**Long-term exposure to source-specific fine particles and mortality – a pooled analysis of 14 European cohorts within the ELAPSE project**

Jie Chen a, Gerard Hoek a, Kees de Hoogh b, c, Sophia Rodopoulou d, Zorana J. Andersen e, Tom Bellander f, g, Jørgen Brandt h, i, Daniela Fecht j, Francesco Forastiere k, l, John Gulliver m, j, Ole Hertel n, Barbara Hoffmann o, Ulla Arthur Hvidtfeldt p, W.M.Monique Verschuren q, r, Karl-Heinz Jöckel s, Jeanette T Jørgensen e, Klea Katsouyanni d, l, Matthias Ketzel h, t, Boel Brynedal u, Karin Leander v, Shuo Liu e, Petter Ljungman v, w, Elodie Faure x, Patrik K.E. Magnusson y, Gabriele Nagel z, Göran Pershagen v, aa , Annette Peters ab, ac, Ole Raaschou-Nielsen p, h, Debora Rizzuto ad, Evangelia Samoli d, Yvonne T. van der Schouw r, Sara Schramm ae, Gianluca Severi af, ag, Massimo Stafoggia k, v, Maciej Strak a, q, Mette Sørensen p, ah, Anne Tjønneland p, ai, Gudrun Weinmayr z, Kathrin Wolf ab, Emanuel Zitt aj, ak, Bert Brunekreef a\*, George D. Thurston al\*

a Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, the Netherlands

b Swiss Tropical and Public Health Institute, Basel, Switzerland

c University of Basel, Basel, Switzerland

d Department of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, Athens, Greece

e Section of Environment and Health, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

f Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

g Centre for Occupational and Environmental Medicine, Region Stockholm, Stockholm, Sweden

h Department of Environmental Science, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark

i iClimate – interdisciplinary Center for Climate Change, Aarhus University, Denmark

j MRC Centre for Environment and Health, School of Public Health, Imperial College London, London, UK

k Department of Epidemiology, Lazio Region Health Service / ASL Roma 1, Rome, Italy

l Environmental Research Group, School of Public Health, Imperial College London, London, UK

m Centre for Environmental Health and Sustainability & School of Geography, Geology and the Environment, University of Leicester, Leicester, UK

n Department of Ecoscience, Aarhus University, Roskilde, Denmark

o Institute for Occupational, Social and Environmental Medicine, Centre for Health and Society, Medical Faculty, Heinrich Heine University Düsseldorf, Germany

p Danish Cancer Society Research Center, Copenhagen, Denmark

q National Institute for Public Health and the Environment, Bilthoven, the Netherlands

r Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands.

s Institute for Medical Informatics, Biometry and Epidemiology, Medical Faculty, University of Duisburg-Essen, Essen, Germany

t Global Centre for Clean Air Research (GCARE), University of Surrey, Guildford GU2 7XH, United Kingdom

u Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

v Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

w Department of Cardiology, Danderyd University Hospital, Stockholm, Sweden

x University Paris-Saclay, UVSQ, Inserm, Gustave Roussy, "Exposome and Heredity" team, CESP UMR1018, 94805, Villejuif, France

y Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

z Institute of Epidemiology and Medical Biometry, Ulm University, Ulm, Germany

aa Centre for Occupational and Environmental Medicine, Region Stockholm, Stockholm, Sweden

ab Institute of Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany

ac Chair of Epidemiology, Ludwig Maximilians Universität München, Munich, Germany

ad Department of Neurobiology, Care Sciences, and Society, Karolinska Institutet, Stockholm, Sweden

Stockholm Gerontology Research Center, Stockholm, Sweden.

ae Institute for Medical Informatics, Biometry and Epidemiology, University of Duisburg-Essen, University Hospital Essen, Germany

af University Paris-Saclay, UVSQ, Inserm, Gustave Roussy, "Exposome and Heredity" team, CESP UMR1018, 94805, Villejuif, France

ag Department of Statistics, Computer Science and Applications “G. Parenti” (DISIA), University of Florence, Italy

ah Department of Natural Science and Environment, Roskilde University, Roskilde, Denmark

ai Diet, Genes and Environment (DGE)

aj Agency for Preventive and Social Medicine (aks), Bregenz, Austria

ak Department of Internal Medicine 3, LKH Feldkirch, Feldkirch, Austria

al Departments of Environmental Medicine and Population Health, Nelson Institute of Environmental Medicine, New York University School of Medicine, Tuxedo, New York, USA

\*These authors contributed equally to this work

Corresponding author: j.chen1@uu.nl

**Acknowledgments**

We thank Marjan Tewis for the data management tasks in creating the pooled cohort database.

**Declaration of interests**

None

**Funding information**

The research described in this article was conducted under contract to the Health Effects Institute (HEI), an organization jointly funded by the United States 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.

**Abstract**

We assessed mortality risks associated with source-specific fine particles (PM2.5) in a large pooled European cohort with 323,782 participants. Cox proportional hazards models were applied to estimate mortality hazard ratios (HRs) for PM2.5 sources identified through a source apportionment analysis. Exposure to 2010 annual average concentrations of source-specific PM2.5 components were assessed at baseline residential addresses. The source apportionment resulted in a plausible identification of five sources, namely traffic, residual oil combustion, soil, biomass & agriculture and industry. In single source analysis, all identified sources were significantly positively associated with increased natural mortality risks. In multi-source analysis, associations with traffic, oil and biomass & agriculture components attenuated but remained statistically significant; associations with soil and industrial components attenuated to unity. On a per interquartile increase basis, the highest association was observed for the traffic component (HR 1.06; 95% CI: 1.04, 1.08 per 2.86-µg/m3 increase) across five identified sources. On a per 1-µg/m3 comparable basis, the residual oil PM2.5 had the highest association (HR 1.13; 95% CI: 1.05, 1.22), which is roughly ten times higher than that for generic PM2.5 mass. Source-specific associations with cause-specific mortality were in general consistent with findings for natural mortality.

**Keywords**

Source apportionment; fine particulate matter (PM2.5); absolute principle component analysis (APCA); mortality

1. **Introduction**

Epidemiological studies around the world have generally reported associations between fine particles (PM2.5) exposure and mortality and morbidity, with variations in the magnitude of effect estimates 1. Part of these discrepancies may be related to the fact that the composition of PM2.5 mass varies in time and space, depending on sources of emission and atmospheric chemistry, which may in turn result in differences in toxicity and risk to health of PM2.5 mass 2-5. Understanding which components of the PM mixture are of greater health concerns than others would help inform targeted policies to control PM2.5 from those sources that contribute most of the toxic components in the PM mixture, as well as allow more accurate assessments of source-specific health impacts.

So far, mixed findings have been reported for associations between health outcomes and long-term exposure to a series of PM2.5 constituents, including secondary inorganic aerosols (sulfate, nitrate), black carbon (BC), metals and organic components 6-11. In our previous analyses within the Effects of Low-level Air Pollution: A Study in Europe (ELAPSE), we found that vanadium (V) within PM2.5 was most consistently associated with increased mortality risks in the pooled cohort 12, whereas potassium (K) and silicon (Si) were most robustly associated with natural mortality in six large administrative cohorts 13. However, analyses on a constituent basis may be difficult to interpret because individual elements can be linked to one or more specific sources (e.g., differing by location-specific source mixtures) and thus have different associated health effects depending on what aerosol mixture they are travelling with (i.e., which source group they are in), while several co-varying elemental markers may more reliably indicate the same source 14-16. For example, iron (Fe) may individually come from either wind-blown soil or a steel mill operation, but it is associated with different elements, depending on which source (e.g., with Si for soil, Mn for steel)17. Looking at Fe in a source-specific tracer group can differentiate the two mixture situations. This factor may help explain inconsistencies found across constituent-based studies in past PM2.5 constituent evaluations (e.g., US. EPA, 2021)18.

Observed health associations with individual PM2.5 trace elements are not necessarily causal, but may actually be serving as markers of associations by the source-specific mixture they travel with. Furthermore, results for two-pollutant model in the elemental context are difficult to interpret when the pollutants simultaneously included in the model are highly correlated (e.g., including nitrogen dioxide (NO2) in a copper (Cu) model would weaken the association with Cu because both pollutants are related to traffic emissions). Instead, assessing health effects of PM components in the group context (e.g., source groupings of tracers) may provide more consistent and interpretable results across studies that are more readily translatable into air quality policy. An analysis within the American Cancer Society Cancer Prevention Study II (ACS CPS-II) suggested exposure to coal combustion-related air pollution explained most associations between PM2.5 mass and increased risk of mortality from all causes, ischemic heart diseases (IHD) and lung cancer 19,20. In the California Teachers study, associations with IHD mortality were observed for sources including gas- and diesel-fueled vehicles, meat cooking and high-sulfur fuel combustion 6. The National Particle Component Toxicity (NPACT) studies identified that secondary sulfate and traffic sources were most consistently associated with adverse health effects 19,21.

In the present study, we performed a further analysis within the ELAPSE pooled cohort to consider the mortality risks associated with long-term PM2.5 exposure on a source-specific basis. By comparing these source-specific results with our previous individual elemental analyses, we expected to have a more complete understanding of the health effects of PM2.5 mixtures.

1. **Materials and Methods**

**2.1 Study population**

The ELAPSE pooled cohort consists of 14 sub-cohorts across six European countries. Detailed information on individual sub-cohorts have been extensively reported 12,22,23. The included sub-cohorts are: *Cardiovascular Effects of Air Pollution and Noise in Stockholm* (CEANS) cohort in Sweden, including *Stockholm Diabetes Prevention Program* (SDPP)24, the *Stockholm Cohort of 60-year-olds* (SIXTY) 25, *Stockholm Screening Across the Lifespan Twin study* (SALT) 26 and *Swedish National Study on Aging and Care in Kungsholmen* (SNACK) 27; the *Diet, Cancer and Health cohort* (DCH) 28 in Denmark; the *Danish Nurse Cohort* (DNC) 29 in Denmark, consisting at baseline of two surveys conducted in 1993 and 1999; the *European Prospective Investigation into Cancer and Nutrition-Netherlands* (EPIC-NL) cohort in the Netherlands, including the *Monitoring Project on Risk Factors* and *Chronic Diseases in the Netherlands* (MORGEN) and Prospect 30; the *Heinz Nixdorf Recall study* (HNR) in Germany 31; the *Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l’Education Nationale* (E3N) in France 32; the *Cooperative Health Research in the Region of Augsburg* (KORA) in Germany, consisting at baseline of two cross-sectional population-representative surveys conducted in 1994–1995 (S3) and 1999–2001 (S4); and the *Vorarlberg Health Monitoring and Prevention Programme* (VHM&PP) in Austria 33. Most cohorts covered a large city and its surrounding areas as study areas. The French E3N cohort and the Danish DNC cohort covered large regions of the country. All included cohort studies were approved by the medical ethics committees in their respective countries.

**2.2 Source apportionment and exposure assessment**

Air pollution measurements for PM2.5 mass, NO2, black carbon (BC) and PM2.5 elemental composition were derived from the ESCAPE monitoring campaign conducted in 19 study areas across Europe. Sampling and analysis methods have been described before 15,34. Briefly, measurements were made at 20 sites in each study area (40 in the large Catalunya and Netherlands/Belgium areas) for three 2-week periods in a 1-year period between October 2008 and April 2011. Monitoring sites were selected to represent pollution levels at regional background, urban background, and street locations using a common sampling protocol. Eight components were *a priori* selected in the ESCAPE study: copper (Cu), Fe, K, nickel (Ni), sulfur (S), Si, V and zinc (Zn) 15,35. Annual average concentrations were calculated based on three 2-week average measurements spread over the seasons (warm, cold and intermediate) with temporal adjustment from a reference background site in each study area. Table S1 documents the distribution of air pollution measurements.

The annual average concentrations of air pollutants were analyzed to estimate source apportioned PM2.5 mass exposures using the absolute principal component analysis (APCA)36. The method involved a) applying PCA to the pollutant data; b) identifying source-related components based on key tracers in each component; c) adjusting PC scores into absolute PC scores; and d) regressing PM2.5 mass on the source-related components, deriving apportionments of PM2.5 mass to each source-related component. Pollution data from a total of 397 sites were included. Before the source identification PCA, an adjusted K index of biomass burning PM2.5 was calculated by subtracting the soil dust-associated K from the total K concentration values (Kw = K - 0.42 x Si). The coefficient was calculated by regressing the K against Si concentrations for those samples with the lowest 10th percentile K/Si ratio samples 37. PM2.5 S was not included in the initial source identification analysis because S is considered as a general marker for fossil fuel sources and may complicate the apportionment of fine mass to sources 38. Also, excluding tracers of secondary formation (i.e., S) from the source apportionment analysis allows a clearer discrimination of the original primary sources of PM2.5 39. S was then apportioned among the sources by regressing it against the identified sources to calculate the ‘unexplained’ secondary mass prior to the PM2.5 mass apportionment regression models 39. Based on the APCA, source-specific compositional profiles were assessed by regressing each pollutant against the mass contributions for all sources in a linear model. Source profiles of S were similarly assessed, though S was not included in the PCA.

Individual exposure to 2010 annual mean concentrations of PM2.5 mass, NO2, BC and PM2.5 elemental composition was assessed at participants’ baseline residential addresses based on Europe-wide land use regression model estimates 40,41. The models were built on ground-based measurements with satellite-derived and chemical transport modeled air pollutant estimates, land use, road and population data as predictors. The models explained a moderate to large fraction of the measured concentration variation at the European scale (i.e., 66% for PM2.5, 58% for NO2, 51% for BC and 41% to 79% across elemental components). Exposure to NO2, BC and individual elemental components was then converted to source apportioned PM2.5 mass exposure estimates based on APCA results. This involved first converting individual exposures to the absolute PC scores for all identified source components, and then multiplied by the regression slopes derived from the PM2.5 mass measurement apportionment.

**2.3 Mortality outcome definition**

Identification of outcomes was based upon linkage to mortality registries. Based on the underlying cause of death recorded on death certificates according to *the International Classification of Diseases, Ninth Revision* (ICD-9; 42) and the *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (ICD-10, 43), we defined mortality from natural causes (ICD-9: 001–779, ICD-10: A00–R99), cardiovascular diseases (ICD-9: 400–440, ICD-10: I10–I70), non-malignant respiratory diseases (ICD-9: 460–519, ICD-10: J00–J99) and lung cancer (ICD-9: 162, ICD-10: C34).

**2.4 Statistical analyses**

Cox proportional hazard regression models were applied to evaluate associations of identified PM2.5 sources with natural and cause-specific mortality, following the general ELAPSE analytical framework applied in our previous paper of elemental exposures 12,23,44. Censoring occurred at the time of the event of interest, death from other causes, emigration, loss to follow-up for other reasons, or at the end of follow-up, whichever came first. The Cox models were stratified by sub-cohorts to account for differences in baseline hazards between the sub-cohorts unexplained by the available covariates and to relax the proportional hazards assumption. The decision to account for between sub-cohort heterogeneity with strata implies that we mostly evaluate within-cohort exposure contrasts. Three confounder models were *a priori* specified with increasing adjustment for individual and area-level covariates: Model 1 adjusted for age (as the time axis), sub-cohort (as strata), sex (as strata), and year of enrollment; Model 2 further added individual-level covariates including marital status (married/cohabiting, divorced/separated, single, widowed), smoking status (never, former, current), smoking duration (years of smoking) for current smokers, smoking intensity (cigarettes/day) for current smokers, squared smoking intensity, body mass index (BMI) categories (<18.5, 18.5–24.9, 25–29.9, and >30 kg/m2), and employment status (employed vs. unemployed); Model 3 further adjusted for neighborhood-level mean income in 2001. Model 3 was considered as the main model. Participants with missing exposure or incomplete information on Model 3 covariates were excluded from all main analyses to allow comparisons between different model results.

Individual PM2.5 sources were included as linear functions in the Cox models as a reasonable summary of the association and to facilitate comparisons with previous studies. Besides the single-source analysis where one PM2.5 source was evaluated at a time, multi-source analyses were also performed with all identified PM2.5 sources included in the model simultaneously. Cumulative risks of all identified sources were estimated assuming additive effects of combined source exposures on mortality. Cumulative Risk Index (CRI) was defined as

where are the log-hazard ratio for the P source exposures estimated at concentrations in a Cox model consisting of all P sources together 45. Hazard ratios (HRs) and 95% confidence intervals (CIs) for interquartile range (IQR) increases in the estimated concentrations of each source were reported. In addition, HRs associated with per 1 µg/m3 increase are presented.

All analyses were performed in R version 3.4.0 using packages: *survival, coxme, Matrix, foreach, multcomp, survey, splines, Hmisc, mfp, VIM, ggplot2, frailtySurv, survsim, eha, stamod, psych*. Statistical significance was based on a 95% confidence interval of effect estimate not including unity.

1. **Results and Discussion**

**3.1 Source apportionment results**

A five-factor solution using Promax rotation was chosen as the optimal solution based on an examination of both the factor eigenvalues (i.e., the data variance explained by the identified component) and the source-related interpretability of the factors. The selection of oblique rotation, instead of orthogonal rotation, allowed the identified source components to be intercorrelated with one another, which is more realistic in real-world settings. Correlations at the monitoring sites between identified sources were low to moderate (Table S2).

Table 1 provides the correlations between the identified source components and the individual air pollutants (i.e., factor loadings). PM2.5 mass and S were not included in the PCA, but their correlations with the identified factors were calculated to aid in the interpretation of the source components. Overall, the source apportionment resulted in a plausible identification of source contributions, with the possible exception of the biomass source where we did not have access to a more specific marker, such as levoglucosan. The first factor was identified as traffic-related particles because of its high loadings on NO2, BC, Cu and Fe. The principle sources of NO2 and BC include the combustion processes from motorized traffic and off-road machinery 46,47, whereas Cu and Fe are considered as markers of brake wear 14,15. The second factor was identified as particles from residual oil combustion, based on its high loadings on both Ni and V, two elements known to be enriched in heavier fuel oils 16. The third factor was identifiable as crustal/soil particles because of its high loading on Si and moderate loading on Fe. Si is a specific tracer for crustal material and Fe is also abundant in crustal dust 14. The fourth component was firstly identified as particles from biomass burning because of its high loading on Kw, which is the most common element used to trace biomass burning 14. The identification of the fourth factor is rather uncertain because of the high explained variance by S in the source profile (Figure 1) and the fact that there is no S in wood. We speculated that this factor may possibly also include windblown soil containing agricultural fertilizers because both K and S are in fertilizers widely used in Europe 48. We therefore identified the fourth factor as particles from biomass & agriculture. The last factor was identified as associated with industrial emissions because of its high loading on Zn. Although Zn was *a priori* selected in ESCAPE to represent non-tailpipe traffic emissions 15, it was not found to correlate with other traffic tracers in this analysis. However, it is also a tracer for particles from industrial sources, and predictors representing industrial Zn emission were picked up in our Europe-wide models and explained a large fraction of the variation in the Zn measurements 49.

The source compositional profiles show that S explained a large proportion of all the identified source-related mass, except for traffic-related particles (Figure 1). Sulfate is a secondary pollutant produced by atmospheric reactions of sulfur dioxide (SO2) emitted by combustion of S-containing liquid and solid fuels. The estimation of trace elements associated with each source component aids in the identification of the source and provides perspective on the mass estimates.

**3.2 Population characteristics and source-specific PM2.5 exposure estimation**

Of the total population of 381,036 subjects, 323,782 (85%) were included in the main analyses and followed up for an average of 19.5 years, contributing to 6,317,235 person-years of follow-up (Table 2). Most of the cohorts started in the mid-1990s and follow-up till 2011-2015. The average age at baseline ranged from 42 to 73. Four sub-cohorts only included female participants and the pooled cohort comprised 66% females. Differences across the cohorts were also observed for the population size, average years of follow-up, socioeconomic status (SES) and lifestyle factors, supporting our decision to account for difference in baseline hazards between sub-cohorts. Detailed baseline characteristics of study population in individual sub-cohorts can be found elsewhere 12,22,23.

Figure 2 and Table S3 show exposure distribution of the identified source-specific PM2.5 concentrations estimated for the study population in the individual sub-cohorts and the pooled cohort. Exposure concentrations generally showed a north-south increasing trend for PM2.5 from traffic and biomass & agriculture and the generic PM2.5 mass. The within-cohort exposure variability was large for PM2.5 from traffic, soil and biomass & agriculture. PM2.5 exposures from residual oil burning and industrial sources were low in all sub-cohorts with relatively small within-cohort exposure contrasts, mainly because of the lack of sources in the study areas. Exception was observed for a small number of subjects within the Dutch EPIC-NL-Morgen cohort with high exposures in PM2.5 from both residual oil burning and industrial sources, likely related to shipping emissions from port cities and emissions from steel industries 50. Soil-related PM2.5 exposures ranged similarly across sub-cohorts, except the relatively high exposures observed for participants within the SNACK cohort located in Kungsholmen, Stockholm. We estimated relatively low PM2.5 concentrations from soil and relatively high PM2.5 concentrations from biomass & agriculture. This suggested that we may not be able to completely disentangle biomass burning associated K from soil dust-associated K even with the adjustment K index calculated. However, the high estimation for biomass contribution was comparable with a PM2.5 source apportionment study conducted in Europe 51. Moreover, soil contribution in PM is predominately in the coarse fraction and the traffic-associated soil mass from road dust may have been picked up by the traffic component and thus resulted in the low contribution of this non-traffic soil component.

Correlations between source-specific PM2.5 exposures at residential addresses were moderate to low (Table S4), allowing proper interpretation of the multi-source analyses. We reported the median of cohort-specific correlations, which is most relevant for our interpretations as the analyses were stratified by sub-cohort. Correlations between sources varied across sub-cohorts. The values were not directly comparable to the correlations between identified sources at the monitoring sites presented in Table S2 as the correlations at monitoring sites were assessed at the European scale.

* 1. **Mortality risks associated with source-specific PM2.5**

During the follow-up, we observed 46,640 (14.4%), 15,492 (4.8%), 2,846 (0.9%) and 3,776 (1.2%) deaths from natural causes, cardiovascular diseases, non-malignant respiratory diseases and lung cancer respectively. Figure 3 and Table S5 show the associations between source-specific PM2.5 and natural and cause-specific mortality.

We observed significantly positive associations with traffic-related PM2.5 for all assessed mortality endpoints in single-source models. In multi-source models, associations decreased slightly for mortality from natural causes, cardiovascular diseases and lung cancer, and remained stable for respiratory mortality with slightly wider CIs. The HRs associated with a per IQR increase in source component concentrations were highest for the traffic component for all assessed endpoints except lung cancer mortality. Consistent findings were reported by the ACS CPS-II and the California Teachers Study for strong associations between excess mortality and traffic-related exposures 6,52. In the previous elemental analyses within the ELAPSE pooled cohort, we observed significantly positive associations for mortality with Cu and Fe in single-pollutant models 12. The associations attenuated substantially in models with further adjustment for NO2, likely reflecting the common traffic source of these pollutants. Similarly, positive associations with mortality were reported in studies analyzing individual tracers of traffic emissions, which became less stable after adjusting for other traffic-related components such as NO2 or organic carbon 7,8,13. The majority of the traffic-related PM2.5 was explained by BC (25.2%) in our study (Figure 1). There is mounting evidence on associations between long-term exposure to BC and adverse health outcomes 46,53. In the ELAPSE pooled cohort, we previously found significantly positive associations for BC exposure with mortality and incidence of stroke, asthma and COPD 22,54-56. Health effects of BC were however difficult to disentangle from NO2 because of their high correlation in concentrations, especially in the developed countries where an important source of both pollutants is diesel powered vehicles. The source apportionment analysis conducted in the present study allowed us to consider the multiple air pollutants from the same source as a group and thus derived more interpretable results.

The oil source component was significantly positive associated with all assessed mortality endpoints in single-source models. HRs reduced moderately and remained statistically significant (borderline significant for CVD mortality) in multi-source models. This is consistent with our previous findings that V was most robustly associated with increased mortality risks 12. When considered on a per 1 µg/m3 comparable basis, the residual oil PM2.5 had the highest associations with all assessed mortality endpoints (HR and 95% CI 1.269 (1.189, 1.354) in single-source model and 1.128 (1.047, 1.215) in multi-source model for natural mortality) (Table S6). This indicates that the natural mortality risk estimate for residual oil PM2.5 is roughly ten times higher than that for PM2.5 mass in general (HR 1.025; 95% CI: 1.021–1.030). Previous findings have been mixed on mortality risks associated with residual oil component and its trace elements 20,57,58, likely because of the small exposure contrasts combined with low concentrations of residual oil-related particles. In quite a few study areas residual oil is a less ubiquitous source than motorized traffic. The large population included in the current analysis allowed us to better detect the potential associations.

The soil component was positively associated with all assessed mortality endpoints in single source models. However, these associations reduced to basically unity in multi-source models. Crustal materials are often abundant in coarse particles. Our finding of null association is consistent with the majority of previous studies showing little evidence for an association between long-term coarse PM exposure and adverse health effects 53,59. The 2019 Integrated Science Assessment (ISA) rated the association between PMcoarse exposure and natural-cause mortality as “suggestive” 60. A recent analysis conducted in the Medicare enrollees reported significantly positive but much smaller association with all-cause mortality for soil component than for combustion-related components 61. The ACS CPS-II found no association with mortality for soil and its elemental tracers (calcium and Si) 52. Si is a specific tracer for crustal material that is a large component of soil and resuspended road dust. However, a distinction between soil and road dust is often not possible because of the overlapping source profiles 14. The California Teachers Study reported adverse cardiovascular associations with long-term exposure to Si, yet the authors interpreted the Si exposure as a proxy either for toxic constituents found in road dust or for exposures to traffic-related pollutants 7. The relatively low concentrations of soil component observed in our study suggests that the traffic-associated soil mass from road dust may have been picked up by the traffic component.

For the biomass & agriculture component, we observed significantly positive associations with mortality from natural causes, cardiovascular diseases and lung cancer, and non-significantly negative association with respiratory mortality in single-source models. In multi-source models, HRs decreased but remained significantly positive for mortality from natural causes and cardiovascular diseases and remained stable for lung cancer mortality. Different from other identified source components, HRs for biomass & agriculture increased from the crude model (Model 1) to the model with further adjustment for individual level potential confounders (Model 2) (Table S7). The increase in HR was because of the negative correlations between biomass & agriculture source component and the individual level covariates, indicating population with higher SES or healthier lifestyle tend to be exposed more to PM2.5 from biomass & agriculture. We identified the biomass & agriculture component because of its high loading on K (Table 1). However, K is not a unique indicator of wood combustion, but can also derived from meat cooking, refuse incineration and agriculture waste combustion 15,37. The California Teachers Study reported positive association between K and IHD mortality 7, whereas null association with mortality for K and biomass combustion source category was found in the ACS CPS-II and the Medicare cohort 52,57.

Industry component was positively associated with all assessed mortality outcomes in single source models, but HRs reduced to unity in multi-source models. The observed null associations may be related to the small exposure contrasts exploited in our analyses. Consistently, no association between steel industry-related PM2.5 and mortality was found in the Medicare cohort 57. The ACS CPS-II reported positive but weak associations with IHD mortality for metal industrial combustion PM2.5 (tracers Pb and Zn) 20.

Overall, the present source-specific analysis and the previous individual elemental analysis revealed that particles from residual oil burning (tracers Ni and V) and traffic-related emissions (tracers NO2, BC, Cu and Fe) were most consistently associated with mortality 12, agreeing with an analysis in the Netherlands based on dispersion model calculated particle source contributions 62.

Cumulatively, we found significantly positive associations by air pollution with mortality from natural causes, cardiovascular diseases and lung cancer. Association with non-malignant respiratory mortality was positive though non-significant. The strongest cumulative risk estimate was for lung cancer mortality. The overall cumulative risk estimates were slightly larger than the generic PM2.5 mass HRs assessed in single-source models, suggesting that we may have characterized the toxicity of the PM2.5 mixture slightly better by considering specific sources compared to generic PM2.5.

**3.4 Strengths and limitations**

One strength is the unique and standardized measurement data for PM2.5 elemental composition collected from 19 study areas across Europe used for source apportionment analysis. The number of studies on PM component-specific health effects has been small partly because of the scarcity of measurements relative to regulated pollutants such as PM2.5 and NO2. Another strength is the large population included in this study with detailed information on individual and area-level covariates. The pooling of 14 sub-cohorts and the harmonization of variables across cohorts allowed enhanced statistical power to detect source-specific associations with mortality, especially for source components that have small exposure contrasts. The source apportionment analysis allowed us to assess health effects of PM components in the group context and provided more interpretable results that are more readily translatable into generalizable air quality policy.

While it is clear from our findings that source-specific PM2.5 differ in their associations with mortality, more source-specific tracers are needed to be more definitive. We were able to include only 9 pollutants in the PCA, whereas source apportionment analyses usually include more tracers. We cannot rule out effects of sources that were not identified by our source apportionment analysis. For example, we could not identify coal burning source, likely because there was not arsenic (As) or selenium (Se) data available to consider in the PCA. Coal combustion PM2.5 was found to be most strongly and robustly associated with IHD mortality in the ACS CPS-II and the association between IHD mortality was roughly five times higher than that for generic PM2.5 mass 20. Also, more (likely organic) tracers such as levoglucosan need to be analyzed in future work to better separate out the biomass contribution 63.

We had limited ability to investigate the spatial and temporal variability of source components. The air pollution data were collected from 397 monitoring sites across Europe and were clustered (20 or 40 sites per study area). Source of trace elements may vary spatially. A sensitivity analysis by study area would be unstable and uninformative for our study because of the small number of monitoring sites within each area. Nevertheless, previous source apportionment analyses suggested that the sources of trace elements were relatively stable across European cities 14. The use of annual average concentrations instead of daily average concentrations hindered the investigation of the temporal variation of source component concentrations, by which we may be able to better identify the source components. For example, the identification of biomass & agriculture source component in our study was rather uncertain, which could be improved by investigating the temporal variation of this source concentrations (i.e., fertilizer use is more frequent in the spring-early summer planting season, whereas concentrations for particles from biomass burning source are higher in winter).

**References**

1. Chen J, Hoek G. Long-term exposure to PM and all-cause and cause-specific mortality: A systematic review and meta-analysis. *Environ Int* 2020: 105974.

2. Laden F, Neas LM, Dockery DW, Schwartz J. Association of fine particulate matter from different sources with daily mortality in six US cities. *Environmental health perspectives* 2000; **108**(10): 941-7.

3. Özkaynak H, Thurston GD. Associations between 1980 US mortality rates and alternative measures of airborne particle concentration. *Risk analysis* 1987; **7**(4): 449-61.

4. Sarnat JA, Marmur A, Klein M, et al. Fine particle sources and cardiorespiratory morbidity: an application of chemical mass balance and factor analytical source-apportionment methods. *Environmental health perspectives* 2008; **116**(4): 459-66.

5. Tsai FC, Apte MG, Daisey JM. An exploratory analysis of the relationship between mortality and the chemical composition of airborne particulate matter. *Inhalation Toxicology* 2000; **12**(sup2): 121-35.

6. Ostro B, Hu J, Goldberg D, et al. Associations of mortality with long-term exposures to fine and ultrafine particles, species and sources: results from the California Teachers Study Cohort. *Environ Health Perspect* 2015; **123**(6): 549-56.

7. Ostro B, Lipsett M, Reynolds P, et al. Long-term exposure to constituents of fine particulate air pollution and mortality: results from the California Teachers Study. *Environ Health Perspect* 2011; **118**(3): 363-9.

8. Badaloni C, Cesaroni G, Cerza F, Davoli M, Brunekreef B, Forastiere F. Effects of long-term exposure to particulate matter and metal components on mortality in the Rome longitudinal study. *Environment international* 2017; **109**: 146-54.

9. Beelen R, Hoek G, Raaschou-Nielsen O, et al. Natural-cause mortality and long-term exposure to particle components: an analysis of 19 European cohorts within the multi-center ESCAPE project. *Environ Health Perspect* 2015; **123**(6): 525-33.

10. Raaschou-Nielsen O, Beelen R, Wang M, et al. Particulate matter air pollution components and risk for lung cancer. *Environment international* 2016; **87**: 66-73.

11. Wang M, Beelen R, Stafoggia M, et al. Long-term exposure to elemental constituents of particulate matter and cardiovascular mortality in 19 European cohorts: results from the ESCAPE and TRANSPHORM projects. *Environ Int* 2014; **66**: 97-106.

12. Chen J, Rodopoulou S, de Hoogh K, et al. Long-Term Exposure to Fine Particle Elemental Components and Natural and Cause-Specific Mortality-a Pooled Analysis of Eight European Cohorts within the ELAPSE Project. *Environ Health Perspect* 2021; **129**(4): 47009.

13. Rodopoulou S, Stafoggia M, Chen J, et al. Long-term exposure to fine particle elemental components and mortality in Europe: Results from six European administrative cohorts within the ELAPSE project. *Science of The Total Environment* 2022; **809**: 152205.

14. Belis CA, Karagulian F, Larsen BR, Hopke PK. Critical review and meta-analysis of ambient particulate matter source apportionment using receptor models in Europe. *Atmospheric Environment* 2013; **69**: 94-108.

15. Tsai MY, Hoek G, Eeftens M, et al. Spatial variation of PM elemental composition between and within 20 European study areas--Results of the ESCAPE project. *Environ Int* 2015; **84**: 181-92.

16. Viana M, Kuhlbusch TAJ, Querol X, et al. Source apportionment of particulate matter in Europe: A review of methods and results. *Journal of Aerosol Science* 2008; **39**(10): 827-49.

17. Cooper JA, Watson JG. Receptor Oriented Methods of Air Particulate Source Apportionment. *Journal of the Air Pollution Control Association* 1980; **30**(10): 1116-25.

18. USEPA. Supplement to the 2019 Integrated Science Assessment for Particulate Matter.(External Review Draft). *EPA/600/R-21/198 Center for Public Health and Environmental Assessment Office of Research and Development US Environmental Protection Agency Research Triangle Park wwwepagov/isa* 2021.

19. Lippmann M, Chen LC, Gordon T, Ito K, Thurston GD. National Particle Component Toxicity (NPACT) Initiative: integrated epidemiologic and toxicologic studies of the health effects of particulate matter components. *Research Report (Health Effects Institute)* 2013; (177): 5-13.

20. Thurston GD, Burnett RT, Turner MC, et al. Ischemic Heart Disease Mortality and Long-Term Exposure to Source-Related Components of U.S. Fine Particle Air Pollution. *Environ Health Perspect* 2016; **124**(6): 785-94.

21. Vedal S, Campen M, McDonald J, et al. National Particle Component Toxicity (NPACT) initiative report on cardiovascular effects. *Research Report (Health Effects Institute)* 2013; (178): 5-8.

22. Strak M, Weinmayr G, Rodopoulou S, et al. Long term exposure to low level air pollution and mortality in eight European cohorts within the ELAPSE project: pooled analysis. *BMJ* 2021; **374**: n1904.

23. Brunekreef B, Strak M, Chen J, et al. Mortality and Morbidity Effects of Long-Term Exposure To Low-Level PM2.5, Black Carbon, NO2 and O3: An Analysis of European Cohorts. *Research Report (Health Effects Institute)* 2021.

24. Eriksson AK, Ekbom A, Granath F, Hilding A, Efendic S, Östenson CG. Psychological distress and risk of pre‐diabetes and Type 2 diabetes in a prospective study of Swedish middle‐aged men and women. *Diabetic Medicine* 2008; **25**(7): 834-42.

25. Wändell P-E, Wajngot A, De Faire U, Hellénius M-L. Increased prevalence of diabetes among immigrants from non-European countries in 60-year-old men and women in Sweden. *Diabetes & metabolism* 2007; **33**(1): 30-6.

26. Magnusson PK, Almqvist C, Rahman I, et al. The Swedish Twin Registry: establishment of a biobank and other recent developments. *Twin Research and Human Genetics* 2013; **16**(1): 317-29.

27. Lagergren M, Fratiglioni L, Hallberg IR, et al. A longitudinal study integrating population, care and social services data. The Swedish National study on Aging and Care (SNAC). *Aging clinical and experimental research* 2004; **16**(2): 158-68.

28. Tjønneland A, Olsen A, Boll K, et al. Study design, exposure variables, and socioeconomic determinants of participation in Diet, Cancer and Health: a population-based prospective cohort study of 57,053 men and women in Denmark. *Scandinavian journal of public health* 2007; **35**(4): 432-41.

29. Hundrup YA, Simonsen MK, Jørgensen T, Obel EB. Cohort profile: the Danish nurse cohort. *International journal of epidemiology* 2012; **41**(5): 1241-7.

30. Beulens JW, Monninkhof EM, Verschuren WM, et al. Cohort profile: the EPIC-NL study. *International journal of epidemiology* 2010; **39**(5): 1170-8.

31. Schmermund A, Möhlenkamp S, Stang A, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. *American heart journal* 2002; **144**(2): 212-8.

32. Clavel-Chapelon F, Group ENS. Cohort profile: the French E3N cohort study. *International journal of epidemiology* 2015; **44**(3): 801-9.

33. Ulmer H, Kelleher C, Fitz‐Simon N, Diem G, Concin H. Secular trends in cardiovascular risk factors: an age‐period cohort analysis of 6 98 954 health examinations in 1 81 350 Austrian men and women. *Journal of internal medicine* 2007; **261**(6): 566-76.

34. Eeftens M, Tsai M-Y, Ampe C, et al. Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2 – Results of the ESCAPE project. *Atmospheric Environment* 2012; **62**: 303-17.

35. de Hoogh K, Wang M, Adam M, et al. Development of land use regression models for particle composition in twenty study areas in Europe. *Environ Sci Technol* 2013; **47**(11): 5778-86.

36. Thurston GD, Spengler JD. A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston. *Atmospheric Environment (1967)* 1985; **19**(1): 9-25.

37. Rahman MM, Begum BA, Hopke PK, Nahar K, Thurston GD. Assessing the PM2.5 impact of biomass combustion in megacity Dhaka, Bangladesh. *Environ Pollut* 2020; **264**: 114798.

38. Thurston GD, Lioy PJ. Receptor modeling and aerosol transport. *Atmospheric Environment (1967)* 1987; **21**(3): 687-98.

39. Thurston GD, Ito K, Lall R. A Source Apportionment of U.S. Fine Particulate Matter Air Pollution. *Atmos Environ (1994)* 2011; **45**(24): 3924-36.

40. Chen J, de Hoogh K, Gulliver J, et al. Development of Europe-Wide Models for Particle Elemental Composition Using Supervised Linear Regression and Random Forest. *Environ Sci Technol* 2020.

41. De Hoogh K, Chen J, Gulliver J, et al. Spatial PM2. 5, NO2, O3 and BC models for Western Europe–Evaluation of spatiotemporal stability. *Environment international* 2018; **120**: 81-92.

42. WHO. International Classification of Diseases, Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death. Geneva, Switzerland: WHO. 1997.

43. WHO. International Statistical Classification of Diseases and Related Health Problems, 10th Revision. <https://icd.who.int/browse10/2016/en>. 2016.

44. Samoli E, Rodopoulou S, Hvidtfeldt UA, et al. Modeling multi-level survival data in multi-center epidemiological cohort studies: Applications from the ELAPSE project. *Environment International* 2021; **147**: 106371.

45. Crouse DL, Peters PA, Hystad P, et al. Ambient PM2.5, O(3), and NO(2) Exposures and Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environ Health Perspect* 2015; **123**(11): 1180-6.

46. Janssen NA, Hoek G, Simic-Lawson M, et al. Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM10 and PM2.5. *Environ Health Perspect* 2011; **119**(12): 1691-9.

47. WHO. Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide: World Health Organization; 2006.

48. EC. European Union: agriculture statistical factsheet. *European Commission, Available at:* [*http://eceuropaeu/agriculture/statistics/factsheets\_en*](http://eceuropaeu/agriculture/statistics/factsheets_en)2021.

49. Chen J, de Hoogh K, Gulliver J, et al. Development of Europe-Wide Models for Particle Elemental Composition Using Supervised Linear Regression and Random Forest. *Environ Sci Technol* 2020; **54**(24): 15698-709.

50. Keys A, Van Hout M, Daniels B. Decarbonisation options for the Dutch steel industry. *PBL Netherlands Environmental Assessment Agency* 2019.

51. Almeida SM, Manousakas M, Diapouli E, et al. Ambient particulate matter source apportionment using receptor modelling in European and Central Asia urban areas. *Environ Pollut* 2020; **266**(Pt 3): 115199.

52. Thurston G, Ito K, Lall R, et al. NPACT Study 4: Mortality and Long-Term Exposure to PM2. 5 and its Components in the American Cancer Society’s Cancer Prevention Study II Cohort. *National Particle Component Toxicity (NPACT) Initiative: Integrated Epidemiologic and Toxicologic Studies of the Health Effects of Particulate Matter Components Health Effect Institute* 2013.

53. Hoek G, Krishnan RM, Beelen R, et al. Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environmental health : a global access science source* 2013; **12**(1): 43.

54. Liu S, Jorgensen JT, Ljungman P, et al. Long-term exposure to low-level air pollution and incidence of chronic obstructive pulmonary disease: The ELAPSE project. *Environ Int* 2021; **146**: 106267.

55. Liu S, Jorgensen JT, Ljungman P, et al. Long-term exposure to low-level air pollution and incidence of asthma: the ELAPSE project. *Eur Respir J* 2021.

56. Wolf K, Hoffmann B, Andersen ZJ, et al. 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. *The Lancet Planetary Health* 2021; **5**(9): e620-e32.

57. Kazemiparkouhi F, Honda T, Eum KD, Wang B, Manjourides J, Suh HH. The impact of Long-Term PM2.5 constituents and their sources on specific causes of death in a US Medicare cohort. *Environ Int* 2022; **159**: 106988.

58. Wang Y, Shi L, Lee M, et al. Long-term Exposure to PM2.5 and Mortality Among Older Adults in the Southeastern US. *Epidemiology* 2017; **28**(2): 207-14.

59. Adar SD, Filigrana PA, Clements N, Peel JL. Ambient coarse particulate matter and human health: a systematic review and meta-analysis. *Current environmental health reports* 2014; **1**(3): 258-74.

60. USEPA. Integrated Science Assessment (ISA) for Particulate Matter (Final Report, 2019). *US Environmental Protection Agency, Washington, DC, EPA/600/R-19/188, 2019 (*[*https://cfpubepagov/ncea/isa/recordisplaycfm?deid=347534*](https://cfpubepagov/ncea/isa/recordisplaycfm?deid=347534)*, last access: 20200128)* 2019.

61. Wang Y, Xiao S, Zhang Y, et al. Long-term exposure to PM2.5 major components and mortality in the southeastern United States. *Environ Int* 2022; **158**: 106969.

62. Fischer PH, Marra M, Ameling CB, et al. Particulate air pollution from different sources and mortality in 7.5 million adults - The Dutch Environmental Longitudinal Study (DUELS). *The Science of the total environment* 2020; **705**: 135778-.

63. Jordan TB, Seen AJ, Jacobsen GE. Levoglucosan as an atmospheric tracer for woodsmoke. *Atmospheric Environment* 2006; **40**(27): 5316-21.