Sanitarie per la Salute in Italia*) (Centro nazionale di epidemiologia) national surveillance survey about lifestyles and personal habits for the years 2008-2013. This sample survey is considered representative of the cohort of Taranto, Massafra and Statte. Among 1,755 subjects recruited in the survey between 2008 and 2013, 620 individuals were members of the cohort and provided questionnaire data. This subset was used to assess the relationship between exposure to industrial pollutants (independent variables in quartiles) and the prevalence of smoking habit (smokers, ex-smokers vs. never smokers), alcohol consumption (Yes/No) and obesity (BMI>30 vs BMI<=30) using a Poisson regression model (Prevalence Rate Ratios) adjusted for age, sex and socio-economic position, and weighted for the age distribution of the original cohort. Lack or presence of association between environmental exposures and individual factors were considered as supporting non-presence or presence of residual confounding from these factors.

Results

A total of 321,356 individuals (51.1% females), were enrolled in the cohort from 1998 to 2010 and followed-up until 2013. During the study period we observed 35,398 deaths for natural causes while 27,260 subjects moved away from the study area and were censored at the time of migration. The main characteristics of the cohort members (age, sex, socioeconomic position, length of residence at recruitment, occupation, and vital status) are described in Table 1.

Figure 1 shows the results of the dispersion model for PM_{10} and SO_2 originating from the industrial plant in the study area (municipalities of Taranto, Massafra and Statte). The areas in the south of the ILVA plant were mostly affected.

Productivity of the steel plant, emissions data and PM_{10} and SO_2 average exposures (lag 0), resulting from the backward and forward extrapolation procedure, are represented in the supplemental Figure S1. In the upper part of the figure, productivity and extrapolated emissions from the steel plant processes are plotted for the period 1965-2013. Emissions follow the trend of productivity until the year 1995, when they decreased, while productivity started to increase until 2008. The average exposures to PM_{10} and SO_2 at lag 0 (in the bottom part of the figure) strictly follow emissions trends and behave similarly for the two pollutants. Both productivity and emissions decreased in the year 2009, possibly due to the economic crisis, and consequently a decrease in the exposure to the pollutants was observed.

The time-varying pollutant average exposure of the past 35 years was attributed to each individual of the cohort. The annual average exposure at lag 0 at baseline (1998) was 9.03μg/m^3 (±SD=9.53) for PM_{10} and 9.09μg/m^3 (±SD=4.81) for SO_2 among 270,833 cohort members recruited at the start of the cohort (Table 2). PM_{10} and SO_2 were highly correlated (r=0.7).

The association between air pollutants and mortality is shown in Table 3. For each 10μg/m^3 increment of PM_{10} and SO_2 at lag 0, we observed an increased risk of

non-accidental mortality (HR=1.04, CI 95% 1.02-1.06, HR=1.09, CI 95% 1.05-1.12, respectively), in particular mortality from heart diseases (HR=1.05, CI 95% 1.02-1.09 and HR=1.11, CI 95% 1.04-1.18, respectively), and from acute myocardial infarction (HR=1.10, CI 95% 1.02-1.19, and HR=1.29, CI 95% 1.10-1.52 for PM_{10} and SO_2, respectively). Malignant neoplasms HR=1.08, CI 95% 1.02-1.15), and lung cancer mortality (HR=1.17, CI 95% 1.03-1.34), showed positive associations with the average concentrations of SO_2. Moreover, mortality for kidney diseases was associated with PM_{10} (HR=1.13, CI 95% 1.02-1.25 for 10μg/m^3 increase). Only weak associations were detected for respiratory diseases mortality and a negative association was found between both air pollutants and mortality from cerebrovascular diseases.

Linearity of the association between PM_{10} at lag 0 and mortality was confirmed for all the causes studied using splines (Figure 2). The exposure-response relationships between SO_2 and cause-specific mortality displayed some deviations from linearity, though with ample confidence bands at higher concentrations (Figure 3).

The latency of the effects on mortality was analyzed estimating different independent models in which one 5-year time window of exposure at the time was used. The effects estimates for these different time-windows of exposure were higher in the most proximal lags (up to 5 years), then decreased so to become almost null and then increased again for exposure occurring in the past, namely 26 years or more (Figure 4).

Table 4 shows the associations between average exposures to PM_{10} and SO_2 and hospital admissions. Both pollutants were positively associated at lag 0 with several outcomes investigated. Increased risks (ranging from 3% to 11% for $10\mu g/m^3$ increment of PM_{10} and 6-35% for $10\mu g/m^3$ increment of SO_2) were found for natural causes, central nervous system, heart, respiratory, and kidney diseases. In addition, SO_2 was associated also with hospital admissions for acute myocardial infarction (HR=1.14, CI 95% 1.06-1.23) and heart failure (HR=1.13, CI 95% 1.06-1.21).

We found positive associations with pediatrics admissions for diseases of the respiratory system (HR=1.11 and HR=1.33, for $10\mu g/m^3$ increases in PM_{10} and SO_2 at lag 0, respectively) and for respiratory infections (HR=1.15 for PM_{10} and HR=1.49 for SO_2). (Table 4, bottom).

We studied the shape of the relationship of hospital admissions with pollutants by estimating penalized splines. The Figure S2 shows the linear trend in the effects of PM_{10} and SO_2 on admissions for most of the conditions.

Table 5 shows the association between industrial pollutants and cancer incidence. We found a positive association with lung cancer incidence (HR=1.29, 95%CI 1.14-1.45, and HR=1.42, 95%CI 1.10-1.84, for 10 μg/m3 increases in PM_{10} and SO_2, respectively) and incidence of kidney cancer (HR=1.32, 95%CI 1.01-1.73 for PM_{10} e HR=2.44, 95%CI 1.38-4.34 for SO_2). Among women, PM_{10} exposure was associated with breast cancer (HR=1.27, 95% CI 1.13-1.41).

Residual confounding

Table S1 in the Supplemental material shows the prevalence of potential confounders (smoking, alcohol, and obesity) by quartiles of the distribution of industrial PM_{10} and SO_2 exposure in the ancillary data set we used. Asjusted prevalence rate ratios are also reported. There is no clear increasing trend of the prevalence of smoking, drinking alcohol and obesity with higher quartiles of exposure. No statistically significant association was found between exposures and individual smoking habits and obesity suggesting that they are not relevant confounders in the present study. Alcohol consumption was related to the 2^{nd} and 4^{th} quartile of PM_{10} only.

Discussion

This study evaluated the effects of environmental exposures on the health status of the residents in the area of Taranto. Results showed associations between exposure to PM_{10} and SO_2 of industrial origin and natural mortality and for specific causes, in particular non accidental mortality, lung cancer, heart disease, acute myocardial infarction and kidney disease. The study showed also higher risks in the first years of exposure and in the far past of more than 25 years before the current exposure. Associations were even stronger considering hospital admissions. The analysis of cancer incidence highlighted a positive relationship between the industrial pollutants and lung and kidney cancer in both genders and breast cancer among women.

We considered modeled PM_{10} and SO_2 concentrations as exposure measures of industrial pollutants on the assumption that the pollution from the plant does not spread uniformly around the site but depends on emissions, prevailing winds and the orography of the area. We used the shape of the concentrations on the ground to rank subjects as more or less exposed, and this shape is of greater importance than are the exact absolute values. In our study SO_2 exposure effects are higher than those of PM_{10} and they are similar in magnitude to the excess risks observed among residents in Paolo VI district by Mataloni et al. (Mataloni et al., 2012). A recent study by Mangia et al. (Mangia et al., 2013) revealed that the estimated SO_2 *footprint* identifies very well the area close to the steel factory, exhibiting higher mean values and positive correlations with wind speed, when the monitoring station is downwind from the industrial site. The monitoring station

located in Paolo VI district recorded the highest SO_2 mean concentration values, compared to the other neighbourhoods, suggesting that SO_2 could be a stronger marker of industrial exposure than PM_{10}.

In this study exposures to PM_{10} and SO_2 from industrial sources were modeled using the annual mean exposure at residence, considering all changes of residence until the end of 2010. Many cohort studies assumed exposure contrasts, estimated at individual level at the study inception (baseline), as representative for long term exposures (Ancona et al., 2015; Beelen et al., 2008; Filleul et al., 2005; Jerrett et al., 2009; Mataloni et al., 2012). This approach can lead to exposure misclassification in case of widespread mobility patterns or changing spatial distribution of exposure over time. Few longitudinal studies (Bentayeb et al., 2015; Lepeule et al., 2012; Tétreault et al., 2016; Wahida et al., 2016) dealt with time-varying exposure assessment for every year of follow-up, and among them only two studies attempted backward reconstruction of individual residential histories (Bentayeb et al., 2015; Wahida et al., 2016). The study conducted in Paris took into account residential mobility by weighting the individual cumulative exposure at census block level by the probability of moving from a census block to another, on a specific year, tabled in a matrix of all possible movements(Wahida et al., 2016). However, in the two studies, aggregation of exposure assessment was at the zip code and census block level, respectively. In both studies (Bentayeb et al., 2015; Wahida et al., 2016), an average cumulative exposure measure was estimated, which in one case (Bentayeb et al., 2015) lead to similar or weaker results compared to the mean annual exposure.

Hazard ratios (HRs) were estimated considering as potential confounders birthplace, SEP, occupational exposure and other environmental exposures besides that under study. We explored the form of the exposure-response relationship between industrial exposures and outcomes by using a semi-parametric smoothed curve, the natural spline (Madouasse, Browne, Huxley, Toni, & Green, 2012).

The reconstruction of the personal exposure history from the moment people became residents in the area, also before the beginning of their follow-up, considering all changes in addresses and migration movements to account for past cumulative exposure, is new in the long-term effects studies context. The level of exposure was assessed on a very fine scale and the exposure estimate came from a well-reasoned procedure, where dispersion estimates were weighted with production and emissions data of the steel plant. All events occurring in decreasing or increasing personal residential exposure, like the economic crisis in 2009, the closure of some comparts after the trial in 2012, and all the personal movement of residence were then evaluated over the previous 35 years of observation. In Table 1 we saw that more than 34% of subjects were residing in the area for more than 30 years, showing that the population is stable, with a percentage of 81-85% of cohort subjects who never changed address during the follow-up. We assumed that the stable population of the cohort has different lifestyle and habits that we took into account correcting the estimates with socioeconomic position index.

Increasing yearly mean exposure in IQR of PM_{10}, $PM_{10-2.5}$ and SO_2 was significantly associated with increase in non-accidental mortality of 12% for PM_{10}, 6% for $PM_{10-2.5}$ and 10% for SO_2. Our study found generally similar effect estimates between time-varying and cumulative exposure, and the analysis of the latency showed that risks of natural mortality were strongest in the first year of exposure and after more than 25 years of permanence in the area, while in the Bentayeb et al.'s study highest effects were found after one year of exposure. However the high correlation (r=0.7) among lags cannot allow us to think in terms of independent effects. Thus the effect seen after 25 years of exposure may depend on the exposure of the previous years.

A more accurate lag- distributed model should be applied, like Lepeule et al. did in six Harvard cities from 1974 to 2009, to inspect the behaviour and trend of time of exposure on the onset of diseases. It is in fact also of interest to explore whether the concentration-response curve has changed over time as particle composition and anthropogenic activities have changed over the 48-years period considered in our study.

One of the strength points in our study was the availability and completeness of mortality and hospital admissions data, which for the latter arrived until the end of follow up. Record linkage procedures attributed almost 98% of the causes of death to individuals and rigorous protocols were adopted to select appropriate hospital admissions.

Another strength of this work is the cohort size (321,356 residents) and the longitudinal study design adopted. Our results were adjusted for several confounders: age, socio-economic position, and variables related to the environmental and occupational context that might otherwise have confounded our results. However, no data were available on the personal habits of the subjects that could have had a role in the diseases investigated, especially cigarette smoking, but also alcohol use, physical activity and obesity. Collecting this information, through telephone interviews or home visits, would have been prohibitive for such a large cohort. The lack of this information may have biased the results because of confounding not controlled in the analysis. However, at least for smoking, alcohol consumption and obesity, the estimates on a cohort sub-group with this information available indicated an absent or negligible association between exposure to PM_{10} and SO_2 and those individual habits. The distribution of smoking addiction among the cohort revealed 31.5% of smokers, compared to a percentage of 27-28% in the Apulia region and 28% in Italy globally. It is therefore extremely unlikely that smoking plays a confounding role in this cohort. Furthermore, it should be noted that many personal habits are associated with socio-economic position. It is thus reasonable to assume that the analysis that adjusted for socio-economic index also took into account others, not measured, individual variables.

The novelty of our study is the individual exposure up to 48 years before the end of follow-up in a multi steps and time varying approach. We were able then to ascertain the residential history of the cohort and to assess the temporal variation

of exposure to PM_{10} and SO_2. However, individual exposures came from models and were not measured directly through the follow-up. This assumes that subjects would not move and stay at their home all day. Since a large proportion of subjects works and spends time at work, this may have introduced a misclassification bias in our study. Another possible source of bias could lead in using a multi steps procedure that, first of all, started from results of a dispersion model estimated from parameters of the year 2010 and meteorology of 2007, and that introduced possible errors in the extrapolation phase, mainly in the one from 1990 backwards, where a constant annual increase of emissions was assumed.

In conclusion, industrial PM_{10} and SO_2 were positively associated with mortality and hospitalizations in the Taranto area, when socioeconomic position and occupational exposures were taken into account, confirming that the industrial pole is an important risk factor for the health status of residents. Both recent and past environmental exposures are responsible for the negative effects observed.

		N	%
Cohort		321,356	100
Gender	Males	157,031	48.9
	Females	164,325	51.1
Age class at enrolment (years)	0	34597	10.8
	1-9	29047	9.0
	10-19	36,224	11.3
	20-29	49,652	15.5
	30-39	45,674	14.2
	40-49	37,811	11.8
	50-59	34,213	10.7
	60-69	26,946	8.4
	70-79	18,502	5.8
	>=80	8,690	2.7
Area based socio-economic position	High	68,693	21.4
	Middle-High	39,095	12.2
	Medium	32,736	10.2
	Middle-Low	58,034	18.1
	Low	112,481	35.0
	Missing	10,317	3.2
Lenght of residence in the area at the enrolment (years)	0-10	105768	32.9
	11-19	46831	14.6
	21-30	50756	15.8
	31-40	111203	34.6
Occupational status			
Employment at the Iron and Steel Industry[a]	No	307,800	95.8
	Yes, laborer	9,633	3.0
	Yes, office worker	3,923	1.2
Employment in the Mechanical Construction Industry[a]	No	304,321	94.7
	Yes	17,035	5.3
Employment in the Naval Construction Industry[a]	No	320,118	99.6
	Yes	1,238	0.4
Vital Status			
	Alive	258,698	80.5
	Dead	35,398	11.0
	Lost to follow-up	27,260	8.5

[a]1974-1997

Table 1: Descriptive characteristics of the cohort members, study period 1998-2013.

Pollutant	Mean	SD	Min	Max	Percentiles				
					5°	25°	50°	75°	95°
PM_{10} (µg/m3)	9.03	9.53	0	85.24	1.05	3.52	7.79	9.35	30.60
SO_2 (µg/m3)	9.09	4.81	0	22.08	1.88	5.53	9.27	11.82	18.18

Table 2: Descriptive data of exposures to PM_{10} and SO_2 of industrial origin at lag 0 among 270,833 cohort members at the baseline in 1998.

Causes of death (ICD-9CM)		PM_{10}		SO_2	
	N	HR*	95%CI	HR*	95%CI
Natural mortality (001-799)	33042	1.04	1.02-1.06	1.09	1.05-1.12
Malignant neoplasms (140-208)	10210	1.03	1.00-1.06	1.08	1.02-1.15
Trachea, bronchus, and lung (162)	2164	1.05	0.99-1.12	1.17	1.03-1.34
Bladder (188)	476	1.03	0.90-1.18	0.98	0.74-1.29
Kidney (189)	116	0.95	0.70-1.30	0.81	0.46-1.45
Lymphatic and hematopoietic tissue (200-208)	879	0.98	0.87-1.09	1.04	0.85-1.28
Diseases of the central nervous system (330-349)	1014	1.05	0.951.16	1.05	0.86-1.29
Diseases of the circulatory system (390-459)	12527	1.02	1.00-1.05	1.04	0.99-1.10
Heart diseases (390-429)	8857	1.05	1.02-1.09	1.11	1.04-1.18
Acute myocardial infarction (410-411)	1275	1.10	1.02-1.19	1.29	1.10-1.52
Cerebrovascular disease (430-438)	2903	0.90	0.85-0.96	0.80	0.72-0.89
Diseases of the respiratory system (460-519)	2741	1.02	0.97-1.08	1.02	0.91-1.14
Respiratory infections (460-466, 480-487)	751	0.90	0.80-1.02	0.85	0.69-1.04
COPD (490-492, 494, 496)	1618	1.03	0.95-1.10	1.04	0.90-1.21
Kidney disease (580-599)	707	1.13	1.02-1.25	1.16	0.93-1.45

*Hazard Ratio (HR) from a Cox model stratified for period of follow-up (3 categories) and sex, adjusted for age (temporal axis), socioeconomic position and occupational status

Table 3: Associations between annual average exposure to PM_{10} and SO_2 and cause-specific mortality. Adjusted hazard ratios (HRs and 95% CI) per 10 µg/m^3 increase of each pollutant, 1998-2013.

Diagnosis (ICD-9 CM)		PM_{10}		SO_2	
	N	HR	95%CI	HR	95%CI
All natural causes (001-799)[a]	193277	1.03	1.02-1.04	1.06	1.04-1.07
Diseases of the central nervous system (330-349)	8890	1.05	1.01-1.08	1.21	1.13-1.30
Diseases of the circulatory system (390-459)	49859	1.04	1.02-1.05	1.06	1.03-1.09
Heart diseases (390-429)	34316	1.05	1.04-1.07	1.10	1.07-1.14
Acute myocardial infarction (410-411)	7253	1.02	0.99-1.06	1.14	1.06-1.23
Heart failure (428)	8952	1.02	0.99-1.06	1.13	1.06-1.21
Cerebrovascular disease (430-438)	13236	1.01	0.98-1.04	0.92	0.87-0.97
Diseases of the respiratory system (460-519)	31091	1.07	1.05-1.08	1.15	1.12-1.19
Respiratory infections (460-466, 480-487)	13654	1.11	1.08-1.13	1.35	1.28-1.42
COPD (490-492, 494, 496)	7474	1.03	1.00-1.06	0.95	0.88-1.01
Asthma (493)	885	0.99	0.90-1.09	0.95	0.78-1.16
Kidney disease (580-599)	13184	1.08	1.05-1.11	1.09	1.04-1.15
Population 0-14 years[b]					
Diseases of the respiratory system (460-519)	9505	1.11	1.09-1.14	1.33	1.26-1.41
Respiratory infections (460-466, 480-487)	6746	1.15	1.11-1.18	1.49	1.39-1.59
Asthma (493)	272	0.77	0.60-0.98	0.55	0.37-0.81

[a]Hazard Ratio (HR) from a Cox model stratified for period of follow-up (3 categories) and sex, adjusted for age (temporal axis), socioeconomic position and occupational status

[b]Hazard Ratio (HR) from a Cox model stratified for period of follow-up (3 categories) and sex, adjusted for age (temporal axis), socioeconomic position

Table 4: Associations between annual average exposure to PM_{10} and SO_2 and hospital admissions. Adjusted hazard ratios (HRs and 95% CI) per 10 µg/m^3 increase of each pollutant, 1998-2014.

Site of cancer	N	PM_{10}		SO_2	
		HR*	95%CI	HR*	95% IC
All site (ICDO3T C00-C809)	8999	1.14	1.09-1.19	1.05	0.97-1.14
Upper respiratory and digestive tract (ICDO3T C00-	144	0.80	0.52-1.23	0.67	0.34-1.31
Oesophagus(ICDO3T C15)	27	0.30	0.06-1.48	0.20	0.04-1.08
Stomach (ICDO3T C16)	284	0.99	0.77-1.28	0.69	0.43-1.11
Colon-rectum and anus (ICDO3T C18-C21)	887	1.11	0.96-1.28	1.00	0.77-1.31
Liver(ICDO3T C22)	340	1.10	0.89-1.37	0.75	0.48-1.15
Gallbladder and biliary tract (ICDO3T C23-C24)	117	1.14	0.80-1.64	0.88	0.41-1.85
Pancreas (ICDO3T C25)	208	1.19	0.90-158	1.19	0.68-2.08
Larynx (ICDO3T C32)	91	1.39	0.99-1.96	1.39	0.62-3.13
Lung incl. trachea and bronchus (ICDO3T C33-C34)	943	1.29	1.14-1.45	1.42	1.10-1.84
Pleural cancer (ICDO3T C384)	89	0.96	0.61-1.52	1.15	0.50-2.64
Bone and cartilage (ICDO3T C40-C41)	22	0.59	0.16-2.22	0.53	0.09-2.96
Malignant melanoma of the skin (ICDO3T C44)	1944	1.15	1.04-1.26	1.08	0.90-1.30
Peripheral nerves, connective and soft tissue	40	1.22	0.66-2.27	0.62	0.17-2.26
Breast (ICDO3T C50)	1137	1.27	1.13-1.41	1.19	0.94-1.51
Prostate (ICDO3T C61)	653	1.09	0.92-1.29	1.06	0.77-1.45
Testis (ICDO3T C62)	42	1.08	0.58-2.01	0.96	0.30-3.11
Kidney (ICDO3T C64)	173	1.32	1.01-1.73	2.44	1.38-4.34
Renal pelvis and urinary organs (ICDO3T C65-C66,	34	0.87	0.34-2.23	0.56	0.13-2.46
Bladder (ICDO3T C67)	415	1.07	0.88-1.32	0.91	0.61-1.35
Brain and central nervous system (ICDO3T C69-	117	1.23	0.87-1.72	0.87	0.42-1.82
Thyroid (ICDO3T C73-C75)	365	0.97	0.75-1.25	0.76	0.49-1.17
Mesothelioma (ICDO3M 9050-9055)	72	0.96	0.57-1.60	0.93	0.36-2.37
Sarcoma Kaposi (ICDO3M 9140)	38	1.35	0.77-2.37	1.39	0.41-4.64
Hodgkin limphoma (ICDO3M 9650-9667)	52	0.98	0.54-1.78	1.56	0.54-4.50
Non-Hodgkin lymphoma (ICDO3M 9590-9596)	31	0.93	0.41-2.11	0.74	0.18-3.06
Multiple Myeloma (ICDO3M 9732)	98	0.91	0.56-1.46	0.76	0.34-1.69
Leukaemias (ICDO3T 9421, ICDO3M 9800-9948)	184	1.11	0.82-1.51	1.21	0.68-2.15

[a]Hazard Ratio (HR) from a Cox model stratified for period of follow-up (2 categories) and sex, adjusted for age (temporal axis), socioeconomic position and occupational status

Table 5: Associations between annual average exposure to PM_{10} and SO_2 and cancer incidence. Adjusted hazard ratios (HRs and 95% CI) per 10 $\mu g/m^3$ increase of each pollutant, 2006-2011.

Figure 1: Study area and pollutants dispersion model (ARPA, year 2010), with geocoded addresses at baseline.

Figure 2: Penalized splines and confidence interval (95%CI) of the relationship between average PM_{10} exposure at lag 0 and mortality for natural mortality, mortality for malignant neoplasms, lung cancer, heart diseases, acute myocardial infarction and kidney diseases.

Figure 3: Penalized splines and confidence interval (95%CI) of the relationship between average SO_2 exposure at lag 0 and mortality for natural mortality, mortality for malignant neoplasms, lung cancer, heart diseases, acute myocardial infarction and kidney diseases.

Figure 4: Distribution of the effects of PM_{10} and SO_2 in 5-years' time windows on natural mortality. Results expressed as percent increase for $10\mu g/m^3$ increment.

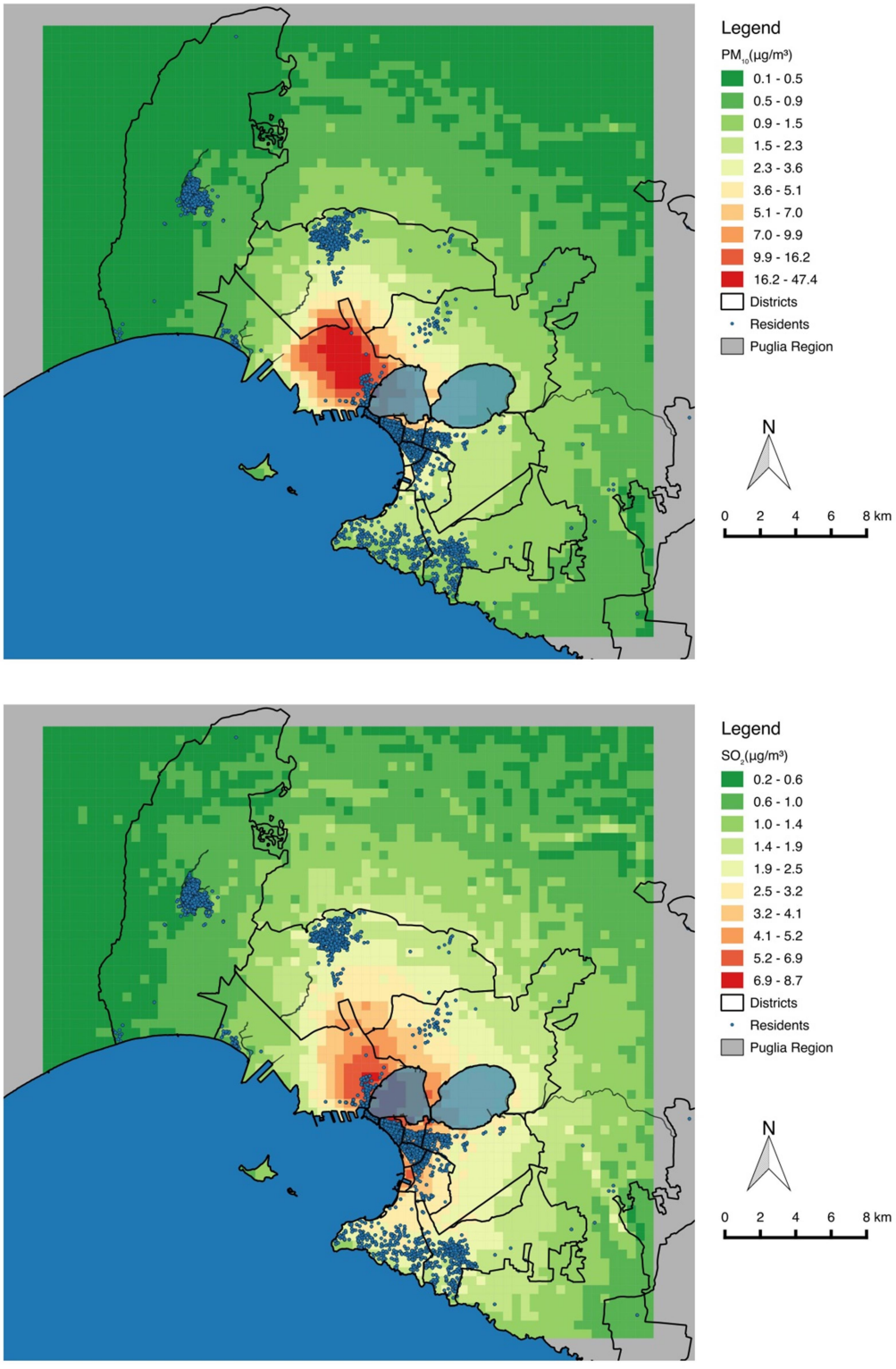
Legend
PM10(µg/m³)
0.1 - 0.5
0.5 - 0.9
0.9 - 1.5
1.5 - 2.3
2.3 - 3.6
3.6 - 5.1
5.1 - 7.0
7.0 - 9.9
9.9 - 16.2
16.2 - 47.4
Districts
Residents
Puglia Region
N
0 2 4 6 8 km
Legend
SO2(µg/m³)
0.2 - 0.6
0.6 - 1.0
1.0 - 1.4
1.4 - 1.9
1.9 - 2.5
2.5 - 3.2
3.2 - 4.1
4.1 - 5.2
5.2 - 6.9
6.9 - 8.7
Districts
Residents
Puglia Region
N
0 2 4 6 8 km

Figure 1

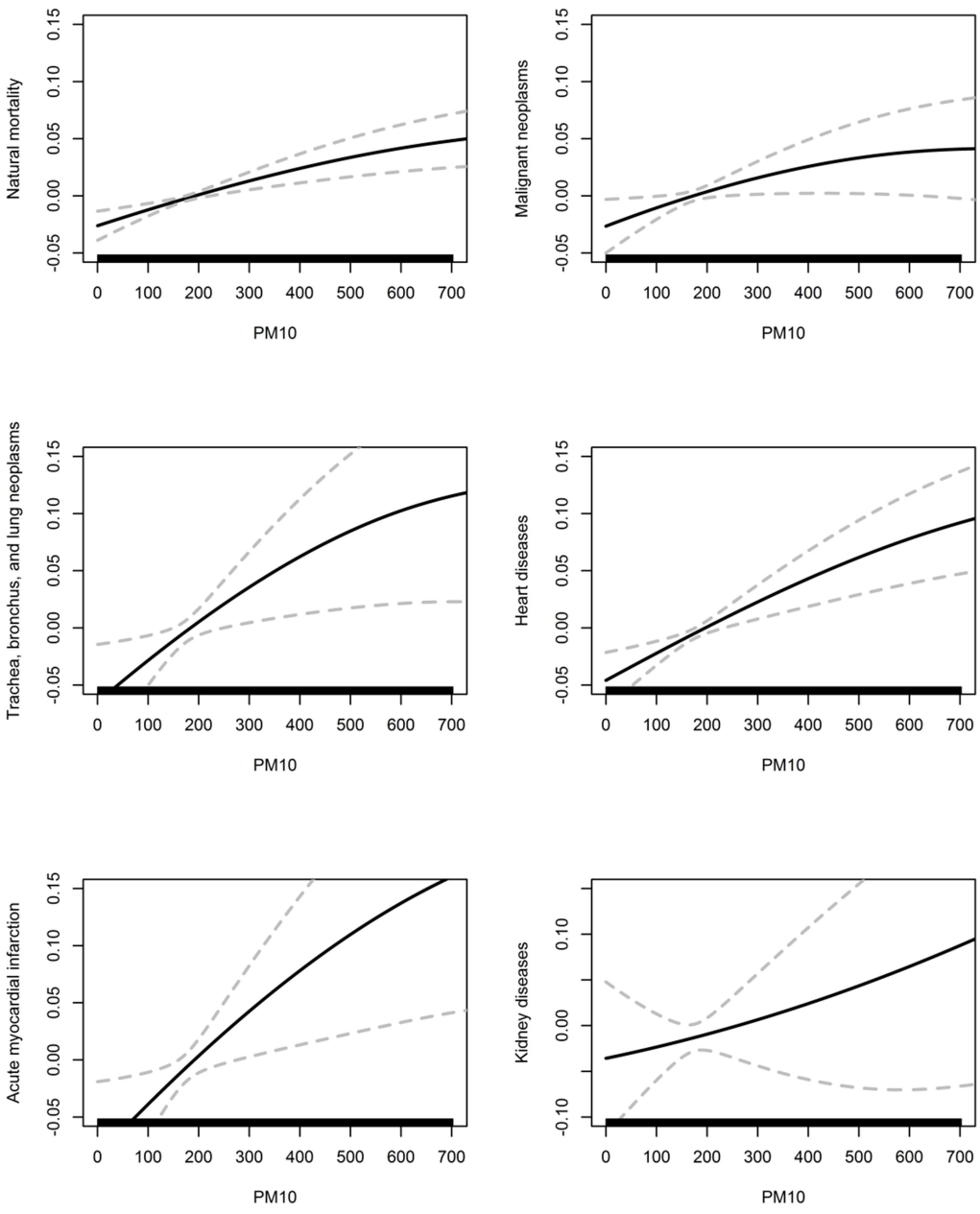
Natural mortality
Malignant neoplasms
Trachea, bronchus, and lung neoplasms
Heart diseases
Acute myocardial infarction
Kidney diseases
PM10
0
100
200
300
400
500
600
700
-0.05
0.00
0.05
0.10
0.15
-0.10

Figure 2

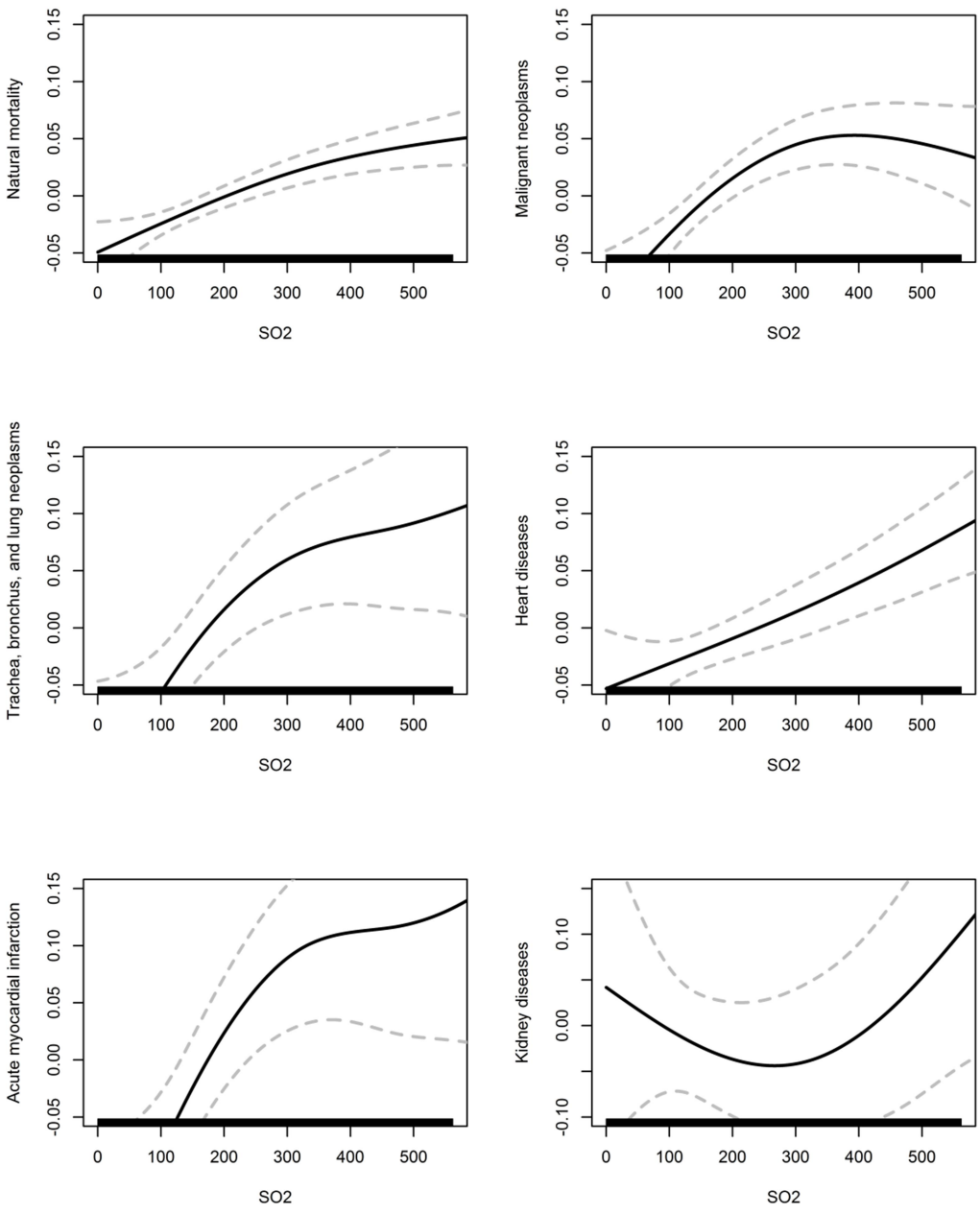
Natural mortality
Malignant neoplasms
Trachea, bronchus, and lung neoplasms
Heart diseases
Acute myocardial infarction
Kidney diseases
SO2
0
100
200
300
400
500
-0.05
0.00
0.05
0.10
0.15
-0.10

Figure 3

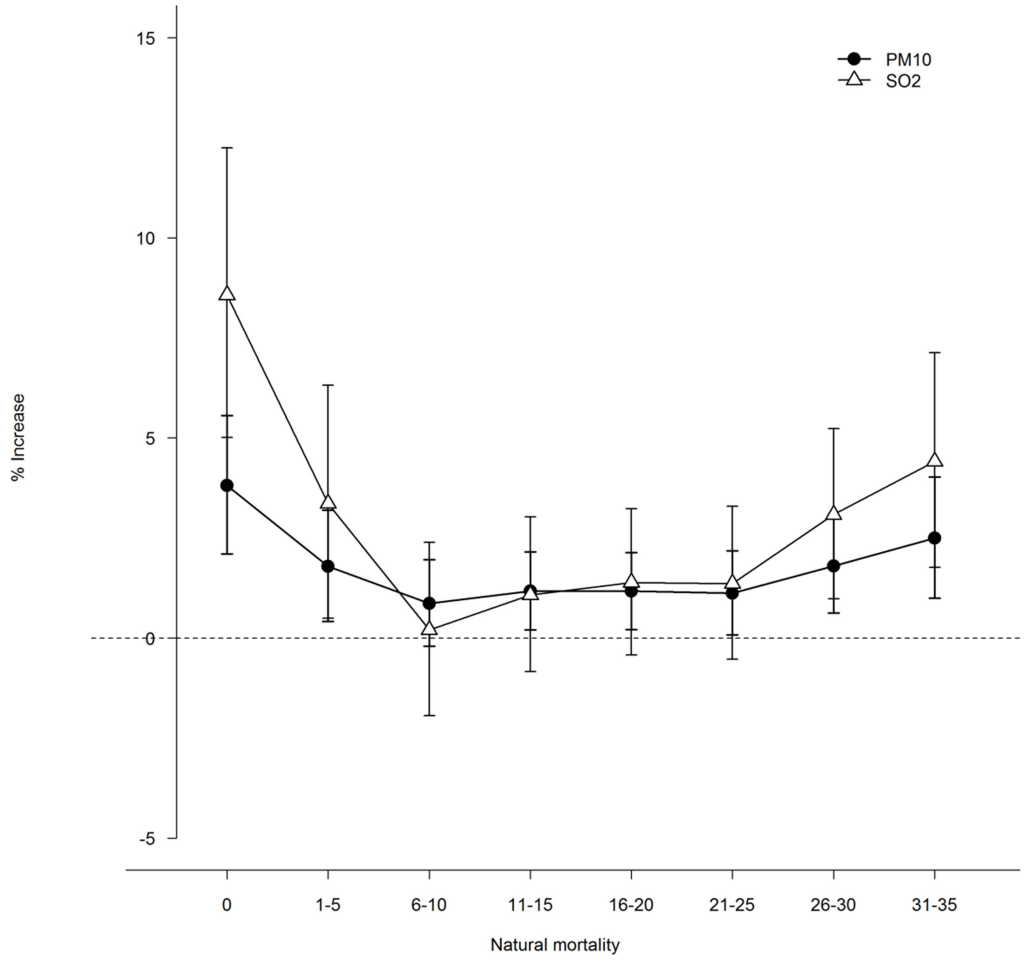
15
10
5
0
-5
% Increase
PM10
SO2
0
1-5
6-10
11-15
16-20
21-25
26-30
31-35
Natural mortality

Figure 4

REFERENCES

1. WHO, Final technical report. 2013.

2. Cancer, I.A.f.R.o., IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Outdoor Air Pollution. Lyon, France, 2014. 109.

3. Hamra, G.B., et al., Outdoor particulate matter exposure and lung cancer: a systematic review and meta-analysis. Environmental health perspectives, 2014. 122(9): p. 906.

4. Pope, C.A., et al., Cardiovascular mortality and long-term exposure to particulate air pollution. Circulation, 2004. 109(1): p. 71-77.

5. Di Gilio, A., et al., An intensive monitoring campaign of PAHs for assessing the impact of a steel plant. Chemosphere, 2017. 168: p. 171-182.

6. Pope 3rd, C., Respiratory disease associated with community air pollution and a steel mill, Utah Valley. American Journal of Public Health, 1989. 79(5): p. 623-628.

7. Pope III, C.A., J. Schwartz, and M.R. Ransom, Daily mortality and PM10 pollution in Utah Valley. Archives of Environmental Health: An International Journal, 1992. 47(3): p. 211-217.

8. Pope 3rd, C., R.W. Hill, and G.M. Villegas, Particulate air pollution and daily mortality on Utah's Wasatch Front. Environmental Health Perspectives, 1999. 107(7): p. 567.

9. Liu, L., et al., Exposure to air pollution near a steel plant and effects on cardiovascular physiology: a randomized crossover study. International journal of hygiene and environmental health, 2014. 217(2): p. 279-286.

10. IARC, monographs on the Evaluation of Carcinogenic Risks to Humans. Overall Evaluation of Carcinogenicity: An Updating of IARC Monographs. 1987. 1-42, Supplement7: p. 224-25.

11. Andjelkovich, D.A., et al., Mortality of iron foundry workers: I. Overall findings. Journal of Occupational and Environmental Medicine, 1990. 32(6): p. 529-540.

12. Hoshuyama, T., et al., Mortality of iron-steel workers in Anshanl China: A retrospective cohort study. International journal of occupational and environmental health, 2006. 12(3): p. 193-202.

13. Finkelstein, M.M. and N. Wilk, Investigation of a lung cancer cluster in the melt shop of an Ontario steel producer. American journal of industrial medicine, 1990. 17(4): p. 483-491.

14. Fletcher, A.C. and A. Ades, Lung cancer mortality in a cohort of English foundry workers. Scandinavian journal of work, environment & health, 1984: p. 7-16.

15. Sorahan, T., A.M. Faux, and M.A. Cooke, Mortality among a cohort of United Kingdom steel foundry workers with special reference to cancers of the stomach and lung, 1946-90. Occupational and environmental medicine, 1994. 51(5): p. 316-322.

16. Park, R.M., et al., Mortality of iron and steel workers in Korea. American journal of industrial medicine, 2005. 48(3): p. 194-204.

17. Bourgkard, E., et al., Lung cancer mortality and iron oxide exposure in a French steel-producing factory. Occupational and environmental medicine, 2008.

18. Rodríguez, V., et al., Lung cancer risk in iron and steel foundry workers: a nested case control study in Asturias, Spain. American journal of industrial medicine, 2000. 38(6): p. 644-650.

19. Brand, R., et al., EPER European Pollutant Emission Register. European Commission, 2004.

20. Cesaroni, G., et al., An index of socioeconomic position based on 2001 Census, Rome. Epidemiologia e prevenzione, 2006. 30(6): p. 352-357.

21. Giua, R., et al., Application of a Lagrangian particle model to the source apportionment for primary macropollutants in Taranto area (South Italy). International Journal of Environment and Pollution, 2014. 55(1-4): p. 31-40.

22. Zanini, G., Il sistema MINNI, modello integrato nazionale per la valutazione degli effetti dell'inquinamento atmosferico e dell'efficacia delle politiche di riduzione delle emissioni di inquinanti atmosferici. Epidemiologia e prevenzione, 2009. 33(6): p. 35-42.

23. Centro nazionale di epidemiologia, s.e.p.d.s.d.I.s.d.s., La sorveglianza PASSI. Il portale dell'epidemiologia per la sanità pubblica.

24. Mataloni, F., et al., Studio di coorte sulla mortalità e morbosità nell'area di Taranto. Epidemiol Prev, 2012. 36(5): p. 237-252.

25. Mangia, C., et al., Spatial variability of air pollutants in the city of Taranto, Italy and its potential impact on exposure assessment. Environmental monitoring and assessment, 2013. 185(2): p. 1719-1735.

26. Beelen, R., et al., Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study). Environmental health perspectives, 2008. 116(2): p. 196.

27. Jerrett, M., et al., A cohort study of traffic-related air pollution and mortality in Toronto, Ontario, Canada. Environmental health perspectives, 2009. 117(5): p. 772.

28. Ancona, C., et al., Mortality and morbidity in a population exposed to multiple sources of air pollution: A retrospective cohort study using air dispersion models. Environmental research, 2015. 137: p. 467-474.

29. Filleul, L., et al., Twenty five year mortality and air pollution: results from the French PAARC survey. Occupational and environmental medicine, 2005. 62(7): p. 453-460.

30. Lepeule, J., et al., Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009. Environmental health perspectives, 2012. 120(7): p. 965.

31. Bentayeb, M., et al., Association between long-term exposure to air pollution and mortality in France: A 25-year follow-up study. Environment international, 2015. 85: p. 5-14.

32. Tétreault, L.-F., et al., Childhood exposure to ambient air pollutants and the onset of asthma: An administrative cohort study in Quebec. Environmental health perspectives, 2016. 124(8): p. 1276.

33. Wahida, K.-T., et al., A Conceptual Framework for the Assessment of Cumulative Exposure to Air Pollution at a Fine Spatial Scale. International journal of environmental research and public health, 2016. 13(3): p. 319.

34. Madouasse, A., et al., A semi-parametric model for lactation curves: Development and application. Preventive veterinary medicine, 2012. 105(1): p. 38-48.

SUPPLEMENTAL MATERIAL

PART 1: backward and forward extrapolation

5) In the first step, we superimposed the PM_{10} and SO_2 concentration maps with cohort geocoded addresses, to assign to each residence the corresponding estimated industrial concentration of air pollutants, with reference to the year 2010.
6) The ILVA steel plant provided productivity data per kiloton year (kton/a) of steel, coke, cast iron and others from 1965 to 2014. In addition, air pollutants emissions for every process of the industry (coke ovens, sintering plant, blast furnace, steel plant, mining parks and transport of materials) were supplied by the Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA) for the years 1990, 1995, 2000 and 2005. For the purposes of this study, only emissions from the steel production process were considered. Emissions for missing years were estimated by backward and forward interpolation of the emission series from 1965 to 2010, weighted with the productivity ratio of the current year to the previous/following one:
    - *1965-1990*: the emissions in year *x* was obtained retrospectively, starting from the first known value in the 1990, weighting with the ratio of the productivity in the current year to the following one (1)

$$E(x) = E(x+1) * \frac{P(x)}{P(x+1)} \qquad (1)$$

      where x indicates the year, E is the estimated emission and P is the productivity

- *1991-2010:* the emission in the year *x* was obtained through the interpolation of the emissions in the time intervals, always of length 5 years, between two observed knowns values (superior and inferior extreme of the interval, for example 2005 and 2010), weighting them with the relative productivity of the previous year (2)

$$E(x) = E(x+1) + \frac{E(sup) - E(inf)}{5} * \frac{P(x)}{P(x-1)} \quad (2)$$

where *E(sup)* and *E(inf)* represent the emissions at the two extremes of the time interval of 5 year length.

- *2010-2014:* the emission in the year x was estimated starting form the last value provided in 2010, weighting prospectively with the relative productivity of the previous year (3)

$$E(x) = E(x-1) * \frac{P(x)}{P(x-1)} \quad (3)$$

7) Once the annual emission series was complete for the entire period 1965-2014, we computed an annual calibration factor (4) as the ratio of emissions on one year to the emissions in the year 2010, the same of the ARPA dispersion model

$$F_x = \frac{\hat{E}_x}{\hat{E}_{2010}} \quad (4)$$

where *x* is the year and $\hat{E}$ is the estimated emission.

This factor was computed to modulate the exposure estimated from the dispersion model with the information acquired on the effective productivity and emissions from the steel plant.

8) The yearly calibration factor was then multiplied to the individual exposure from the dispersion model for every residence and every year. In this way we obtained

an annual time-varying exposure from the industry based on the spatial pattern of the dispersion model and the temporal pattern estimated by the annual series of the productivity and the emissions of the steel plant.

	PM_{10}				SO_2			
Risk factors	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**
Smoking								
Never smokers (%)	46.75	51.35	50.3	40.88	46.05	50.99	51.92	40.69
Smokers (%)	31.82	27.03	30.3	37.23	32.24	27.15	30.13	36.55
Ex-smokers (%)	21.43	21.62	19.39	21.9	21.71	21.85	17.95	22.76
Alcohol consumption (%)	52.6	66.9	52.12	62.77	54.61	62.91	55.77	60
Obesity (BMI>=30kg/m^2) (%)	16.88	12.16	10.3	16.06	15.79	13.25	11.54	14.48
Adjusted models[a] -RR(95%CI)								
Smokers [b]	1.00	0.87 (0.62-1.22)	0.97 (0.72-1.32)	1.24 (0.92-1.68)	1.00	0.88 (0.63-1.24)	0.94 (0.69-1.27)	1.24 (0.92-1.66)
Ex-smokers [b]	1.00	0.97 (0.64-1.46)	0.85 (0.57-1.27)	0.98 (0.65-1.47)	1.00	0.90 (0.60-1.37)	0.74 (0.49-1.12)	0.95 (0.64-1.41)
Alcohol consumption [c]	1.00	1.32 (1.09-1.59)	1.02 (0.82-1.26)	1.25 (1.02-1.52)	1.00	1.20 (0.99-1.46)	1.06 (0.86-1.30)	1.15 (0.95-1.40)
Obesity [d]	1.00	0.66 (0.37-1.17)	0.59 (0.33-1.04)	0.87 (0.51-1.50)	1.00	0.76 (0.43-1.34)	0.69 (0.39-1.23)	0.82 (0.47-1.45)

[a] Separated Poisson models adjusted for age, gender and socioeconomic status index.
[b] Reference category "non-smokers"; [c] reference category "No alcohol consumption"; [d] reference category "Not obese (BMI<30kg/m^2)"
Abbreviations: Qx= quartile; BMI=Body Mass Index; RR= Relative Risks; CI=Confidence Interval

Table S1: Association between smoking, alcohol consumption, obesity and quartiles of PM_{10} and SO_2 exposures.

Figure S1: Temporal trend of the steel productivity, emissions and pollutants concentrations at lag 0, study period 1965-2014.

Figure S2: Penalized splines and confidence interval (95%CI) of the relationship between average PM_{10} and SO_2 exposure at lag 0, and hospital admissions for neurological disorders, heart, respiratory and kidney diseases.

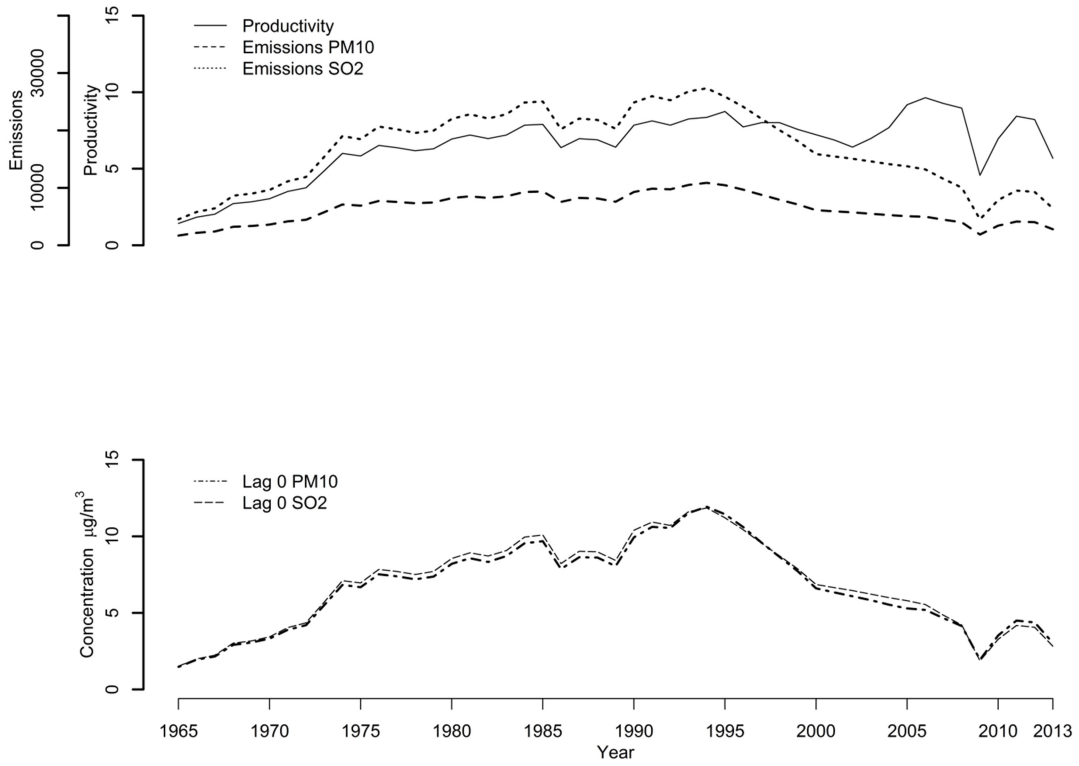

Figure S1: Temporal trend of the steel productivity, emissions and pollutants concentrations at lag 0, study period 1965-2014.

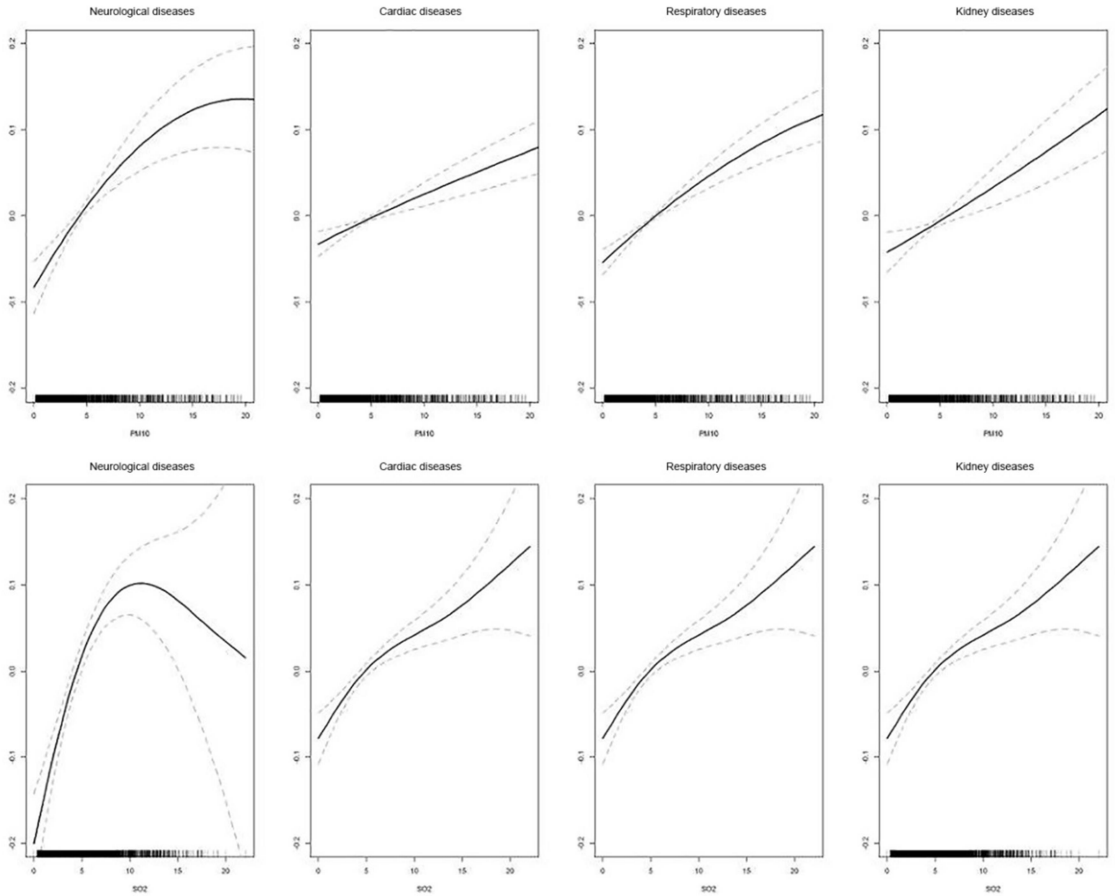

Figure S2: Penalized splines and confidence interval (95%CI) of the relationship between average PM_{10} and SO_2 exposure at lag 0, and hospital admissions for neurological disorders, heart, respiratory and kidney diseases.

Annex B- Paper on DID methodology

Title: Estimation of Causal Effects of Industrial Pollution on Mortality in the Taranto Area, Southern Italy

Authors: Simona Leogrande, Ester Rita Alessandrini, Massimo Stafoggia, Angela Morabito, Carla Ancona, Lucia Bisceglia, Roberto Giua, Antonia Mincuzzi, Sante Minerba, Stefano Spagnolo, Giorgio Assennato and Francesco Forastiere, on behalf of the CSA Puglia Study Group

Affiliations

Local Health Service Taranto, Italy (Simona Leogrande, Antonia Mincuzzi, Aldo Minerba)

Department of Epidemiology, Lazio Regional Health Service, Rome – ASL Roma 1, Italy (Ester Rita Alessandrini, Massimo Stafoggia, Carla Ancona, Francesco Forastiere)

ARPA Puglia, Bari, Italy (Roberto Giua, Angela Morabito, Stefano Spagnolo, Giorgio Assennato)

AReS Puglia, Bari, Italy (Lucia Bisceglia, CSA Puglia Study Group)

Contact information

Simona Leogrande

Address: V.le Virgilio 31, 74121 – Taranto, Italy

Email: simona.leogrande@asl.taranto.it

Abstract (250 words max)

Background: A large steel plant close to the urban area of Taranto (Southern Italy) has been operating since the sixties. Several epidemiological studies conducted in the past 20 years have reported an excess of mortality and morbidity from various causes at the town level, possibly due to air pollution from the plant. However, the assessment of the causal relationship between air pollutants emitted from the plant and adverse health outcomes has been controversial. We applied a variant of the "difference-in-difference" (DID) approach to examine the relationship between changes in exposure to industrial PM_{10} from the plant and changes in cause-specific mortality rates.

Methods: We examined a dynamic cohort of all subjects (321,356 individuals) resident in the Taranto area in 1998-2010 and followed them up for mortality till 2014. In this work, we included in the analysis only deaths occurring on 2008-2014 and calculated annual death rates by district of residence and age class. PM_{10} and NO_2 concentrations measured at fixed monitoring stations were used, together with the results of a dispersion model, to estimate annual average population exposures to PM_{10} of industrial origin for each year, district and age class. The method presumably removes "by design" all confounding from temporal and individual covariates.

Results: During 2008-2014, we observed a total of 15,303 natural deaths in the cohort. We estimated an increased risk in natural mortality (1.86%, 95% Confidence Interval [CI]: -0.06, 3.83%, relative to 1 $\mu g/m^3$ annual change of industrial PM_{10}) mainly driven by respiratory causes (8.74%, 95% CI: 1.50, 16.51%). The associations were stronger in elderly (65+ years).

Conclusions: The study indicates an effect of industrial PM_{10} on natural mortality in the study area, especially in the elderly population. The potential effect of individual confounders has been removed and we postulate that the relationship is causal.

Main text

Introduction

Environmental epidemiology is an observational discipline which investigates the relationship between environmental exposures and health outcomes without any intervention of the researcher in exposure assignment, opposite to what happens in the experimental study designs where randomization guarantees comparability of the exposed and unexposed populations with regards to external factors. The "causal inference" methods to observational studies, in the spirit of mimicking randomized trials, has been applied to several fields in biomedicine [1, 2 3 4 5 6 7] but the applications in environmental epidemiology are sparse. [8 9 10 11 12] There are many techniques commonly used in traditional epidemiology to control known and measured confounding factors (restriction, stratification, matching, standardization, regression models), the problem is that none can completely rule out the presence of residual confounding, i.e. lack of comparability between exposed and unexposed groups for factors that are not controlled at all or are controlled for but are measured inaccurately. [12] Of course, the use of sensitivity analyses to test the robustness of the results to alternative modeling formulations is of some use, but the problem of residual bias from potential un- (or mis-) measured confounding remains. A promising alternative is to apply methods that try to remove confounding "by design", e.g. by trying to mimic randomized clinical trials. [13]

In this study we have adopted a "differences in differences" approach to investigate the relationship between exposure to air pollution from a specific

industrial source and mortality in the nearby population. The method has been applied in the past in the econometric literature and it can be seen as a "before/after study with a control group". The essence of the design is that differences in exposures across time are related to differences in rates of diseases in the same populations so that the role of potential individual and behavioral factors are cancelled out as the comparisons are occurring within populations. Of course, variability of exposure across time is essential to appreciate differences in disease occurrence. A recent study in the USA has used a variant of this method to evaluate the causal effects of long-term $PM_{2.5}$ exposure on mortality in New Jersey. [10 12]

The present study has been conducted in the area of Taranto, South of Italy, where a large industrial area has been operating in close proximity of the resident population since several decades and the possible health effects due to the pollution have been under scrutiny, [14-22] debate, [23-25] and legal case[26] in the last few years.

METHODS

Study area and enrolment of the cohort

The study area is 445.17 km^2 large, is located in the Apulia region (south-east of Italy) and includes the municipalities of Taranto, Massafra and Statte. The municipality of Taranto (area: 249.86 km^2, population on 1st January 2014: 203,257, Source: Italian Institute of Statistics-Istat) is the main town of the province, overlooking the Ionian Sea and for its geographical location is known as the "city of the two seas", in fact it stretches between two seas: "Mar Grande" (Big

sea) and "Mar Piccolo" (Little sea). Because of its location and geographical structure, it was considered a strategic point and became the location of an industrial, commercial and military harbor. It is characterized by the presence of a large industrial area which is placed in the northwest of the city and includes: a refinery, a cement plant, and the ILVA steel plant operating since 1965, which represents the largest European site for the steel production.

The ILVA is an "integrated cycle" plant, so its main production are: coke, sinter, pig iron, solid steel, hot rolled coils, cold rolled coils, hot galvanized coils, hot rolled heavy plates, black or coated welded pipes. A power plant is also integrated in the area. On 30th November 2015 it had 15,487 directly employed workers and there were estimated 8,000/10,000 additional workers in the satellite activities (Source: Ilva website). The surface extension of the ILVA steel plant is about 15.45 km^2 of which about 10.45 km^2 in the municipality of Taranto and about 5 km^2 in the nearby city of Statte. The latter one is placed in the northwest of the city of Taranto (Statte area: 67.32 km^2, population on 1st January 2014: 14,190, Source: Istat) where there is also the municipality of Massafra (area: 128 km^2, population on 1st January 2014: 32,780, Source: Istat).

In a previous study,[18] an open residential cohort has been enrolled, including all subjects resident in the study area at 01/01/1998 and those who entered in the three municipalities until 31/12/2010. Data on resident population (vital status, residential address and changes in address over time) were provided by the General Registry Offices of the three cities and were used after careful quality

control procedures. We used that cohort and updated the vital status until 31/12/2014. Causes of death were provided by the Local Health Authority in Taranto. For the aims of the study, we retained only the latest period: 2008-2014, as it matched the availability of environmental data and was short enough to reduce potential residual confounding from temporal covariates (see later for the assumption of the D-D model). All the residential addresses of the members of the cohort were geocoded using the ARCGIS software. On that basis, we assigned to each subject the corresponding district and census tract of residence. The list and a brief description of the 9 districts (and the corresponding census tracts) are illustrated in the Supplementary material.

Exposure Assessment

Our objective was to estimate individual exposure to PM_{10} from industrial origin for each year of the study. To this aim, we combined information from different sources. We considered four monitoring stations of the Regional Agency for the Environmental Protection (ARPA) network operating in the study area for the study period 2008-2014, which measured concentrations of nitrogen dioxide (NO_2) and particulate matter 10 micrometers or less in diameter (PM_{10}). One station was located in an urban area of Taranto ("Via Alto Adige"), another one in a suburban setting away from the industrial area ("Via Ugo Foscolo" in the Talsano district), to measure background concentrations. The remaining two stations were located near the industrial area: "Via Machiavelli" and "Via Archimede", in the Tamburi district at the border of the industrial area. These data were used to describe temporal variability in exposure over the study area.

The regional Environmental Protection Agency (ARPA Puglia) developed, for the year 2010, a dispersion model of PM_{10} emitted from the ILVA steel plant using a Lagrangian particle dispersion model (SPRAY) built on the available information about measured emissive characteristics, orography and meteorology.[27] This model provided an estimate of the spatial distribution of ILVA-related PM_{10} concentrations, valid for the year 2010, which we assumed remained unaltered over the years of the study. We combined annual measurements from the monitoring stations and the data of the 2010 dispersion model map in order to estimate industrial PM_{10} exposures at district level for each year of the period 2008-2014 according to a methodology illustrated in the Supplementary material.

Mortality data

Causes of death were ascertained using a record linkage procedure with the mortality registry of the Local Health Authority in Taranto. In this registry, causes of death were coded using the *International Classification of Diseases, 9th Revision* (ICD IX) for the period 2008-2010 and *10th Revision* (ICD X) for the period 2011-2014.

We analyzed the following causes of death: natural causes (ICD IX 001-799, ICD X A00-R99), diseases of the circulatory system (ICD IX 390-459, ICD X I00-I99), heart diseases (ICD IX 390-429, ICD X I00-I51), and respiratory diseases (ICD IX 460-519, ICD X J00-J99).

Statistical analysis

For each year (seven: 2008-2014), area-unit (eleven: 9 Taranto districts + Massafra and Statte) and age class (four: 0-34, 35-64, 65-74, >75 years) we computed counts of cause-specific deaths. Furthermore, using cohort information and individual residential history, we estimated person-years to be used as denominators to calculate mortality rates or as "offset" in the multivariate Poisson regression analysis described below.

We defined the following model:

$$\ln\left[E\left(Y_{q,t,e}\right)\right] = \beta_0 + \beta_1 I_q + \beta_2 T + \beta_3 I_e + \beta_4 I_q * T + \beta_5 I_e * T + \beta_6 PM_{10\ q,t,e} + \ln\left(P_{q,t,e}\right) \quad (3)$$

where:

- $Y_{q,t,e}$ represents the number of deaths in area-unit q, year t and age class e;
- $PM_{10\ q,t,e}$ is the mean concentration of the industrial PM_{10} in the same stratum;
- $P_{q,t,e}$ is an offset term which represents person-years at risk;
- I_q, I_e define dummy variables for area units and age classes, respectively;
- T is a continuous variable for year, which we modeled linearly.

Furthermore:

- β_0 is the intercept term;
- β_1, β_2, β_3 are regression coefficients adjusting for confounding induced by factors varying across districts (β_1) and age classes (β_3) when T=0, and over time (β_2) in the reference stratum of district and age group. Their meaning is of little interest;

- β_4 removes potential confounding introduced by factors, known or unknown, which might display different linear time trends across area units;
- β_5, similarly to β_4, adjusts for potential confounders which display different linear time trends across age groups;
- β_6 represents the causal effect of industrial PM_{10}.

We can interpret the model in (3) as a variant of the difference-in-differences method, it is similar to the method proposed by Wang et al. [10] The idea underlying model (3) is that a causal effect of PM is obtained by removing confounding from spatio-temporal covariates by design. This is achieved through the introduction of age-specific and district-specific linear trends in the regression model. For example, if socio-economic status or lifestyle factors (smoking, diet, etc.) have changed differently across districts or age groups over 2008-2014, and if such changes have been collinear with concurrent changes in air pollution, the linear trends introduced in the model should account for that, provided that such changes have been linear. The drawback of the model, on the other side, is that only fluctuations of PM around linear trends are contrasted to concurrent fluctuations in mortality rates, with consequent decrease of statistical power.

Results were presented as percent increase risk of death, and 95% confidence intervals, relative to 1μg/m^3 variation of industrial PM_{10}. This unit of measure has been chosen because easy to interpret, and close to the interquartile range at population level (1.6 μg/m^3).

We tested if the association was modified by age. In this regard we considered two age classes: <65 years, 65+ years. The effect modification was tested by stratification. Analytically, we calculated the difference between the strata coefficients. Assuming that the difference between the coefficients was distributed as a normal distribution with mean zero and variance equal to the sum of the strata variances (e.g. covariance=0), we evaluated the p-value of the relative effect modification (REM).[14]

The follow-up and the statistical analysis were performed using SAS 9.0 (SAS Institute Inc., Cary, NC) and Stata 13 (StataCorp LLC, College Station, Texas) and geographical data were analyzed using Arcgis (Esri, Redlands, California) and Qgis.

RESULTS

Figure 1 illustrates the study area, divided into 11 small area units: the two municipalities of Massafra and Statte and 9 districts of the city of Taranto. The dots represent the ARPA monitoring stations, with a circle around the stations considered for the estimation of the industrial PM_{10}. Figure 2 shows the 2010 pollutant dispersion model of the ILVA steel plant for PM_{10}. Figure 3 displays the productivity (kton/year) of the ILVA plant in the period under study divided by type: steel, coke, sinter and pig iron. Figure 4 illustrates the annual average PM_{10} (measured overall concentrations) assessed in the four monitoring stations.

A total of 321,356 subjects resident in the study area were originally enrolled in the cohort; a total of 262,375 individuals were still alive and resident at the start of

our study at January 1st, 2008. At 31/12/2014 238,473 people (74.2%) were alive and resident in the area, 37,736 subjects (11.7%) were dead, 45,147 individuals (14.1%) were emigrated outside the study area and so considered lost to follow up. For the purposes of the present study, we only included residents in the period 2008-2014, because we had concurrent data on air pollution monitoring stations only in the latest period.

Table 1 displays the person-years of follow-up and the number of deaths in the study period, by age, district and year. We estimated 1,726,353 person-years of follow-up, whose distribution is different across the area units. There were 15,303 natural deaths, 71.8% of which from subjects 75+ years old, and 0.8% from those below 35 years of age. Our enrolment period was up till 2010, therefore the cohort was open in the first three years of the study period (2008-2010) and closed afterwards. In particular, in the year 2008 there were 262,375 people, with a -0.23% change in cohort residents between 2009 and 2008. In contrast, for the last four years the changes were more significant albeit not extreme, and between 2014 and 2013 the percentage change in the population was -2.07%. We accounted of changing population over time by computing, and using as offset in the analysis, annual person-years of observations.

For descriptive purposes, we divided the study area into three sub-areas only: "Tamburi", close to the plant, "Isola, Borgo", intermediate, and "Other" which includes all the districts and municipalities different from the previous two and away from the plant. For each of the three area units we calculated the absolute

change in the estimated industrial PM_{10} between the yearly value and the mean for the all period, this is represented in Figure 5. In Figure 6 we represent the percent change of natural mortality rate (per 1,000 person-years) between the yearly value and the mean for the all period. We can observe that the annual trends in mortality rates resemble the trends of the industrial PM_{10} in the "Tamburi" and to less extent in the "Isola, Borgo" districts, i.e. the districts mostly influenced by industrial emissions, whereas the patterns of exposure and mortality in the other areas deviate.

On the basis of the model reported in equation (3), we observed (Table 2) a percent increase of natural mortality of 1.86% (95% Confidence Interval (CI): -0.06, 3.83%) relative to 1 $\mu g/m^3$ variation of industrial PM_{10}. In particular, we found a 2.37% (95% CI: 0.31, 4.47%) increase in natural mortality (REM p-value= 0.22) among subjects 65+ years old (Figure 5). Furthermore, we found a 8.74% (95% CI: 1.50, 16.51%) increase in respiratory mortality. For the same outcome there was no evidence of effect modification by age (REM p-value= 0.96). Effect estimates for circulatory and cardiac diseases were positive but affected by larger standard errors.

DISCUSSION

We found evidence for a causal link between industrial PM_{10} and mortality in the study area. In particular, annual exposure to industrial PM_{10} increased mortality for natural and respiratory causes, especially in the elderly population.

Several studies conducted in the area of Taranto have found clear evidence of an impairment of the environmental status and health of the resident population. In particular, it is worth mentioning a multicenter study and health impact assessment,[14-15] ecological studies based on a priori evidence,[16-17] a cohort study,[18] descriptive studies,[19 20 21 22] biomonitoring studies.[28 29 30] However, the issue of the possible causal link between environmental exposures from the plant and adverse health outcomes has been disputed. So, we decided to use a new and original method to assess the effect of exposure to industrial PM_{10} on mortality in the Taranto area. The novelty of the approach has been to try to remove all known and unknown confounders "by design". This has been achieved by focusing on annual fluctuations of PM around area-specific and age-specific time trends, instead of exploiting the full range of PM variability over space and time. In addition, in order to further remove potential confounding from long-term time trends, we restricted the study period to a short time window of only seven years.

In this paper we used the cohort enrolled in a previous study[18] and we updated it with the same tools. This study has, to our judgement, several strengths: careful reconstruction of the individual residential histories, good quality geocoding, high completeness in the assignment of causes of death. It should be noted that the cohort lacks data on individual risk factors (such as smoking and other lifestyle characteristics, individual estimates of exposure to meteorological parameters, etc.). However, such factors might have confounded the estimates under investigation only under the assumption that they varied differently across age groups and districts, and that such differences were not adequately captured by

linear trends. We believe it to be unlikely. Further support to this is provided by a recent re-analysis of the cohort study,[31] which applied indirect adjustment methods on data collected from an external survey. That study showed that the associations between industrial pollutants and mortality/morbidity were not biased by unmeasured individual confounders, including smoking and body-mass index.

Our analyses started from individual data, so for each subject of the cohort we had information about his/her residential history and so about his/her exposure. Therefore, we took into account the variation in time and space of the person-years and we could estimate the exposure for each year, district and age class. Some assumptions have been made. It was assumed that the changes over time in the temperature were the same in the districts (or, if different, were adequately described by linear shapes), so it was not necessary to control for this factor. Furthermore, we did not control for socioeconomic status, because it was assumed that the variations in time between area units were stable, or at least were captured by our linear time trends. Finally, it should be noticed that, for estimating the industrial PM_{10} component, the NO_2 concentrations have been used as a proxy measure for the contribution of traffic and local sources, an assumption which could be invalid if NO_2 is affected by industry as well.

In conclusion, under the model assumptions, with new and original methods the present study supports the evidence of a causal link between industrial PM_{10} and mortality of the population living in the Taranto area.

TABLE 1. Study population: number of deaths for natural causes and number of mortality person-years during the study period by age classes, area units and years.

	Number of deaths		**Person-years of follow-up**	
	n.	%	n.	%
Total	15,303	100.0	1,726,353	100.0
Age class				
0-34	116	0.8	621,334	36.0
35-64	1,771	11.6	737,862	42.7
65-74	2,427	15.9	191,020	11.1
>74	10,989	71.8	176,138	10.2
Area units				
Isola, Borgo	2,363	15.4	197,931	11.5
Italia Montegranaro	2,296	15.0	190,803	11.1
Lama, San Vito, Carelli	675	4.4	127,533	7.4
Massafra	1,569	10.3	201,381	11.7
Paolo VI	710	4.6	122,082	7.1
Salinella	1,192	7.8	160,348	9.3
Solito Corvisea	1,391	9.1	157,536	9.1
Statte	664	4.3	96,152	5.6
Talsano	1,129	7.4	176,840	10.2
Tamburi, Lido azzurro	1,152	7.5	117,166	6.8
Tre Carrare, Battisti	2,162	14.1	178,580	10.3
Year				
2008	2,201	14.38	255,446	14.80
2009	2,177	14.23	254,638	14.75
2010	2,212	14.45	253,491	14.68
2011	2,236	14.61	248,402	14.39
2012	2,173	14.20	243,077	14.08
2013	2,111	13.79	238,050	13.79
2014	2,193	14.33	233,248	13.51

TABLE 2. Number of deaths, percent increase of risk (I.R.%) and 95% confidence intervals (C.I. 95%), relative to 1 μg/m^3 variation of industrial PM_{10} during the study period: 2008-2014.

Causes of death (ICD IX)	Number of deaths	I.R. %	95% C.I.	
Natural causes (001-799)	15,303	1.86	-0.06	3.83
Circulatory system diseases (390-459)	5,721	0.70	-2.35	3.84
Heart diseases (390-429)	4,346	1.91	-1.55	5.50
Respiratory diseases (460-519)	1,150	8.74	1.50	16.51

FIGURE 1

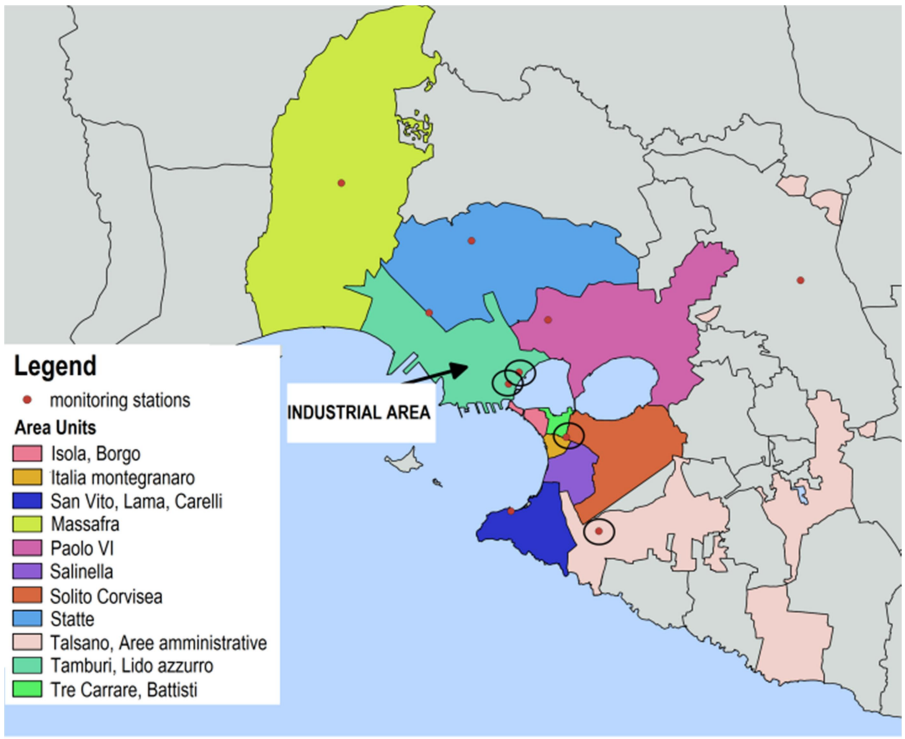

FIGURE 2

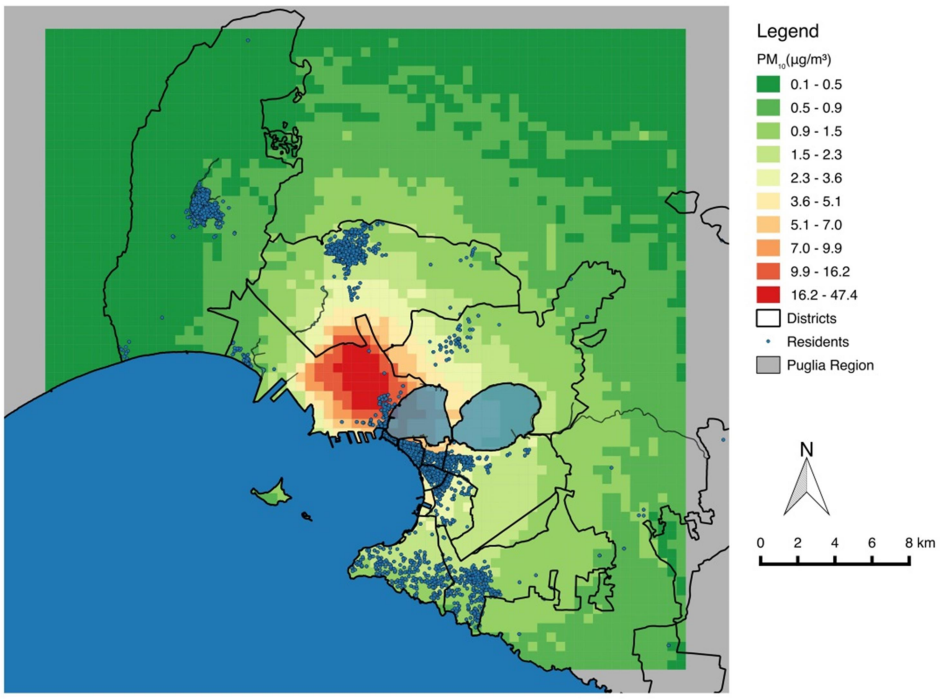

FIGURE 3

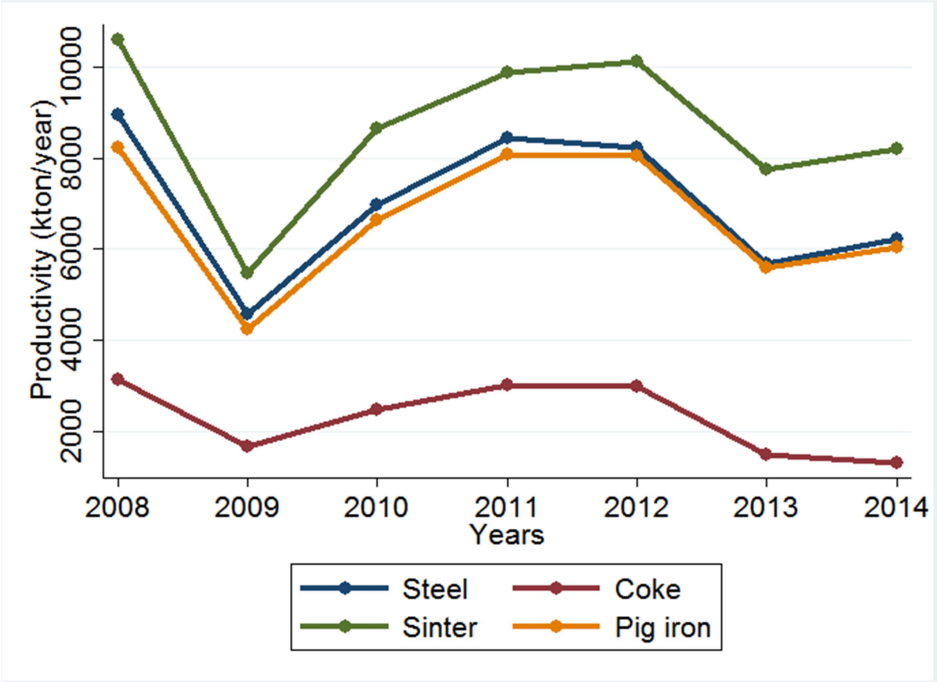

FIGURE 4

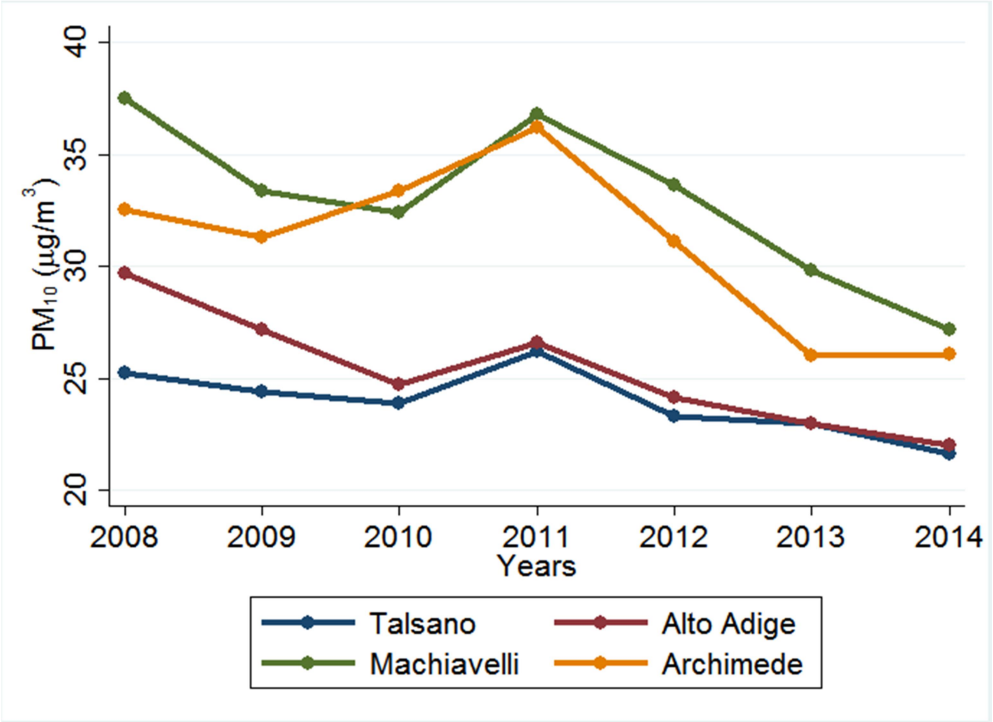

FIGURE 5

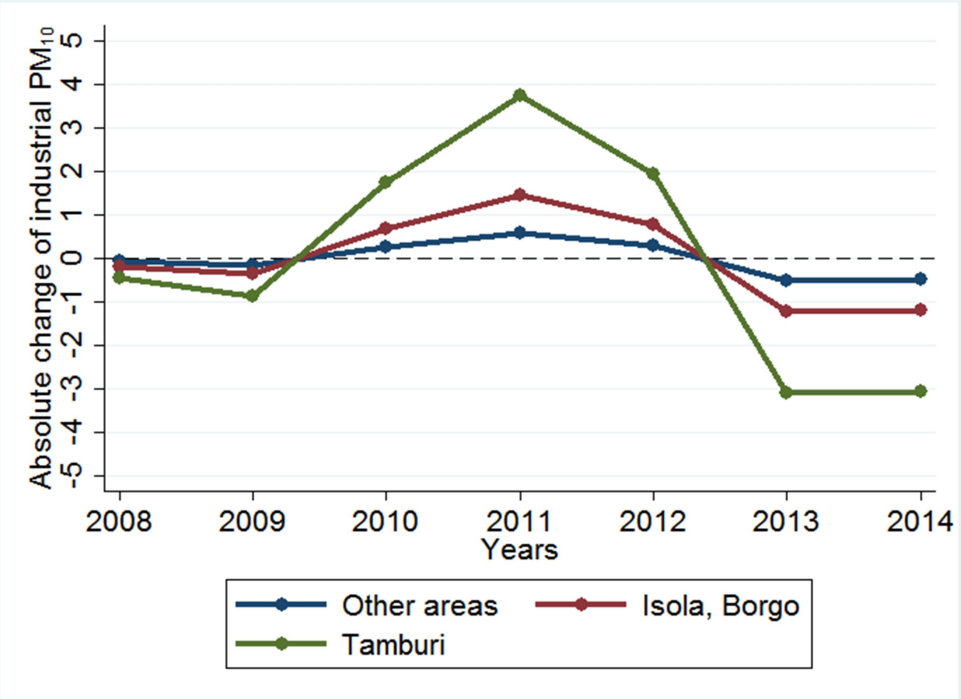

FIGURE 6

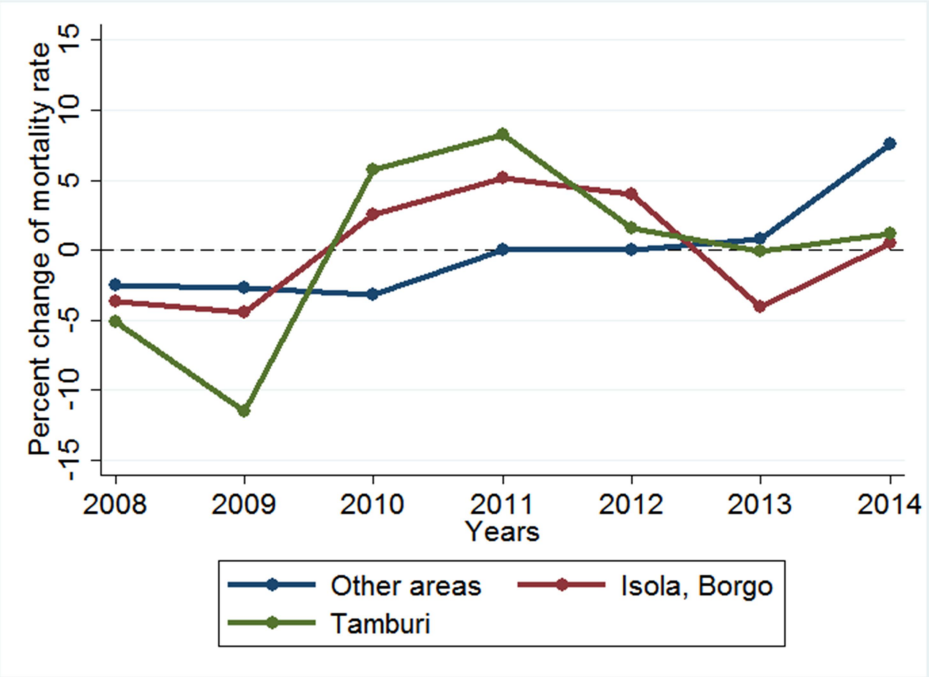

FIGURE 7

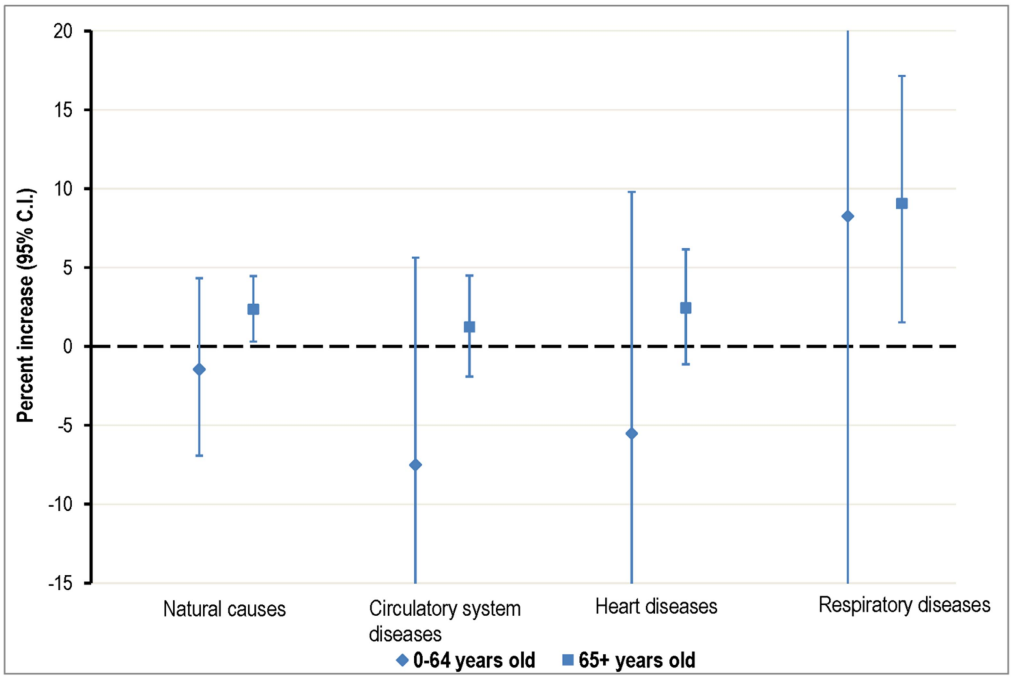
20
15
10
5
0
-5
-10
-15
Percent increase (95% C.I.)
Natural causes
Circulatory system diseases
Heart diseases
Respiratory diseases
0-64 years old
65+ years old

FIGURE LEGENDS

FIGURE 1. Study area, monitoring stations and area units of the study.

FIGURE 2. Results of the 2010 dispersion model for PM_{10} ($\mu g/m^3$).

FIGURE 3. Productivity (kton/year) of the ILVA Steel Plant by year and product.

FIGURE 4. PM_{10} ($\mu g/m^3$) concentrations at the fixed monitors by year.

FIGURE 5. Absolute change of exposure to industrial PM_{10} by year and area (annual average/average of the period)

FIGURE 6. Percent change of natural mortality rate (per 1,000 person-years) by year and area (Tamburi, Isola Borgo, and other)

FIGURE 7. Percent increase of risk of mortality (95% confidence intervals, C.I. 95%) relative to 1 $\mu g/m^3$ variation of industrial PM_{10} during the study period (2008-2014) by age class.

Supplementary material

The districts of Taranto

The municipality of Taranto was divided into nine small-area units (called districts in the paper), according to census tracts:

1. Isola-Borgo (census tracts: 1-93, 279-400);
2. Italia Montegranaro (census tracts: 401-492);
3. Paolo VI (census tracts: 222-278, 1522, 1527, 1528, 1529, 1533);
4. Salinella (census tracts: 705-801, 955, 956, 1550);
5. San Vito, Lama, Carelli (census tracts: 886-921, 922, 923, 954, 957, 1553);
6. Solito Corvisea (census tracts: 617-704, 952, 1512, 1514, 1523, 1524, 1526);
7. Talsano (census tracts: 802-885, 925, 936, 953, 1108, 1516, 1518, 1519, 1532, 1535, 1539, 924, 926-933, 934, 935, 937, 940-950);
8. Tamburi, Lido azzurro (census tracts: 94-221, 983, 1483, 1511, 1546, 1547, 1548);
9. Tre Carrare, Battisti (census tracts: 493-616).

Some of the districts are of importance because of the industrial contamination. In the northwest of the town there is the district called "Tamburi, Lido Azzurro" that includes the industrial area. Near the latter one, and in particular near the area of mineral parks, lies the residential area of the Tamburi district. The area called "Isola, Borgo" consists of two districts: Isola and Borgo. Isola is also known as

"Città Vecchia" and is the oldest part of the city; Borgo is the first district of the "new city". The district named "Paolo VI" is located in the north of the city of Taranto and it was built around the mid-sixties intended as a residential area for the steel plant workers and their families. A study from Mangia et al.[32] showed that in the city of Taranto the areas near the industrial area are the most polluted ones. The studies conducted by Gariazzo et al.[33] [34] showed that the maximum fallout of industrial emissions zones are Tamburi and Borgo.

Moving away from the Borgo district there are the areas "Tre carrare, Battisti" and "Italia Montegranaro", which together with Borgo constitute the town's commercial and the residential heart of the city. Proceeding to the south there is the vast district "Solito corvisea" and the peripheral district "Salinella". The last five districts mentioned are crossed from the main urban roads. The city ends with the peripheral areas "San Vito, Lama Carelli" and "Talsano, Aree amministrative" that can be considered the background areas since they are distant from the industrial area and only moderately affected by traffic pollution.

Exposure Assessment

To estimate industrial PM_{10} exposures for each year of the period 2008-2014, we employed the following methodology:

1. Industrial PM_{10} evaluation in the Tamburi industrial district.

Assuming that the annual mean concentration of PM_{10} measured at one station can be partitioned into three components: background (B), industrial (I) and traffic (V), we developed a set of linear systems (one for each year $t \in \{2008, \dots, 2014\}$) of three equations in three variables:

$$\begin{cases} V_{Mt} + B_t + I_{Mt} = C_{Mt} \\ \Delta_{Ft} V_{Mt} + B_t + \gamma_1 I_{Mt} = C_{Ft} \\ \Delta_{At} V_{Mt} + B_t + \gamma_2 I_{Mt} = C_{At} \end{cases} \quad (1)$$

- The first equation defines C_{Mt}, the annual average PM_{10} concentration measured on year t at the two stations placed in the Tamburi industrial area: "Via Machiavelli" and "Via Archimede". $C_{Mt}$was partitioned into three components: B_t background (assumed the same all over the sites), V_{Mt} traffic in that site and I_{Mt} industrial in that site;
- The second equation partitions C_{Ft}, the annual mean concentration of PM_{10} measured on year t at the "Via Ugo Foscolo" (Talsano district) station, as the sum of three addenda: B_t background component, $V_{Ft} = \Delta_{Ft} V_{Mt}$ traffic component and $I_{Ft} = \gamma_1 I_{Mt}$ industrial component in the year t. The traffic component in Talsano was estimated as the product of the traffic component at Tamburi (V_{Mt}) and Δ_{Ft}, the ratio of annual mean NO_2 concentrations in Talsano over annual mean NO_2 concentrations in the two stations of Tamburi (NO_2 has been used as a proxy measure for contribution from traffic and local sources). 14 Similarly, the industrial component in Talsano was estimated as the product of the industrial component at Tamburi (I_{Mt}) to γ_1, ratio of the PM_{10} predicted value from the ARPA dispersion model at Talsano to the corresponding predicted value at Tamburi;
- The third equation partitions C_{At} , annual mean PM_{10} concentrations in Via Alto Adige, as the sum of three components, background, traffic and industry. These were estimated based on the same assumptions made in the second equation.

More explicitly, in the system (1), we have assumed that the background component B_t is the same in all monitoring stations, since it represents background pollution levels originated by common large-scale sources plus transboundary transport processes. Furthermore, as already explained, we have assumed that: a) it was possible to estimate relative contributions of traffic at different points as ratios between NO_2 concentrations in those points, and b) it was possible to estimate relative contributions of industry at different points as ratios between predicted PM_{10} concentrations from the ARPA dispersion model. Solving the system (1) it was possible to estimate I_{Mt}: the industrial component in the year t, in the Tamburi district.

2. Individual industrial PM_{10} evaluation

For each individual i in the cohort, and each year t between 2008 and 2014, we estimated annual industrial PM_{10} concentrations at his/her residential address based on the following formula:

$$I_{i,t} = P_i \cdot I_{Mt} \qquad (2)$$

where P_i represents the ratio of the individual predicted PM_{10} value from the dispersion model over the average of the predicted values in the two stations "Via Machiavelli" and "Via Archimede". This formula assumes that the relative contribution of the ILVA steel plant in two points, as estimated by their ratio, remains constant over the years.

3. Industrial PM_{10} evaluation for each year, district and age class

Finally, we estimated the exposure to industrial PM_{10} by year, district and age class by averaging the individual contributions ($I_{i,t}$).

REFERENCES

1. Gillespie IA, Floege J, Gioni I, et al. Propensity score matching and persistence correction to reduce bias in comparative effectiveness: the effect of cinacalcet use on all-cause mortality. Pharmacoepidemiol Drug Saf. 2015;24:738-747.

2. Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology. 2008;19:766-779.

3. Shimizu M, Yi S, Tuot S, et al. The impact of a livelihood program on depressive symptoms among people living with HIV in Cambodia. Glob Health Action. 2016;9.

4. Lu B, Marcus S. Evaluating long-term effects of a psychiatric treatment using instrumental variable and matching approaches. Health Serv Outcomes Res Methodol. 2012;12:288-301.

5. Jacob BJ, Sutradhar R, Moineddin R, Baxter NN, Urbach DR. Methodological approaches to population based research of screening procedures in the presence of selection bias and exposure measurement error: colonoscopy and colorectal cancer outcomes in Ontario. BMC Med Res Methodol. 2013;13(1):59. doi: 10.1186/1471-2288-13-59.

6. Lu M. The productivity of mental health care: an instrumental variable approach. J Ment Health Policy Econ. 1999;2:59-71.

7. Afendulis CC, He Y, Zaslavsky AM, Chernew ME. The Impact of Medicare Part D on Hospitalization Rates. Health Serv Res. 2011;46:1022-1038.

8. Galiani S, Gertler P, Schargrodsky E. Water for Life: The Impact of the Privatization of Water Services on Child Mortality. Journal of Political Economy. 2005;113(1):83-120.

9. Branas CC, Cheney RA, MacDonald JM, Tam VW, Jackson TD, Ten Have TR. A difference-in-differences analysis of health, safety, and greening vacant urban space. Am J Epidemiol. 2011;174:1296-1306.

10. Wang Y, Kloog I, Coull BA, Kosheleva A, Zanobetti A, Schwartz JD. Estimating Causal Effects of Long-Term PM2.5 Exposure on Mortality in New Jersey. Environ Health Perspect. 2016;124:1182-1188.

11. Benmarhnia T, Bailey Z, Kaiser D, Auger N, King N, Kaufman JS. A Difference-in-Differences Approach to Assess the Effect of a Heat Action Plan on Heat-Related Mortality, and Differences in Effectiveness According to Sex, Age, and Socioeconomic Status (Montreal, Quebec). Environ Health Perspect. 2016;124:1694-1699.

12. Kioumourtzoglou MA, Schwartz J, James P, Dominici F, Zanobetti A. PM2.5 and Mortality in 207 US Cities: Modification by Temperature and City Characteristics. Epidemiology. 2016;27:221-227.

13. Rothman KJ. Epidemiology: An Introduction. New York: Oxford University Press; 2012.

14. Biggeri A, Bellini P, Terracini B. Meta-analysis of the Italian studies on short-term effects of air pollution--MISA 1996-2002. Epidemiol Prev. 2004;28:4-100.

15. Berti G, Galassi C, Faustini A, Forastiere F. Air pollution and health: epidemiological surveillance and prevention. Epidemiol Prev. 2009;33:1-144.

16. Pirastu R, Iavarone I, Pasetto R, Zona A, Comba P, SENTIERI Working Group. SENTIERI Project - Mortality study of residents in Italian polluted sites: RESULTS. Epidemiol Prev. 2011;35:1-204.

17. Pirastu R, Ancona C, Iavarone I, et al. SENTIERI Project. Mortality study of residents in Italian polluted sites: evaluation of the epidemiological evidence. Epidemiol Prev. 2010;34:1-100.

18. Mataloni F, Stafoggia M, Alessandrini E, Triassi M, Biggeri A, Forastiere F. A cohort study on mortality and morbidity in the area of Taranto, Southern Italy. Epidemiol Prev. 2012;36:237-252.

19. Mincuzzi A, Minerba S, Tafuri S, et al. IESIT - Indagine epidemiologica nel sito inquinato di Taranto. Lecce: CLIO EDU; 2013.

20. Martinelli D, Mincuzzi A, Minerba S, et al. Malignant cancer mortality in Province of Taranto (Italy). Geographic analysis in an area of high environmental risk. J Prev Med Hyg. 2009;50:181-190.

21. Comba P, Pirastu R, Conti S, et al. Environment and health in Taranto, southern Italy: epidemiological studies and public health recommendations. Epidemiol Prev. 2012;36:305-320.

22. Vigotti MA, Mataloni F, Bruni A, Minniti C, Gianicolo EA. Mortality analysis by neighbourhood in a city with high levels of industrial air pollution. Int J Public Health. 2014;59:645-653.

23. Bianchi F. Environmental epidemiology and communication in the Taranto crisis (commentary). Epidemiol Prev. 2012(36):332-336.

24. Michelozzi P. Environmental disaster in Taranto, southern Italy: the contribution of epidemiology (editorial). Epidemiol Prev. 2012;36:231-233.

25. ILVA di Taranto: cosa ne dicono gli epidemiologi? (commentary). Epidemiol Prev. 2012;36:1.

26. Ministero dell'ambiente e della tutela del territorio e del mare. Perizia Epidemiologica. http://aia.minambiente.it/DomandeAIADocumenti.aspx?id=333. Accessed 2017.

27. ARPA Puglia. Rapporto sulla Valutazione del Danno Sanitario Stabilimento ILVA di Taranto. 2013. http://www.arpa.puglia.it/web/guest/vds. Accessed 2017.

28. Bustaffa E, Minichilli F, Bianchi F. Studies on markers of exposure and early effect in areas with arsenic pollution: methods and results of the project SEpiAs. Epidemiol Prev. 2014;38:1-97.

29. Iavarone I, De Felip E, Ingelido AM, et al. Exploratory biomonitoring study among workers of livestock farms of the Taranto Province. Epidemiol Prev. 2012;36:321-331.

30. Serinelli M, Bisceglia L, Vimercati L, et al. Exposure assessment to heavy metals in general population in an area at high environmental risk through biological monitoring (Abstract). Occup Environ Med. 2011;68:A94.

31. Centro Salute Ambiente. Studio di coorte sugli effetti delle esposizioni ambientali ed occupazionali sulla morbosità e mortalità della popolazione residente a Taranto. 2016. https://www.sanita.puglia.it/web/csa. Accessed 2017.

32. Mangia C, Gianicolo EA, Bruni A, Vigotti MA, Cervino M. Spatial variability of air pollutants in the city of Taranto, Italy and its potential impact on exposure assessment. Environ Monit Assess. 2013;185:1719-1735.

33. Gariazzo C, Papaleo V, Pelliccioni A, Calori G, Radice P, Tinarelli G. Application of a Lagrangian particle model to assess the impact of harbour, industrial and urban activities on air quality in the Taranto area, Italy. Atmospheric Environment. 2007;41:6432-6444.

34. Gariazzo C. Analisi modellistica dell'inquinamento atmosferico di origine industriale e antropica nell'area di Taranto. Roma: Texmat; 2006.

Annex C - List of indicators from the Italian 2001 census

NOME CAMPO	DEFINIZIONE
COD_PRO	Codice numerico della Provincia
COD_COM	Codice numerico del Comune
PRO_COM	Codice numerico che identifica univocamente il 'comune' nell'ambito del territorio nazionale. Il valore è ottenuto dalla concatenazione del codice provinciale e comunale
SEZ2001	Codice che identifica univocamente la sezione di censimento 2001 nell'ambito del territorio nazionale. Il valore è ottenuto dalla concatenazione del PRO_COM con il campo SEZIONE
SEZIONE	Codice che identifica univocamente la sezione di censimento 2011 nell'ambito di ciascun comune.
P1	Popolazione residente - TOTALE
P2	Popolazione residente - Maschi
P3	Popolazione residente - Femmine
P4	Popolazione residente - Celibi/nubili
P5	Popolazione residente - Coniugati/e
P6	Popolazione residente - Separati/e legalmente
P7	Popolazione residente - Vedovi/e
P8	Popolazione residente - Divorziati/e
P9	Popolazione residente - Maschi celibi
P10	Popolazione residente - Maschi coniugati o separati di fatto
P11	Popolazione residente - Maschi separati legalmente
P12	Popolazione residente - Maschi vedovi
P13	Popolazione residente - Maschi divorziati
P14	Popolazione residente - età < 5 anni
P15	Popolazione residente - età 5 - 9 anni
P16	Popolazione residente - età 10 - 14 anni
P17	Popolazione residente - età 15 - 19 anni
P18	Popolazione residente - età 20 - 24 anni
P19	Popolazione residente - età 25 - 29 anni
P20	Popolazione residente - età 30 - 34 anni
P21	Popolazione residente - età 35 - 39 anni
P22	Popolazione residente - età 40 - 44 anni
P23	Popolazione residente - età 45 - 49 anni
P24	Popolazione residente - età 50 - 54 anni
P25	Popolazione residente - età 55 - 59 anni
P26	Popolazione residente - età 60 - 64 anni
P27	Popolazione residente - età 65 - 69 anni
P28	Popolazione residente - età 70 - 74 anni
P29	Popolazione residente - età > 74 anni
P30	Popolazione residente - Maschi - età < 5 anni
P31	Popolazione residente - Maschi - età 5 - 9 anni
P32	Popolazione residente - Maschi - età 10 - 14 anni

P33	Popolazione residente - Maschi - età 15 - 19 anni
P34	Popolazione residente - Maschi - età 20 - 24 anni
P35	Popolazione residente - Maschi - età 25 - 29 anni
P36	Popolazione residente - Maschi - età 30 - 34 anni
P37	Popolazione residente - Maschi - età 35 - 39 anni
P38	Popolazione residente - Maschi - età 40 - 44 anni
P39	Popolazione residente - Maschi - età 45 - 49 anni
P40	Popolazione residente - Maschi - età 50 - 54 anni
P41	Popolazione residente - Maschi - età 55 - 59 anni
P42	Popolazione residente - Maschi - età 60 - 64 anni
P43	Popolazione residente - Maschi - età 65 - 69 anni
P44	Popolazione residente - Maschi - età 70 - 74 anni
P45	Popolazione residente - Maschi - età > 74 anni
P46	Popolazione residente di 6 anni e più -TOTALE
P47	Popolazione residente di 6 anni e più - Laurea o diploma universitario o terziario di tipo non universitario
P48	Popolazione residente di 6 anni e più - Diploma di scuola secondaria superiore
P49	Popolazione residente di 6 anni e più - Media inferiore
P50	Popolazione residente di 6 anni e più - Licenza elementare
P51	Popolazione residente di 6 anni e più - Alfabeti
P52	Popolazione residente di 6 anni e più - Analfabeti
P53	Popolazione residente di 6 anni e più - Maschi - TOTALE
P54	Popolazione residente di 6 anni e più - Maschi - Laurea o diploma universitario o terziario di tipo non universitario
P55	Popolazione residente di 6 anni e più - Maschi - Diploma di scuola secondaria superiore
P56	Popolazione residente di 6 anni e più - Maschi - Media inferiore
P57	Popolazione residente di 6 anni e più - Maschi - Licenza elementare
P58	Popolazione residente di 6 anni e più - Maschi - Alfabeti
P59	Popolazione residente di 6 anni e più - Maschi - Analfabeti
P60	Forze lavoro - TOTALE
P61	Forze lavoro - Occupati
P62	Forze lavoro - Disoccupati e altre persone in cerca di occupazione
P64	Forze lavoro - Maschi
P65	Forze lavoro - Maschi - Occupati
P66	Forze lavoro - Maschi - Disoccupati e altre persone in cerca di occupazione
P68	Occupati per sez A,B - Agricoltura Totale
P69	Occupati per sez C,E - Industria (Estrazione, Produzione energia)
P70	Occupati per sez D - Industria (Manifatturiere)
P71	Occupati per sez F - Industria (Costruzioni)
P72	Occupati - Industria Totale
P73	Occupati per sez G,H - Altre attività (Commercio/riparazioni, Alberghi/ristoranti)
P74	Occupati per sez I - Altre attività (Trasporti/comunicazioni)
P75	Occupati per sez J - Altre attività (Intermediazione)

P76	Occupati per sez K - Altre attività (Immobiliari, professionali, imprenditoriali)
P77	Occupati per sez L - Altre attività (Pubblica Amm., difesa, assicur. sociale)
P78	Occupati per sez M - Altre attività (Istruzione)
P79	Occupati per sez N - Altre Attività (Sanità, Servizi sociali)
P80	Occupati per sez O,P,Q - Altre attività (Servizi pubblici/domestici, org. extraterritoriali)
P81	Occupati per sez - Altre attività Totale
P82	Occupati - Maschi -per sez A,B - Agricoltura totale
P83	Occupati - Maschi -per sez C,E - Industria (Estrazione, Produzione energia)
P84	Occupati - Maschi -per sez D - Industria (Manifatturiere)
P85	Occupati - Maschi -per sez F - Industria (Costruzioni)
P86	Occupati - Maschi -- Industria totale
P87	Occupati - Maschi -per sez G,H - Altre attività (Commercio/riparazioni, Alberghi/ristoranti)
P88	Occupati - Maschi -per sez I - Altre attività (Trasporti/comunicazioni)
P89	Occupati - Maschi -per sez J - Altre attività (Intermediazione)
P90	Occupati - Maschi -per sez K - Altre attività (Immobiliari, professionali, imprenditoriali)
P91	Occupati - Maschi -per sez L - Altre attività (Pubblica Amm., difesa, assicur. sociale)
P92	Occupati - Maschi -per sez M - Altre attività (Istruzione)
P93	Occupati - Maschi -per sez N - Altre Attività (Sanità, Servizi sociali)
P94	Occupati - Maschi -per sez O,P,Q - Altre attività (Servizi pubblici/domestici, org. extraterritoriali)
P95	Occupati - Maschi -- Altre attività Totale
P96	Occupati - Imprenditori e liberi professionisti
P97	Occupati - Lavoratori in proprio
P98	Occupati - Coadiuvanti
P99	Occupati - Lavoratori dipendenti
P100	Occupati - Imprenditori e liberi professionisti in Agricoltura
P101	Occupati - Lavoratori in proprio in Agricoltura
P102	Occupati - Coadiuvanti in Agricoltura
P103	Occupati - Lavoratori dipendenti in Agricoltura
P104	Occupati - Imprenditori e liberi professionisti in Industria
P105	Occupati - Lavoratori in proprio in Industria
P106	Occupati - Coadiuvanti in Industria
P107	Occupati - Lavoratori dipendenti in Industria
P108	Occupati - Imprenditori e liberi professionisti in Altre attività
P109	Occupati - Lavoratori in proprio in Altre attività
P110	Occupati - Coadiuvanti in Altre attività
P111	Occupati - Lavoratori dipendenti in Altre attività
P112	Occupati - Maschi - Imprenditori e liberi professionisti
P113	Occupati - Maschi - Lavoratori in proprio
P114	Occupati - Maschi - Coadiuvanti

P115	Occupati - Maschi - Lavoratori dipendenti
P116	Occupati - Maschi - Imprendenditori e liberi professionisti in Agricoltura
P117	Occupati - Maschi - Lavoratori in proprio in Agricoltura
P118	Occupati - Maschi - Coadiuvanti in Agricoltura
P119	Occupati - Maschi - Lavoratori dipendenti in Agricoltura
P120	Occupati - Maschi - Imprenditori e liberi professionisti in Industria
P121	Occupati - Maschi - Lavoratori in proprio in Industria
P122	Occupati - Maschi - Coadiuvanti in Industria
P123	Occupati - Maschi - Lavoratori dipendenti in Industria
P124	Occupati - Maschi - Imprenditori e liberi professionisti in Altre attività
P125	Occupati - Maschi - Lavoratori in proprio in Altre attività
P126	Occupati - Maschi - Coadiuvanti in Altre attività
P127	Occupati -Maschi - Lavoratori dipendenti in Altre attività
P128	Non appartenente alle forze lavoro - TOTALE
P129	Non appartenente alle forze lavoro - Maschi
P130	Non forze lavoro - casalinghi/e
P131	Non forze lavoro - studenti
P132	Non forze lavoro - Maschi - Studenti
P133	Non forze lavoro - Ritirati dal lavoro
P134	Non forze lavoro - Maschi - Ritirati dal lavoro
P135	Non forze lavoro - Altra condizione
P136	Non forze lavoro - Maschi - Altra condizione
P137	Popolazione residente che si sposta giornalmente nel comune di dimora abituale
P138	Popolazione residente che si sposta giornalmente fuori del comune di dimora abituale
A1	Abitazioni totali
A2	Abitazioni occupate da persone residenti
A3	Abitazioni occupate solo da persone non residenti
A4	Abitazioni vuote
A5	Altri tipi di alloggio-TOTALE
A6	Stanze in totale
A7	Stanze in abitazioni occupate da persone residenti
A9	Abitazioni occupate da persone residenti in proprietà
A10	Abitazioni occupate da persone residenti in affitto
A11	Abitazioni occupate da persone residenti ad altro titolo
A12	Abitazioni occupate da persone residenti con una stanza
A13	Abitazioni occupate da persone residenti con 2 stanze
A14	Abitazioni occupate da persone residenti con 3 stanze
A15	Abitazioni occupate da persone residenti con 4 stanze
A16	Abitazioni occupate da persone residenti con 5 stanze
A17	Abitazioni occupate da persone residenti con 6 o più stanze
A18	Abitazioni totali fornite acqua potabile
A19	Abitazioni totali fornite di gabinetto
A20	Abitazioni totali fornite di vasca da bagno e/o doccia

A21	Abitazioni occupate da persone residenti fornite di una linea telefonica fissa attiva
A22	Abitazioni totali senza acqua potabile e gabinetto
A23	Superficie delle abitazioni totali
A24	Abitazioni occupate da persone residenti fornite di impianto di riscaldamento - TOTALE
A25	Abitazioni totali fornite di impianto di riscaldamento centralizzato
A44	Superficie delle abitazioni occupate da persone residenti
E1	Edifici e complessi di edifici - Totale
E2	Edifici e complessi di edifici utilizzati
E3	Edifici ad uso abitativo
E4	Edifici e complessi di edifici (utilizzati) per alberghi, uffici, commercio e industria, comunicazioni e trasporti
E6	Edifici ad uso abitativo in muratura portante
E7	Edifici ad uso abitativo in calcestruzzo armato
E9	Edifici ad uso abitativo costruiti prima del 1919
E10	Edifici ad uso abitativo costruiti tra il 1919 e il 1945
E11	Edifici ad uso abitativo costruiti tra il 1946 e il 1961
E12	Edifici ad uso abitativo costruiti tra il 1962 e il 1971
E13	Edifici ad uso abitativo costruiti tra il 1972 e il 1981
E14	Edifici ad uso abitativo costruiti tra il 1982 e il 1991
E15	Edifici ad uso abitativo costruiti dopo il 1991
E16	Edifici ad uso abitativo con un piano
E17	Edifici ad uso abitativo con 2 piani
E18	Edifici ad uso abitativo con 3 piani
E19	Edifici ad uso abitativo con 4 piani o più
E20	Edifici ad uso abitativo con un interno
E21	Edifici ad uso abitativo con 2 interni
E22	Edifici ad uso abitativo da 3 a 10 interni
E23	Edifici ad uso abitativo con più di dieci interni
E24	Totale interni in edifici ad uso abitativo
PF1	Famiglie totale
PF2	Totale componenti delle famiglie
PF3	Famiglie 1 componente
PF4	Famiglie 2 componenti
PF5	Famiglie 3 componenti
PF6	Famiglie 4 componenti
PF7	Famiglie 5 componenti
PF8	Famiglie 6 e oltre componenti
PF9	Componenti delle famiglie residenti di 6 e oltre componenti
ST01	Stranieri residenti in Italia - Europa
ST02	Stranieri residenti in Italia - Africa
ST03	Stranieri residenti in Italia - America
ST04	Stranieri in Italia - Asia
ST05	Stranieri in Italia - Oceania
ST06	Apolidi residenti in Italia

ST07	Stranieri residenti in Italia - Totale