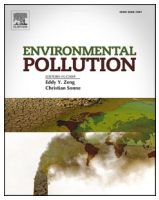
Environmental Pollution 327 (2023) 121515

Contents lists available at ScienceDirect

Environmental Pollution

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



Comparison of traditional Cox regression and causal modeling to investigate the association between long-term air pollution exposure and natural-cause mortality within European cohorts☆

Kathrin Wolf[a,*], Sophia Rodopoulou[b], Jie Chen[c], Zorana J. Andersen[d], Richard W. Atkinson[e], Mariska Bauwelinck[f], Nicole A.H. Janssen[g], Doris Tove Kristoffersen[h], Youn-Hee Lim[d], Bente Oftedal[i], Maciek Strak[c,g], Danielle Vienneau[j,k], Jiawei Zhang[d], Bert Brunekreef[c], Gerard Hoek[c], Massimo Stafoggia[l,m,1], Evangelia Samoli[b,1]

[a] Institute of Epidemiology, Helmholtz Zentrum München GmbH - German Research Center for Environmental Health, Neuherberg, Germany
[b] Department of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, Athens, Greece
[c] Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, the Netherlands
[d] Section of Environmental Health, Department of Public Health, University of Copenhagen, Copenhagen, Denmark
[e] Population Health Research Institute, St George's, University of London, London, UK
[f] Interface Demography, Department of Sociology, Vrije Universiteit Brussel, Brussels, Belgium
[g] National Institute for Public Health and the Environment, Bilthoven, the Netherlands
[h] Cluster for Health Services Research, Norwegian Institute of Public Health, Oslo, Norway
[i] Department of Air Quality and Noise, Norwegian Institute of Public Health, Oslo, Norway
[j] Swiss Tropical and Public Health Institute, Allschwil, Switzerland
[k] University of Basel, Basel, Switzerland
[l] Department of Epidemiology, Lazio Region Health Service / ASL Roma 1, Rome, Italy
[m] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

ARTICLE INFO

Keywords:
Causal inference
Inverse probability weighting
Air pollution
Health effects
Fine particulate matter
Nitrogen dioxide

ABSTRACT

Most studies investigating the health effects of long-term exposure to air pollution used traditional regression models, although causal inference approaches have been proposed as alternative. However, few studies have applied causal models and comparisons with traditional methods are sparse. We therefore compared the associations between natural-cause mortality and exposure to fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2) using traditional Cox and causal models in a large multicenter cohort setting. We analysed data from eight well-characterized cohorts (pooled cohort) and seven administrative cohorts from eleven European countries. Annual mean $PM_{2.5}$ and NO_2 from Europe-wide models were assigned to baseline residential addresses and dichotomized at selected cut-off values ($PM_{2.5}$: 10, 12, 15 μg/m^3; NO_2: 20, 40 μg/m^3). For each pollutant, we estimated the propensity score as the conditional likelihood of exposure given available covariates, and derived corresponding inverse-probability weights (IPW). We applied Cox proportional hazards models i) adjusting for all covariates ("traditional Cox") and ii) weighting by IPW ("causal model"). Of 325,367 and 28,063,809 participants in the pooled and administrative cohorts, 47,131 and 3,580,264 died from natural causes, respectively. For $PM_{2.5}$ above vs. below 12 μg/m^3, the hazard ratios (HRs) of natural-cause mortality were 1.17 (95% CI 1.13–1.21) and 1.15 (1.11–1.19) for the traditional and causal models in the pooled cohort, and 1.03 (1.01–1.06) and 1.02 (0.97–1.09) in the administrative cohorts. For NO_2 above vs below 20 μg/m^3, the HRs were 1.12 (1.09–1.14) and 1.07 (1.05–1.09) for the pooled and 1.06 (95% CI 1.03–1.08) and 1.05 (1.02–1.07) for the administrative cohorts. In conclusion, we observed mostly consistent associations between long-term air pollution exposure and natural-cause mortality with both approaches, though estimates partly differed in individual

☆ This paper has been recommended for acceptance by Dr Alessandra De Marco.

* Corresponding author. Institute of Epidemiology – Research Group "Environmental Risks", Ingolstädter Landstr. 1, 85764 Neuherberg, Germany.
E-mail address: kathrin.wolf@helmholtz-muenchen.de (K. Wolf).

[1] Shared last authorship.

https://doi.org/10.1016/j.envpol.2023.121515
Received 23 January 2023; Received in revised form 14 March 2023; Accepted 24 March 2023
Available online 24 March 2023

cohorts with no systematic pattern. The application of multiple modelling methods might help to improve causal inference.

299 of 300 words.

1. Introduction

Evidence is increasing that associations between long-term exposure to outdoor air pollution and mortality and morbidity also persist at low concentrations (e.g. lower than the current limit values set by the European Union (EU) (2008/50/EC Directive on Ambient Air Quality and Cleaner Air for Europe and 2004/107/EC Directive on heavy metals and polycyclic aromatic hydrocarbons in ambient air), the US Environmental Protection Agency (EPA) standards, and World Health Organization (WHO) 2005 air quality guidelines) (Brauer et al., 2022; Brunekreef et al., 2021; Dominici et al., 2022; Burnett et al., 2018; Chen and Hoek, 2020). Consequently, the WHO has recently lowered their guidelines considerably, e.g. from 10 to 5 μg/m^3 for annual average $PM_{2.5}$ and from 40 to 10 μg/m^3 for annual average NO_2 (World Health Organization (WHO), 2021).

In our previous analyses within the Effects of Low-Level Air Pollution: A Study in Europe (ELAPSE) project, we observed associations between air pollution and natural-cause mortality in the well-characterized ELAPSE pooled cohort as well as in seven population-based administrative cohorts. Thereby, we used traditional Cox regression adjusting for an extensive set of individual and area-level covariates (Strak et al., 2021; Stafoggia et al., 2022). The associations persisted when reducing our study population to participants exposed to concentrations below the current EU limits of 25 μg/m^3 for $PM_{2.5}$ and 40 for NO_2, but also below the stricter EPA standard of 12 μg/m^3 and WHO 2005 guideline for $PM_{2.5}$ of 10 μg/m^3, and 20 μg/m^3 for NO_2 which was suggested by the Health Risks of Air Pollution In Europe (HRAPIE) (Health risks of air pollution, 2013) project.

While some argue that traditional approaches (which include potential confounders as covariates in the regression model) do not inform causality, others argue against, suggesting that an integration of results from different methods is essential to improve causal inference (Pearce et al., 2019; Wu et al., 2020). Causal inference models that mimic randomized controlled trials have been suggested to be: less susceptible to sources of bias, particularly confounding; more robust to model misspecification and; if correctly specified, may identify causal relationships (Dominici et al., 2022; Higbee et al., 2020; Little and Rubin, 2000; Rubin, 2008). However, several main assumptions need to be fulfilled: consistency, positivity and no unmeasured confounders (Dominici et al., 2022; Makar et al., 2017). Consistency, sometimes also referred to as no interference or stable unit treatment value assumption, states that the potential outcome for a given observation is not affected by the exposure (or treatment assignment) of any other observation. Positivity (or overlap) implies that each individual has a positive chance of receiving any exposure level (or each level of treatment) independent of the set of potential confounders. Last, no unmeasured confounders assumes that the available covariates are sufficient to adjust for residual confounding. A major advantage of causal inference methods over traditional multivariable regression models is the splitting of the design and analyses stage. In the design stage, the approximation of a randomized study can be quantified and visualized, which should then allow an evaluation of causality of results of the analyses stage (Wu et al., 2020; Rubin, 2008).

The aim of the study was to compare the associations between natural-cause mortality and exposure to fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2) above vs. below selected cut-off values using traditional Cox and causal models in a large multicenter cohort setting.

2. Materials and methods

We followed a causal inference approach suggested by Makar and colleagues using inverse probability weighting (IPW) (Makar et al., 2017). Thereby, the exposure data is dichotomized at selected cut-off values. Then, for each binary pollutant exposure, a propensity score (PS) is calculated as the conditional likelihood of exposure given all available covariates. Finally, corresponding IPWs are derived and included in a weighted Cox proportional hazards model. Although this method implies the assumption that the exposure is binary and time-invariant, we chose this approach over others using continuous exposures (Wu et al., 2020; Higbee et al., 2020; Naimi et al., 2014). Reasons for the binary approach included simplicity in assumptions testing, interpretation and computational complexity, particularly as we aimed to include as many cohorts from our previous analyses as possible. Thereby, we did not intend to duplicate our previous findings, which specifically related to the question of health effects in the low exposure ranges. We rather aimed to supplement them by contrasting high versus low exposures using a previously applied causal modelling approach allowing easy interpretation. We a-priori defined the cut-off levels ($PM_{2.5}$: 10, 12 and 15 μg/m^3; NO_2: 20 and 40 μg/m^3) on the basis of the current US standard ($PM_{2.5}$: 12 μg/m^3), the WHO 2005 Air Quality Guidelines ($PM_{2.5}$: 10 μg/m^3; NO_2: 40 μg/m^3), the WHO HRAPIE cut-off value for health impact assessment (NO_2: 20 μg/m^3), but also dependent on the cohorts' overall distribution (Supplementary Fig. S1). However, to obtain stable estimates, we only applied the respective cut-offs to those cohorts for which more than 5% of the concentrations were both above and below the cut-offs.

For a detailed description of the pooled and administrative cohorts, respective outcomes and exposure assessment, covariates, and statistical protocols, we refer to our previous publications (Strak et al., 2021; Stafoggia et al., 2022). All included cohort studies were approved by the medical ethics committees or appropriate institutional review boards complying with all relevant national, state, and local regulations.

2.1. Study populations and natural-cause mortality definition

The ELAPSE pooled cohort consists of eight population-based cohorts (14 subcohorts) from six European countries: Sweden (Stockholm County), Denmark (Copenhagen and Aarhus, and nationwide), France (nationwide), the Netherlands (four cities), Germany (Ruhr and Augsburg areas), and Austria (Vorarlberg region) (Strak et al., 2021). Most cohorts were enrolled from large cities and surrounding regions. Recruitment and baseline examinations of most cohorts took place in the 1990s or early 2000s, and participants were followed up until death, emigration out of the study area, or end of follow-up in 2011–15, whichever came first. All cohorts harmonised their outcome and covariate data according to a common codebook, and transferred it to the Institute for Risk Assessment Sciences, Utrecht University, the Netherlands, where it was further checked, pooled, and stored on a secure server for our analyses. In addition, we analysed data from six population-based nationwide cohorts in Belgium, Denmark, England, the Netherlands, Norway, and Switzerland, and from one citywide cohort in Rome, Italy (Stafoggia et al., 2022). We enrolled participants aged 30 years or older between 2000 and 2011 based on data from population or census registries and followed them up until death, emigration out of the study area, or end of follow-up in 2011–17, whichever came first. Since data needed to be analysed in a national secure environment in all countries, we harmonised the definition of covariates and standardised the analytical strategy by centrally developing and distributing common R scripts to all cohort analysts.

For all cohorts, we defined mortality from natural causes according to the International Classification of Diseases, 9th revision (ICD-9) or

10th revision (ICD-10) as ICD-9: 001–779 or ICD-10: A00-R99, based on the underlying cause of death recorded on death certificates in mortality registries.

The present analysis and wider ELAPSE project was done in accordance with the Declaration of Helsinki. The original cohort studies were approved by the relevant authorities complying with all relevant national, state, and local regulations, and written informed consent was obtained from all participants before enrolment.

2.2. Air pollution exposure assessment

We centrally modeled annual mean concentrations of $PM_{2.5}$ and NO_2 by land use regression (LUR) models for Western Europe for the year 2010 (de Hoogh et al., 2018). The model was developed by regressing routine monitoring data from the AirBase network of the European Environmental Agency on satellite observations, chemical transport model estimates, land use, and road data and validated the models with 5-fold cross-validation (de Hoogh et al., 2018). All models performed well with an R^2 (fraction of variance explained by the respective models) of 66% for $PM_{2.5}$ and 58% for NO_2. We then applied the models to a 100 m * 100 m grid to compile concentration maps and assigned the respective exposures to the baseline residential addresses of our cohort participants. To apply the binary approach, we dichotomized the individually assigned exposures at the a-priori selected cut-off levels.

2.3. Covariates

Information on covariates was only available at baseline. In the pooled cohort, we included age, sex, year of enrolment, smoking status, duration and intensity of smoking (linear and squared for intensity), body mass index, marital status, employment status and neighbourhood or municipal level mean income in 2001 to adjust for potential confounders (Supplementary Table S1). In addition, we adjusted for subcohorts to account for differences not fully captured by the available covariates (Samoli et al., 2021). Since not all of these covariates were available for the administrative cohorts, the confounder model differed across cohorts. All cohorts included at least age (timescale), sex (strata), some individual-level socio-economic variable (though different between cohorts) and multiple area-level socioeconomic status (SES) indicators such as income, education, and unemployment rates at both the regional and neighbourhood scale (Supplementary Table S2). We defined regions as large-scale areas, such as counties and administrative regions, which was applicable for the six national cohorts only. Neighbourhoods were defined as smaller units, representing parts of a city, with about 1000–10,000 people, with some differences across cohorts. In addition, the administrative cohorts included all available individual-level variables to achieve maximal adjustment. However, only five cohorts had information on marital status and occupation, four cohorts had data on country of origin and education and only the English cohort had individual data on smoking and body-mass index (BMI).

For all analyses, we excluded participants with missing information on covariates or missing residential address at baseline.

2.4. Traditional cox regression with binary exposures

We reran the main covariate models from our previous analyses (Strak et al., 2021; Stafoggia et al., 2022) replacing the continuous exposures with the binary pollutants (as specified above) to enable a direct comparison with the binary causal inference models. Our previous comparisons of different approaches to account for subcohorts in the pooled cohort showed stable effect estimates (Samoli et al., 2021). We therefore included a categorical covariate in the model instead of the previously used strata term to align the model further with the causal model. Also, we extensively checked the proportional hazards assumptions (Grambsch and Therneau, 1994) by investigating the log-log plots, scaled Schoenfeld residuals against time as well as a global test from R survival package (cox.zph) in previous analyses. We observed some deviation from the proportional hazards assumption for BMI and smoking for some of the outcomes. However, sensitivity analyses incorporating these in strata resulted in almost identical estimates (Hvidtfeldt et al., 2021; Wolf et al., 2021).

2.5. Causal modeling framework

In the **design stage**, we applied IPW to compute participant-specific weights and to derive a weighted sample for which the distribution of covariates is balanced with regard to the binary exposure. Thus, for each of the selected cut-off levels, we fitted a logistic regression with the binary exposure as response and i) only an intercept (null model), and ii) all available covariates as explanatory variables (PS). We then calculated the weights as the ratio of the predictions from the null model i) vs. the predictions from the PS model ii) for the “treated” subjects (those above the selected cut-off values). For the "non-treated" participants (below the selected cut-off values), we calculated the weights as the ratio of the complementary predictions (Supplementary Methods M1). We assumed consistency a priori, because ambient exposures are exogenous and it is unlikely that individual mortality risks are affected by exposures of other individuals in the cohort. To check the positivity assumptions, we examined that there is an overlap of PS distributions above vs. below the selected cut-off values. The assumption of no unmeasured confounding cannot be tested. For the pooled cohort, we adjusted for the main set that was available in all cohorts. For each administrative cohort, we included all available individual- and area-level covariates. In our previous analyses applying traditional Cox, we conducted intensive sensitivity analyses (including adjustment for additional covariates, e.g. education, diet, and occupational status, in the pooled cohort and indirect adjustment for smoking and BMI in the administrative cohorts) (Strak et al., 2021; Stafoggia et al., 2022). Since results were very robust, we assume that our main set of covariates is adequate to adjust for confounding. Covariate balance before and after conditioning for IPW was assessed by standardized mean differences (as close to 0 as possible) and Kolmogorov-Smirnov-tests (0 indicates perfect identical distributions and 1 indicates perfect separation thus values close to 0 are indicative of balance) of the distributions of the covariates above vs. below the selected cut-off (Greifer, 2022).

In the **analyses stage**, we applied weighted Cox proportional hazards models with weights equal to the IPWs derived in the design stage to estimate the causal hazard rate ratios (HR) of exposure to above vs below the pollutant-specific cut-off for natural-cause mortality. For the pooled cohort, we adjusted for the subcohorts (as categorical variable) in the analysis stage rather than the design stage. We assumed the subcohort to be less important for predicting exposure, whereas adjusting for differences in outcome diagnosis coding and covariates is essential. We truncated the left and right tail of the weights to mitigate the effect of extremely large or small weights. Specifically, after some preliminary checks following Cole and Hernan (2008), we truncated at the 1st and 99th quantiles of the distribution of the standardized weights to limit the truncation as much as possible. The 95% confidence intervals were calculated based on adjusted standard errors applying robust variance estimators to account for the standardized weights (Therneau and Grambsch, 2000). Finally, we pooled the cohort-specific estimates of the administrative cohorts with fixed and random effects meta-analyses. For the latter, we used the restricted maximum likelihood estimator of the between-cohorts variance following our previous analyses (Stafoggia et al., 2022; Veroniki et al., 2016).

2.6. Sensitivity analyses

To test the robustness of our causal estimates, we alternatively truncated the standardized weights at the 5th and 95th quantiles of the distribution and repeated the analysis. We did not truncate further, e.g. at the 10th and 90th quantiles as applied in Makar et al. (2017), since

preliminary checks in the pooled and Rome cohort indicated that the most extreme weights were already captured with the initial 1st and 99th quantiles truncation.

All analyses were conducted with R software (version 3.6.0) and were based on common scripts developed and distributed by the ELAPSE statistical group.

3. Results

The pooled cohort consisted of 325,367 participants with complete data (contributing 6,339,553 person-years of observation) of whom 47,131 died from natural causes (Table 1). The seven administrative cohorts included data from 28,063,809 participants in total (257,456,655 person-years), of whom 3,580,264 died from natural causes. Mean age at enrollment was 48.7 years in the pooled cohort and ranged from 52.6 (Belgian) to 58.9 (English) years in the administrative cohorts. The percentage of females was considerably higher in the pooled cohort with 66.0% compared to 49.4–54.5% in the administrative cohorts, since four of the subcohorts included only women (Supplementary Tables S1 and S2). Modeled annual mean concentrations for the selected cut-off values for the European region are displayed in Fig. 1. Mean air pollution concentrations were highest in the Belgian ($PM_{2.5}$) and Roman (NO_2) cohorts and lowest in the Norwegian cohort (both pollutants) with considerable variation within and across cohorts (Table 1 and Supplementary Fig. S1).

The $PM_{2.5}$ cut-off value of 12 μg/m^3 could not be applied to the Belgian, Dutch and Roman cohort since less than 5% of the exposure data fell below this cut-off, whereas less than 5% of the data was above 15 μg/m^3 for the Norwegian cohort (and again below 15 μg/m^3 for the Belgian and Roman cohort; see Supplementary Table S3 for the number of persons, person-years at risk and cases above and below the selected cut-off values). For NO_2, the Norwegian, Danish, English and Swiss cohort did not exhibit enough data above 40 μg/m^3 and the Belgian and Roman cohort lacked enough data below 20 μg/m^3. For the pooled cohort, only subcohorts with more than 5% of data above and below the respective cut-off values were included. We also used 30 μg/m^3 for NO_2 as alternative cut-off value for the pooled cohort since only four subcohorts offered enough data for the cut-off value of 40 μg/m^3.

All cohorts showed a reasonable overlap of propensity score distributions above vs. below the selected cut-off values indicating that the positivity assumption was met (Supplementary Figs. S2 and S3). Covariates were partly imbalanced before conditioning but balance (standardized mean differences and/or Kolmogorov-Smirnov-tests close to 0) was achieved after conditioning for individual-level covariates, especially when truncating at the 1st and 99th percentiles (Supplementary Figs. S4 and S5). However, the balance plots indicated some deviations for some of the regional and neighbourhood socio-economic variables.

Analyses in the pooled cohort showed consistent associations with natural-cause mortality for both pollutants and both approaches (Table 2; though the main causal model was not significant for the NO_2 cut-off value of 30 μg/m^3). However, the effect sizes partly differed, e.g. they were considerably higher for the causal models compared to traditional Cox for the $PM_{2.5}$ cut-off value of 15 μg/m^3, quite similar for 12 μg/m^3 and considerably lower for both NO_2 cut-offs.

The pattern was rather heterogeneous in the administrative cohorts (Table 3). When increasing $PM_{2.5}$ exposure from levels below to above 15 μg/m^3, both approaches showed increased HR for natural-cause mortality for the majority of administrative cohorts except the English cohort with mostly comparable effect estimates (only the Danish causal estimate was twice the traditional Cox estimate). For the 12 μg/m^3 cut-off value, all traditional Cox estimates pointed to positive associations though both the English and Swiss cohort showed an inverse association in the causal models. The results were consistent for NO_2 with all cohorts showing a clear positive association and similar effect sizes of the two approaches for the 20 μg/m^3 cut-off value. For the 40 μg/m^3 cut-off value, the estimates were heterogenous with an inverse association in the Belgian cohort and null or positive associations in the Dutch and Roman cohorts (positive for the Dutch causal model and the Roman traditional Cox). Thus, the meta-analytical results indicated only a weak association for $PM_{2.5}$ but a clear signal for NO_2 with an HR of 1.06 (95% CI 1.03–1.08) for the traditional Cox model and of 1.05 (95% CI 1.02–1.07) for the causal model when increasing levels below to above 20 μg/m^3 (Fig. 2). We could only investigate the 10 μg/m^3 cut-off value for $PM_{2.5}$ within the Norwegian cohort, for which we observed HRs of 1.04 (1.03–1.05) and of 1.03 (95% CI 1.03–1.04) for the traditional and causal models, respectively (Supplementary Table S4).

3.1. Sensitivity analyses

Truncation at the 1st and 99th percentile resulted in considerably better covariate balance compared to truncation at the 5th and 95th percentile for most of the cohorts and covariates (Supplementary Figs. S4 and S5). Only the Danish ($PM_{2.5}$ cut-off 15), English ($PM_{2.5}$ cut-off 15), Dutch ($PM_{2.5}$ cut-off 15; NO_2 cut-off 20) and Rome cohort ($PM_{2.5}$ cut-off 15; NO_2 cut-off 40) indicated a better balance for some of the area-level variables for the truncation at the 5th and 95th percentile. However, there was no clear pattern across the pollutants or cut-offs and the causal HR estimates from the weighted Cox models were mostly similar except for the Swiss $PM_{2.5}$ cut-off value of 15 μg/m^3 (Tables 2 and 3 and Supplementary Fig. S6).

4. Discussion

With this large multicenter analysis of the ELAPSE pooled cohort and seven administrative cohorts, we aimed to compare the associations between natural-cause mortality and long-term exposure to $PM_{2.5}$ and NO_2 using a traditional Cox regression model and a causal modelling framework using IPW. We observed mostly consistent associations with both methods, though effect sizes partly differed in individual cohorts, with no particular pattern of larger effect estimates in either method. While the pooled cohort showed consistent associations for both pollutants, the pattern was rather heterogeneous in the administrative

Table 1
Description of the ELAPSE pooled cohort and seven administrative cohorts.

Characteristic	Pooled[a]	Norwegian	Danish	English	Dutch	Belgian	Swiss	Roman
Number of participants	325,367	2,309,001	3,083,227	1,368,740	10,376,406	5,474,548	4,188,175	1,263,712
Follow-up	1985–2015	2001–2016	2000–2015	2011–2017	2008–2012	2001–2011	2000–2014	2001–2015
Person-years at risk	6,339,553	32,531,421	42,586,464	9,084,293	50,034,558	54,575,223	53,344,296	15,300,400
Natural-cause mortality, N	47,131	524,592	714,629	145,988	590,832	707,146	661,534	235,543
Mean (SD) age at enrollment	48.7 (13.4)	53.9 (15.9)	53.0 (15.1)	58.9 (12.8)	53.6 (15.1)	52.6 (15.2)	52.7 (15.2)	55.1 (15.4)
Percentage of women	66.0%	50.9%	51.7%	52.5%	51.3%	49.4%	52.0%	54.5%
$PM_{2.5}$ (μg/m^3), mean (SD)	15.0 (3.2)	8.3 (2.6)	12.4 (1.6)	13.1 (1.4)	16.3 (1.4)	18.6 (1.6)	15.9 (2.4)	16.7 (0.9)
NO_2 (μg/m^3), mean (SD)	25.0 (8.1)	15.3 (7.8)	20.3 (7.9)	28.0 (6.9)	31.4 (7.1)	30.4 (7.3)	23.7 (7.4)	32.9 (6.1)
N of individual/area-level confounders	8/1	6/6	5/6	4/1	5/8	6/8	7/8	5/5

[a] Descriptive numbers are given for the main model 3 dataset of the pooled cohort (Strak et al., 2021). Numbers slightly differ for the different cut-off values due to an inclusion criteria of at least 5% of data that needed to be available above and below the respective cut-off value.

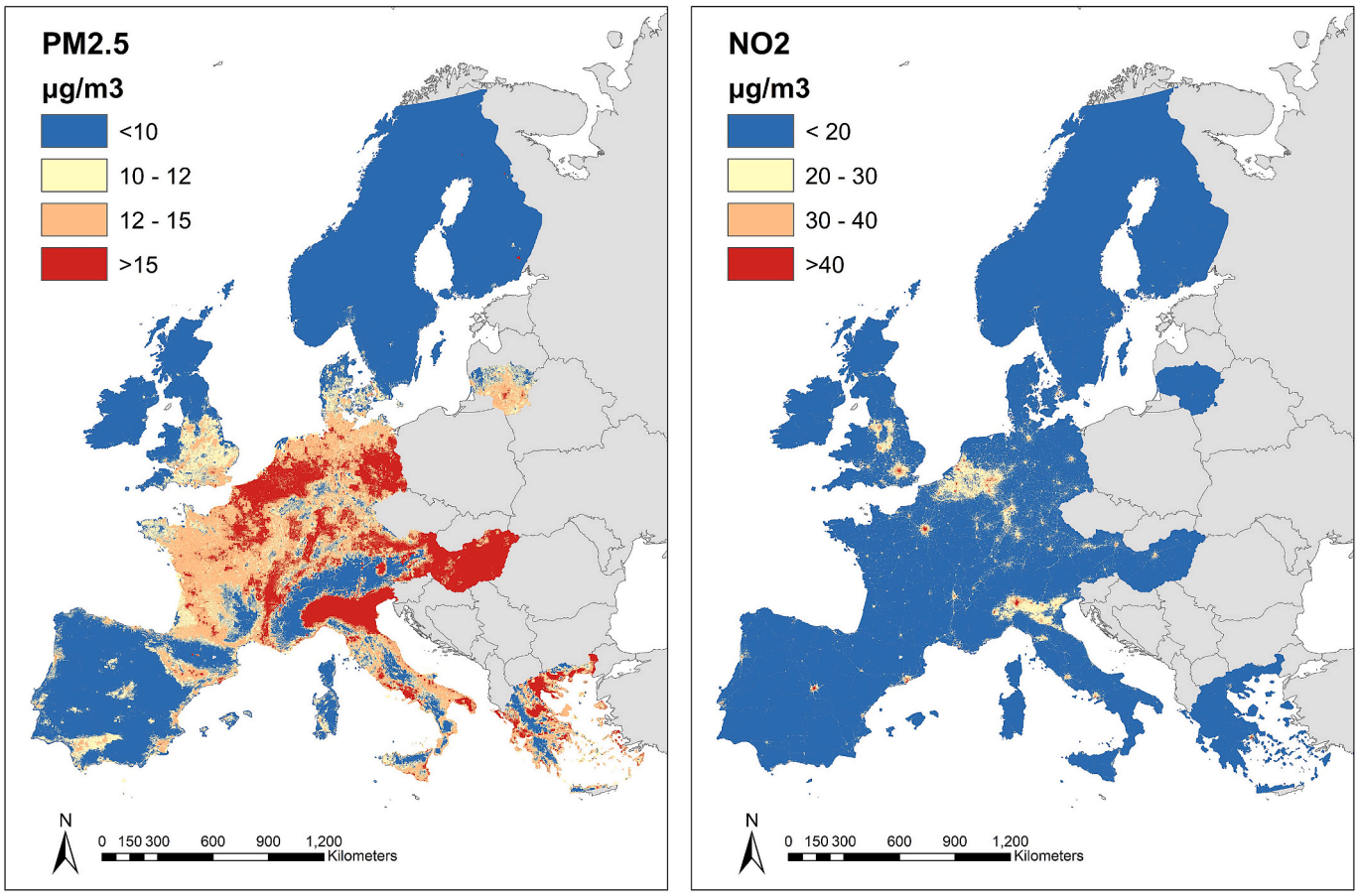


Fig. 1. Modeled annual mean concentrations for 2010 for $PM_{2.5}$ (left) and NO_2 (right) for the European region.

Table 2
Hazard ratios (and 95% confidence intervals) of the association between residential air pollutants above versus below the specified cut-off values and natural-cause mortality in the pooled cohort[a]. Results from traditional Cox regression, main causal model and sensitivity analysis.

	$PM_{2.5}$		NO_2	
Model	15 µg/m^3	12 µg/m^3	30 µg/m^3	20 µg/m^3
% above/below cut-off	51/49	86/14	41/59	69/31
Traditional Cox[b]	1.10 (1.08, 1.13)	1.17 (1.13, 1.21)	1.08 (1.05, 1.12)	1.12 (1.09, 1.14)
Main Causal[c]	1.22 (1.20, 1.25)	1.15 (1.11, 1.19)	1.02 (0.99, 1.06)	1.07 (1.05, 1.09)
Sensi Causal[d]	1.21 (1.18, 1.23)	1.16 (1.12, 1.20)	1.04 (1.01, 1.07)	1.08 (1.06, 1.10)

[a] Subcohorts for cut-off levels $PM_{2.5}$ 15 µg/m^3 (DCH, DNC-1993, DNC-1999, E3N, KORA-S3, KORA-S4, VHM_PP), $PM_{2.5}$ 12 µg/m^3 (DCH, DNC-1993, DNC-1999, VHM_PP), NO_2 30 µg/m^3 (CEANS-SALT, CEANS-SIXTY, CEANS-SNACK, DCH, DNC-1993, DNC-1999, E3N, EPIC_NL-Morgen, EPIC_NL-Prospect), NO_2 20 µg/m^3 (CEANS-SALT, CEANS-SDPP, CEANS-SIXTY, CEANS-SNACK, DCH, DNC-1993, DNC-1999, E3N, KORA-S3, KORA-S4, VHM_PP).

[b] Main model as in Strak et al., 2021 but sex and subcohort id as factors and continuous pollutants replaced with binary pollutants as factors.

[c] Main causal model included only subcohort id and the pollutant as factors weighted by IPW truncated at the 1st and 99th quantiles.

[d] Sensitivity causal model included only subcohort id and the pollutant as factors weighted by IPW truncated at the 5th and 95th quantiles.

cohorts and the meta-analytical results indicated only a weak association for $PM_{2.5}$ but a clear signal for NO_2.

The ELAPSE project was specifically designed to investigate the exposure response functions in the low exposure ranges (Brunekreef et al., 2021). However, with this analysis we aimed to explore a causal inference method that allows standardized application and testing in all ELAPSE cohorts to provide a comprehensive picture across the different cohort settings. We therefore decided for an IPW approach using binary exposures (Makar et al., 2017) over others, more advanced methods using continuous exposures (Wu et al., 2020; Higbee et al., 2020; Naimi et al., 2014). Makar and colleagues applied the binary IPW approach by contrasting $PM_{2.5}$ concentrations above vs. below the current US standard of 12 µg/m^3 in an US cohort of Medicare beneficiaries. They observed no association with all-cause mortality (HR 0.97; 95% CI 0.90–1.04) but a significant association with all-cause hospitalizations (1.07; 1.03–1.10) (Makar et al., 2017). Yet, they had only one year of follow-up and the number of deaths was comparably small. When restricting the cohort to individuals with exposure levels below 12 µg/m^3, they found an indication for associations with mortality (HR 1.11; 95% CI 0.97–1.28) and the HR for hospitalizations was doubled (1.15; 1.08–1.23) for a cut-off value of 8 µg/m^3. In comparison, we observed a HR of 1.15 (1.11–1.19) in the pooled cohort and HRs ranging from 0.97 (0.95–0.99) to 1.09 (1.07–1.12) in the administrative cohorts. However, not all cohorts could be included for all cut-off values since no or not enough data was available above or below the selected cut-off values. Only the Norwegian cohort allowed to investigate a lower cut-off value of 10 µg/m^3, for which we observed a HR of 1.03 (1.03–1.04) whereas Makar et al. reported a HR of 1.09 (0.98–1.22) for all-cause mortality.

Several papers using advanced causal modeling techniques for continuous exposures compared their approaches with traditional methods (Dominici et al., 2022; Wu et al., 2020; Higbee et al., 2020). Higbee and colleagues investigated $PM_{2.5}$ associations with all-cause

Table 3
Hazard ratios (and 95% confidence intervals) of the association between residential air pollutants above versus below the specified cut-off values and natural-cause mortality in the administrative cohorts. Results from traditional Cox regression, main causal model and sensitivity analysis.

		$PM_{2.5}$		NO_2	
Cohort	Model	15 μg/m^3	12 μg/m^3	40 μg/m^3	20 μg/m^3
Norwegian	% above/ below cut-off	0/100	6/94	0/100	26/74
	Traditional Cox[a]	–	1.03 (1.01, 1.04)	–	1.04 (1.03, 1.04)
	Main Causal[b]	–	1.09 (1.07, 1.12)	–	1.03 (1.02, 1.04)
	Sensi Causal[c]	–	1.07 (1.06, 1.08)	–	1.02 (1.01, 1.03)
Danish	% above/ below cut-off	7/93	59/41	0/100	43/57
	Traditional Cox[a]	1.04 (1.03, 1.05)	1.07 (1.07, 1.08)	–	1.11 (1.10, 1.12)
	Main Causal[b]	1.08 (1.04, 1.12)	1.06 (1.05, 1.08)	–	1.10 (1.08, 1.12)
	Sensi Causal[c]	1.11 (1.08, 1.15)	1.07 (1.05, 1.08)	–	1.09 (1.07, 1.11)
English	% above/ below cut-off	7/93	81/19	0/100	87/13
	Traditional Cox[a]	0.98 (0.96, 1.00)	1.01 (1.00, 1.03)	–	1.05 (1.03, 1.07)
	Main Causal[b]	0.96 (0.94, 0.98)	0.98 (0.96, 0.99)	–	1.03 (1.01, 1.05)
	Sensi Causal[c]	0.99 (0.97, 1.01)	0.98 (0.96, 0.99)	–	1.05 (1.03, 1.07)
Dutch	% above/ below cut-off	82/18	100/0	11/89	94/6
	Traditional Cox[a]	1.03 (1.02, 1.03)	–	1.00 (0.99, 1.01)	1.06 (1.04, 1.07)
	Main Causal[b]	1.03 (1.02, 1.05)	–	1.02 (1.01, 1.04)	1.05 (1.03, 1.07)
	Sensi Causal[c]	1.04 (1.03, 1.05)	–	1.04 (1.03, 1.05)	1.04 (1.03, 1.05)
Belgian	% above/ below cut-off	97/3	100/0	11/89	96/4
	Traditional Cox[a]	–	–	0.99 (0.98, 1.00)	–
	Main Causal[b]	–	–	0.98 (0.97, 0.99)	–
	Sensi Causal[c]	–	–	0.99 (0.98, 1.00)	–
Swiss	% above/ below cut-off	73/27	94/6	0/100	68/32
	Traditional Cox[a]	1.01 (1.01, 1.02)	1.01 (1.00, 1.02)	–	1.04 (1.03, 1.05)
	Main Causal[b]	1.01 (1.00, 1.02)	0.97 (0.96, 0.99)	–	1.03 (1.03, 1.04)
	Sensi Causal[c]	0.99 (0.99, 1.00)	0.96 (0.95, 0.96)	–	1.02 (1.02, 1.03)
Roman	% above/ below cut-off	97/3	100/0	10/90	97/3
	Traditional Cox[a]	–	–	1.02 (1.01, 1.03)	–
	Main Causal[b]	–	–	1.01 (0.99, 1.03)	–
	Sensi Causal[c]	–	–	0.99 (0.97, 1.02)	–

[a] Main model as in Stafoggia et al., 2022 but sex as factors and continuous pollutants replaced with binary pollutants as factors.
[b] Main causal model included only the pollutant as factors weighted by IPW truncated at the 1st and 99th quantiles.
[c] Sensitivity causal model included the pollutant as factors weighted by IPW truncated at the 5th and 95th quantiles.

and cardiopulmonary mortality in the US National Health Interview Survey and tested multiple distributions and weight generation techniques (Higbee et al., 2020). Their results generally provided consistent and robust estimates that were relatively insensitive to the choice of IPW. Also, the estimates were similar to those of an unweighted multivariable Cox model, although they indicated marginally lower point estimates and higher standard errors. Another study from the US investigating long-term $PM_{2.5}$ exposure and all-cause mortality among more than 68.5 million Medicare enrollees (65 years of age or older) compared two traditional approaches (Cox models, Poisson regression) with three causal inference methods

applying a generalized propensity score (GPS) to adjust for confounding by i) matching by GPS, ii) weighting by GPS, and iii) adjustment by GPS (Dominici et al., 2022; Wu et al., 2020). While all approaches lead to comparable estimates in the full population, the traditional models showed considerably higher estimates when restricting the population to $PM_{2.5}$ exposures below 12 μg/m^3.

Our study has several strengths. We combined results from eight well-characterized adult cohorts and seven large administrative cohorts from eleven European countries, with more than 28 million participants included in the analysis. The multicenter nature and inclusion of European-wide cohorts reduced the likelihood of bias from spatial patterns in unmeasured covariates that correspond to air pollution patterns. While the well-characterized adult cohorts offered a broad range of potential confounders and allowed pooled analyses, the registry based administrative cohorts had less selective recruitment and dropout compared with the traditional cohort studies. We used a previously applied straightforward causal inference approach that allowed simple assumptions testing and interpretation as well as application in all ELAPSE cohorts due to its low computational complexity. Moreover, we also investigated the sensitivity of our results by applying more stringent weights truncation and traditional Cox models with binary exposures both resulting in comparable estimates.

A major limitation of our study is the dichotomization of continuous exposures which was prerequisite to apply the approach suggested by Makar and colleagues (Makar et al., 2017). However, in our previous traditional Cox analyses we comprehensively investigated and described the full range of the continuous exposures and their respective health effects (Brunekreef et al., 2021; Stafoggia et al., 2022; Wolf et al., 2021). The main aim of this analysis was therefore not to duplicate our previous results, but to complement them by applying an alternative analysis approach that allowed general application in all our cohorts (though not for all selected cut-off values) and straightforward interpretation. Adequate confounder control is one of the main assumptions of causal inference methods but also traditional approaches. We cannot rule out the risk of residual confounding from unmeasured covariates, especially in the large administrative cohorts. However, we ran a set of sensitivity

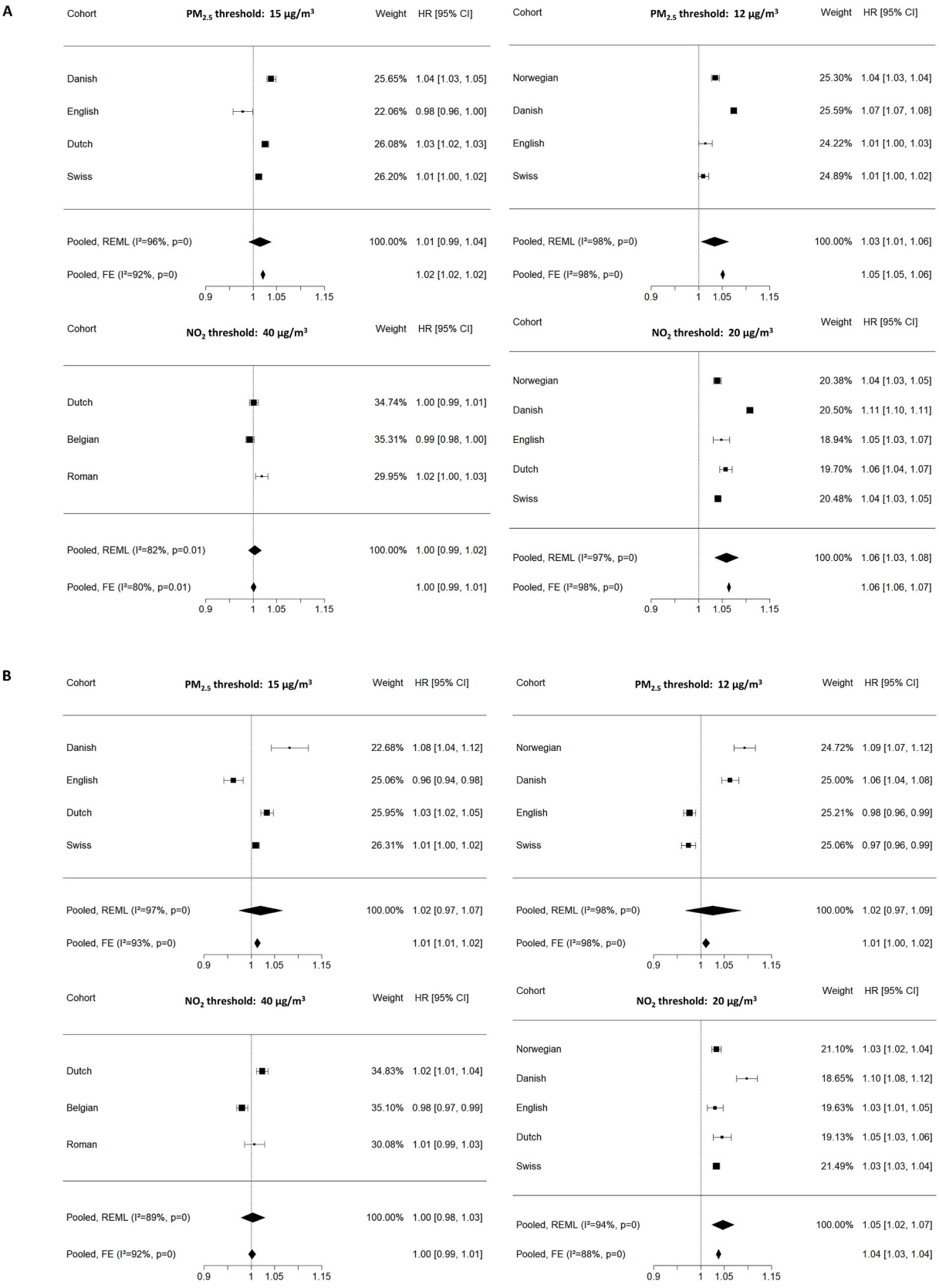


Fig. 2. Cohort-specific and meta-analytical associations between air pollutants above vs below the given cut-off values and natural-cause mortality derived from (A) traditional Cox model and (B) causal Cox models weighted by IPW. The size of the squares is proportional to the cohort-specific weight in the meta-analysis. Diamonds are centered on the point estimate and extend to the 95% CIs.

analyses including additional covariate adjustment in those cohorts with respective information available and did not observe substantial deviations from our main results (Strak et al., 2021; Stafoggia et al., 2022). In the administrative cohorts, we previously applied indirect adjustment which generally showed robust results (Stafoggia et al., 2022). Moreover, the investigation of the interplay of the exposures was out of the scope of this paper which focused on the comparison of two different statistical methods adjusting for potential confounders. We documented two-pollutant effects of the traditional Cox models with continuous exposures previously (Strak et al., 2021; Stafoggia et al., 2022). Causal methods for two- or multipollutant models are not yet established and should be investigated in future analyses.

5. Conclusion

In conclusion, our results provide further evidence that traditional and causal modelling methods produce robust results for adverse health effects of air pollution. Thus, the application of multiple modelling methods and documentation of the robustness of the results might help to improve causal inference.

Author statement

Kathrin Wolf: Conceptualization, Methodology, Writing - Original Draft, Review & Editing. **Sophia Rodopoulou:** Methodology, Software, Formal analysis, Writing - Review & Editing. **Jie Chen:** Methodology, Formal analysis, Writing - Review & Editing. **Zorana J Andersen:** Investigation, Writing - Review & Editing. **Richard W Atkinson:** Investigation, Formal analysis, Writing - Review & Editing. **Mariska Bauwelinck:** Investigation, Formal analysis, Writing - Review & Editing. **Nicole A H Janssen:** Investigation, Writing - Review & Editing. **Doris Tove Kristoffersen:** Investigation, Formal analysis, Writing - Review & Editing. **Youn-Hee Lim:** Investigation, Formal analysis, Writing - Review & Editing. **Bente Oftedal:** Investigation, Writing - Review & Editing. **Maciek Strak:** Investigation, Formal analysis, Writing - Review & Editing. **Danielle Vienneau:** Investigation, Formal analysis, Writing - Review & Editing. **Jiawei Zhang:** Investigation, Formal analysis, Writing - Review & Editing. **Bert Brunekreef:** Methodology, Writing - Review & Editing. **Gerard Hoek:** Methodology, Writing - Review & Editing. **Massimo Stafoggia:** Conceptualization, Methodology, Formal analysis, Writing - Original Draft, Review & Editing. **Evangelia Samoli:** Conceptualization, Methodology, Writing - Original Draft, Review & Editing.

Sources of financial support

This work was supported by Health Effects Institute (HEI) research agreement (grant No 4954-RFA14-3/16-5-3). Research described in this article was conducted under contract to the HEI, an organisation jointly funded by the US Environmental Protection Agency (EPA) (assistance award No R-82811201) and certain motor vehicle and engine manufacturers. The contents of this article do not necessarily reflect the views of HEI, or its sponsors, nor do they necessarily reflect the views and policies of the EPA or motor vehicle and engine manufacturers.

Sharing of data and computing code

The exposure maps are available on request from Kees de Hoogh (c.dehoogh@swisstph.ch). The cohort data could not be shared among the ELAPSE project members including named authors, nor can the data be shared externally due to strict national data protection regulations and the General Data Protection Regulation of the EU. The ELAPSE study protocol is available online (http://www.elapseproject.eu/). A detailed statistical analysis plan is available from the corresponding author on reasonable request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

We thank Marjan Tewis for compiling the pooled cohort. We acknowledge the services of Statistics Norway and the Norwegian Cause of Death Registry.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2023.121515.

References

Brauer, M., Brook, J.R., Christidis, T., Chu, Y., Crouse, D.L., Erickson, A., Hystad, P., Li, C., Martin, R.V., Meng, J., et al., 2022. Mortality–Air Pollution Associations in Low Exposure Environments (MAPLE): Phase 2, p. 119.

Brunekreef, B., Strak, M., Chen, J., Andersen, Z.J., Atkinson, R., Carey, I., Cesaroni, G., Forastiere, F., Fecht, D., Gulliver, J., et al., 2021. Mortality and Morbidity Effects of Long-Term Exposure to Low-Level PM2.5, BC, NO2, and O3: an Analysis of European Cohorts in the ELAPSE Project. in, p. 150.

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope 3rd, C.A., Apte, J.S., Brauer, M., Cohen, A., Weichenthal, S., et al., 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proc. Natl. Acad. Sci. U. S. A. 115, 9592–9597. https://doi.org/10.1073/pnas.1803222115.

Chen, J., Hoek, G., 2020. Long-term exposure to PM and all-cause and cause-specific mortality: a systematic review and meta-analysis. Environ. Int. 143, 105974 https://doi.org/10.1016/j.envint.2020.105974.

Cole, S.R., Hernan, M.A., 2008. Constructing inverse probability weights for marginal structural models. Am. J. Epidemiol. 168, 656–664. https://doi.org/10.1093/aje/kwn164.

de Hoogh, K., Chen, J., Gulliver, J., Hoffmann, B., Hertel, O., Ketzel, M., Bauwelinck, M., van Donkelaar, A., Hvidtfeldt, U.A., Katsouyanni, K., et al., 2018. Spatial PM2.5, NO2, O3 and BC models for western Europe – evaluation of spatiotemporal stability. Environ. Int. 120, 81–92. https://doi.org/10.1016/j.envint.2018.07.036.

Dominici, F., Zanobetti, A., Schwartz, J., Braun, D., Sabath, B., Wu, X., 2022. In: Inst, R. R.H.E. (Ed.), Assessing Adverse Health Effects of Long- Term Exposure to Low Levels of Ambient Air Pollution: Implementation of Causal Inference Methods, vol. 79.

Grambsch, P.M., Therneau, T.M., 1994. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 81, 515–526. https://doi.org/10.1093/biomet/81.3.515.

Greifer, N., 2022. Covariate balance tables and plots. Guide. Cobalt Pack. Accessed 11/12/2022 https://cran.r-project.org/web/packages/cobalt/vignettes/cobalt.html.

Health risks of air pollution in Europe – HRAPIE project. In: Recommendations for Concentration–Response Functions for Cost–Benefit Analysis of Particulate Matter, Ozone and Nitrogen Dioxide, 2013. WHO Regional Office for Europe, Copenhagen.

Higbee, J.D., Lefler, J.S., Burnett, R.T., Ezzati, M., Marshall, J.D., Kim, S.Y., Bechle, M., Robinson, A.L., Pope 3rd, C.A., 2020. Estimating long-term pollution exposure effects through inverse probability weighting methods with Cox proportional hazards models. Environ. Epidemiol. 4, e085. https://doi.org/10.1097/EE9.0000000000000085.

Hvidtfeldt, U.A., Severi, G., Andersen, Z.J., Atkinson, R., Bauwelinck, M., Bellander, T., Boutron-Ruault, M.-C., Brandt, J., Brunekreef, B., Cesaroni, G., et al., 2021. Long-term low-level ambient air pollution exposure and risk of lung cancer – a pooled analysis of 7 European cohorts. Environ. Int. 146. https://doi.org/10.1016/j.envint.2020.106249.

Little, R.J., Rubin, D.B., 2000. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. Annu. Rev. Publ. Health 21, 121–145. https://doi.org/10.1146/annurev.publhealth.21.1.121.

Makar, M., Antonelli, J., Di, Q., Cutler, D., Schwartz, J., Dominici, F., 2017. Estimating the causal effect of low levels of fine particulate matter on hospitalization. Epidemiology 28, 627–634. https://doi.org/10.1097/EDE.0000000000000690.

Naimi, A.I., Moodie, E.E., Auger, N., Kaufman, J.S., 2014. Constructing inverse probability weights for continuous exposures: a comparison of methods. Epidemiology 25, 292–299. https://doi.org/10.1097/ede.0000000000000053.

Pearce, N., Vandenbroucke, J.P., Lawlor, D.A., 2019. Causal inference in environmental epidemiology: old and new approaches. Epidemiology 30, 311–316. https://doi.org/10.1097/EDE.0000000000000987.

Rubin, D.B., 2008. For objective causal inference, design trumps analysis. Ann. Appl. Stat. 2 https://doi.org/10.1214/08-aoas187.

Samoli, E., Rodopoulou, S., Hvidtfeldt, U.A., Wolf, K., Stafoggia, M., Brunekreef, B., Strak, M., Chen, J., Andersen, Z.J., Atkinson, R., et al., 2021. Modeling multi-level survival data in multi-center epidemiological cohort studies: applications from the ELAPSE project. Environ. Int. 147, 106371 https://doi.org/10.1016/j.envint.2020.106371.

Stafoggia, M., Oftedal, B., Chen, J., Rodopoulou, S., Renzi, M., Atkinson, R.W., Bauwelinck, M., Klompmaker, J.O., Mehta, A., Vienneau, D., et al., 2022. Long-term exposure to low ambient air pollution concentrations and mortality among 28 million people: results from seven large European cohorts within the ELAPSE project. Lancet Planet. Health 6, e9–e18. https://doi.org/10.1016/s2542-5196(21)00277-1.

Strak, M., Weinmayr, G., Rodopoulou, S., Chen, J., de Hoogh, K., Andersen, Z.J., Atkinson, R., Bauwelinck, M., Bekkevold, T., Bellander, T., et al., 2021. Long term exposure to low level air pollution and mortality in eight European cohorts within the ELAPSE project: pooled analysis. Br. Med. J. 374, n1904. https://doi.org/10.1136/bmj.n1904.

Therneau, T.M., Grambsch, P.M., 2000. Modeling Survival Data: Extending the Cox Model. Paper/Poster Presented.

Veroniki, A.A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., Kuss, O., Higgins, J.P., Langan, D., Salanti, G., 2016. Methods to estimate the between-study variance and its uncertainty in meta-analysis. Res. Synth. Methods 7, 55–79. https://doi.org/10.1002/jrsm.1164.

Wolf, K., Hoffmann, B., Andersen, Z.J., Atkinson, R.W., Bauwelinck, M., Bellander, T., Brandt, J., Brunekreef, B., Cesaroni, G., Chen, J., et al., 2021. Long-term exposure to low-level ambient air pollution and incidence of stroke and coronary heart disease: a pooled analysis of six European cohorts within the ELAPSE project. Lancet Planet. Health 5, e620–e632. https://doi.org/10.1016/S2542-5196(21)00195-9.

World Health Organization (WHO), 2021. WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. In: World Health Organization. Geneva.

Wu, X., Braun, D., Schwartz, J., Kioumourtzoglou, M.A., Dominici, F., 2020. Evaluating the impact of long-term exposure to fine particulate matter on mortality among the elderly. Sci. Adv. 6, eaba5692 https://doi.org/10.1126/sciadv.aba5692.