Environmental Research

Long-term exposure to fine particle elemental components and lung cancer incidence in the ELAPSE pooled cohort

--Manuscript Draft--

Manuscript Number:	
Article Type:	Research paper
Section/Category:	Environmental Health & Risk Assessment
Keywords:	air pollution; fine particulate matter; elemental components; lung cancer incidence; pooled cohort
Corresponding Author:	Ulla Arthur Hvidtfeldt Kraeftens Bekaempelse DENMARK
First Author:	Ulla Arthur Hvidtfeldt
Order of Authors:	Ulla Arthur Hvidtfeldt
	Jie Chen
	Zorana Jovanovic Andersen
	Richard Atkinson
	Mariska Bauwelinck
	Tom Bellander
	Jørgen Brandt
	Bert Brunekreef
	Giulia Cesaroni
	Hans Concin
	Daniela Fecht
	Francesco Forastiere
	Carla H. van Gils
	John Gulliver
	Ole Hertel
	Gerard Hoek
	Barbara Hoffmann
	Kees de Hoogh
	Nicole Janssen
	Karl-Heinz Jöckel
	Jeanette Therming Jørgensen
	Klea Katsouyanni
	Matthias Ketzel
	Jochem O. Klompmaker
	Alois Lang
	Karin Leander
	Shuo Liu
	Petter L.S. Ljungman

	Patrik K.E. Magnusson
	Amar Jayant Mehta
	Gabriele Nagel
	Bente Oftedal
	Göran Pershagen
	Raphael Simon Peter
	Annette Peters
	Matteo Renzi
	Debora Rizzuto
	Sophia Rodopoulou
	Evangelia Samoli
	Per Everhard Schwarze
	Gianluca Severi
	Torben Sigsgaard
	Massimo Stafoggia
	Maciej Strak
	Danielle Vienneau
	Gudrun Weinmayr
	Kathrin Wolf
	Ole Raaschou-Nielsen
Abstract:	Background An association between long-term exposure to fine particulate matter (PM 2.5) and lung cancer has been established in previous studies. PM 2.5 is a complex mixture of chemical components from various sources and little is known about whether certain components contribute specifically to the associated lung cancer risk. The present study builds on recent findings from the “Effects of Low-level Air Pollution: A Study in Europe” (ELAPSE) collaboration and addresses the potential association between specific elemental components of PM 2.5 and lung cancer incidence. Methods We pooled seven cohorts from across Europe and assigned exposure estimates for eight components of PM 2.5 representing non-tail pipe emissions (copper (Cu), iron (Fe), and zinc (Zn)), long-range transport (sulfur (S)), oil burning/industry emissions (nickel (Ni), vanadium (V)), crustal material (silicon (Si)), and biomass burning (potassium (K)) to cohort participants’ baseline residential address based on 100 m by 100 m grids from newly developed hybrid models combining air pollution monitoring, land use data, satellite observations, and dispersion model estimates. We applied stratified Cox proportional hazards models, adjusting for potential confounders (age, sex, calendar year, marital status, smoking, body mass index, employment status, and neighborhood-level socio-economic status). Results The pooled study population comprised 306,550 individuals with 3,916 incident lung cancer events during 5,541,672 person-years of follow-up. We observed a positive association between exposure to all eight components and lung cancer incidence, with adjusted HRs of 1.10 (95% CI 1.05, 1.16) per 50 ng/m 3 PM 2.5 K, 1.09 (95% CI 1.02, 1.15) per 1 ng/m 3 PM 2.5 Ni, 1.22 (95% CI 1.11, 1.35) per 200 ng/m 3 PM 2.5 S, and 1.07 (95% CI 1.02, 1.12) per 200 ng/m 3 PM 2.5 V. Effect estimates were largely unaffected by adjustment for nitrogen dioxide (NO 2). After adjustment

	for PM 2.5 mass, effect estimates of K, Ni, S, and V were slightly attenuated, whereas effect estimates of Cu, Si, Fe, and Zn became null or negative. Conclusions Our results point towards an increased risk of lung cancer in connection with sources of combustion particles from oil and biomass burning and secondary inorganic aerosols rather than non-exhaust traffic emissions. Specific limit values or guidelines targeting these specific PM 2.5 components may prove helpful in future lung cancer prevention strategies.
Suggested Reviewers:	Bertil Forsberg Professor, Umeå University bertil.forsberg@umu.se Expert in the field of Air pollution and health
	Dan Crouse Consulting Senior Scientist, Health Effects Institute dcrouse@healtheffects.org Experienced within the field of health effects of fine particulate matter
	Bart Ostro University of California bostro@pacbell.net Expert in the field of ambient PM2.5 and health
	C Arden Pope III Brigham Young University-Provo: Brigham Young University cap3@byu.edu Expert in the field of PM2.5 and health
	Francine Laden Professor, Harvard University T H Chan School of Public Health francine.laden@channing.harvard.edu Expert in the field of environmental epidemiology of cancer

Danish Cancer Society

October, 2020

Danish Cancer Society
Research Center
Work, Environment and Cancer
(WEC)

Strandboulevarden 49
DK-2100 Copenhagen

Tlf +45 35 25 75 00
www.cancer.dk

UNDER PROTECTION OF
HER MAJESTY THE QUEEN

Dear editors,

Enclosed, please find the manuscript entitled "Long-term exposure to fine particle elemental components and lung cancer incidence in the ELAPSE pooled cohort" which we propose for publication in Environmental Research.

This study is part of the Effects of Low-level Air Pollution: A Study in Europe (ELAPSE) collaboration based on a large pooled cohort of more than 300,000 participants across Europe and close to 4,000 incident cases of lung cancer. The individual cohorts for pooling were selected among the most well characterized ESCAPE cohorts. It builds on a new study (currently in second revision with Environment International), in which we investigated the association between ambient exposure to NO_2, $PM_{2.5}$, BC, and O_3 and lung cancer incidence. The findings of this study indicated an effect of $PM_{2.5}$ even at levels lower than the EU limit value of 25 $\mu g/m^3$ and possibly even below the WHO Air Quality Guideline value of 10 $\mu g/m^3$.

This specific study on fine particulate elemental components and lung cancer incidence pools seven cohorts from across Europe with a total of 306,104 participants and approx. 4,000 incident lung cancers (close to 2,000 more than in ESCAPE). We developed new hybrid models combining air pollution monitoring, land use data, satellite observations, and dispersion model estimates for this project.

All of the authors have read and approved the paper and it has not been published previously nor is it being considered by any other peer-reviewed journal. All authors certify to have participated sufficiently in the work to take public responsibility for the appropriateness of the design and method, and the collection, analysis, and interpretation of the data.

Yours sincerely,

Ulla Arthur Hvidtfeldt,
MSc, PhD

Highlights

- Exposure to $PM_{2.5}$ is associated with a higher risk of lung cancer
- $PM_{2.5}$ is a complex mixture of components from various sources
- We observed positive associations between all components and lung cancer
- Combustion particles and secondary inorganic aerosols may be of special importance

Long-term exposure to fine particle elemental components and lung cancer incidence in the ELAPSE pooled cohort

Ulla Arthur Hvidtfeldt[a]
[a]Danish Cancer Society Research Center, Strandboulevarden 49, 2100 Copenhagen, Denmark
ullah@cancer.dk

Jie Chen[b]
[b]Institute of Risk Assessment Sciences, University of Utrecht, P.O. Box 80177, NL 3508 TD Utrecht, The Netherlands
j.chen1@uu.nl

Zorana Jovanovic Andersen[c]
[c]Section of Environmental Health, Department of Public Health, University of Copenhagen, Øster Farimagsgade 5, 1014 Copenhagen, Denmark
zorana.andersen@sund.ku.dk

Richard Atkinson[d]
[d]Population Health Research Institute and MRC-PHE Centre for Environment and Health, St George's, University of London, London, UK
atkinson@sgul.ac.uk

Mariska Bauwelinck[e]
[e]Interface Demography – Department of Sociology, Vrije Universiteit Brussel, Pleinlaan 2, 1050, Brussels, Belgium
mariska.bauwelinck@vub.ac.be

Tom Bellander[f,g]
[f]Institute of Environmental Medicine, Karolinska Institutet, Box 210, SE-171 77 Stockholm, Sweden
[g]Center for Occupational and Environmental Medicine, Region Stockholm, Solnavägen 4, plan 10 SE-113 65, Stockholm, Sweden
tom.bellander@ki.se

Jørgen Brandt[h,i]
[h]Department of Environmental Science, Aarhus University, Frederiksborgvej 399, P.O.Box 358, 4000 Roskilde, Denmark
[i]iClimate – Aarhus University interdisciplinary Centre for Climate Change, Frederiksborgvej 399, P.O.Box 358, 4000 Roskilde, Denmark
jbr@envs.au.dk

Bert Brunekreef[b]
[b]Institute of Risk Assessment Sciences (IRAS), University of Utrecht, P.O. Box 80177, NL 3508 TD Utrecht, The Netherlands
B.Brunekreef@uu.nl

Giulia Cesaroni[j]
[j]Department of Epidemiology, Lazio Region Health Service / ASL Roma 1, via Cristoforo Colombo 112, 00147 Rome, Italy
g.cesaroni@deplazio.it

Hans Concin[k]
[k]Agency for Preventive and Social Medicine, Rheinstraße 61, 6900 Bregenz, Austria
Hans.Concin@aks.or.at

Daniela Fecht[l]
[l]UK Small Area Health Statistics Unit, MRC Centre for Environment and Health, School of Public Health, Imperial College London, W2 1PG London, UK
d.fecht@imperial.ac.uk

Francesco Forastiere[m,n]
[m]Institute for Biomedical Research and Innovation (IRIB), National Research Council, 90146 Palermo, Italy.
[n]Environmental Research Group, Imperial College, London W12 0BZ, UK.
fran.forastiere@gmail.com

Carla H. van Gils[o]
[o]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, P.O. Box 85500, 3508 GA Utrecht, the Netherlands
C.vanGils@umcutrecht.nl

John Gulliver[p]
[p]Centre for Environmental Health and Sustainability & School of Geography, Geology and the Environment, University of Leicester, Leicester, LE1 7RH, UK
jg435@leicester.ac.uk

Ole Hertel[q]
[q]Department of Environmental Science, Aarhus University, Frederiksborgvej 399, P.O.Box 358, 4000 Roskilde, Denmark
oh@envs.au.dk

Gerard Hoek[b]
[b]Institute of Risk Assessment Sciences (IRAS), University of Utrecht, P.O. Box 80177, NL 3508 TD Utrecht, The Netherlands
G.Hoek@uu.nl

Barbara Hoffmann[r]
[r]Institute of Occupational, Social and Environmental Medicine, Medical Faculty, Heinrich Heine University, Gurlittstraße 55, 40223 Dusseldorf, Germany
b.hoffmann@uni-duesseldorf.de

Kees de Hoogh[s,t]
[s]Swiss Tropical and Public Health Institute, Socinstrasse 57, 4051 Basel, Switzerland
[t]University of Basel, Petersplatz 1, Postfach, 4001 Basel, Switzerland
c.dehoogh@swisstph.ch

Nicole Janssen[u]
[u]National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
nicole.janssen@rivm.nl

Karl-Heinz Jöckel[v]
[v]Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, University Duisburg-Essen, 45147 Essen, Germany
k-h.Joeckel@uk-essen.de

Jeanette Therming Jørgensen[c]
[c]Section of Environmental Health, Department of Public Health, University of Copenhagen, Øster Farimagsgade 5, 1014 Copenhagen, Denmark
jethe@sund.ku.dk

Klea Katsouyanni[x,y]
[x]Dept. of Hygiene, Epidemiology and Medical Statistics, National and Kapodistrian University of Athens, Medical School 75, Mikras Asias street 115 27 Athens, Greece
[y]NIHR HPRU Health Impact of Environmental Hazards, School of Public Health, Imperial College, London
kkatsouy@med.uoa.gr

Matthias Ketzel[h,z]
[h]Department of Environmental Science, Aarhus University, Frederiksborgvej 399, P.O.Box 358, 4000 Roskilde, Denmark
[z]Global Centre for Clean Air Research (GCARE), University of Surrey, Guildford, United Kingdom
mke@envs.au.dk

Jochem O. Klompmaker[b,u]
[k]Institute of Risk Assessment Sciences (IRAS), University of Utrecht, P.O. Box 80177, NL 3508 TD Utrecht, The Netherlands
[u]National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
jochem.klompmaker@rivm.nl

Alois Lang[aa]
[aa]Cancer registry Vorarlberg, Agency for Preventive and Social Medicine, Rheinstraße 61, Bregenz 6900, Austria
alois.lang@vol.at

Karin Leander[f]
[f]Institute of Environmental Medicine, Karolinska Institutet, Box 210, SE-171 77 Stockholm, Sweden
karin.leander@ki.se

Shuo Liu[c]
[c]Section of Environmental Health, Department of Public Health, University of Copenhagen, Øster Farimagsgade 5, 1014 Copenhagen, Denmark
shli@sund.ku.dk

Petter L.S. Ljungman[f,bb]
[f]Institute of Environmental Medicine, Karolinska Institutet, Box 210, SE-171 77 Stockholm, Sweden
[bb]Department of Cardiology, Danderyd University Hospital, Stockholm, Sweden
petter.ljungman@ki.se

Patrik K.E. Magnusson[cc]
[cc]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, SE-171 77 Stockholm, Sweden
patrik.magnusson@ki.se

Amar Jayant Mehta[dd,ee]
[dd]Statistics Denmark, Sejrøgade 11, 2100 Copenhagen, Denmark.
[ee]Section of Epidemiology, Department of Public Health, Faculty of Health and Medical Sciences, University of Copenhagen, Øster Farimagsgade 5, 1014 Copenhagen, Denmark.
amar.mehta@sund.ku.dk

Gabriele Nagel[ff,k]
[ff]Institute of Epidemiology and Medical Biometry, Ulm University, Helmholtzstr. 22, 89081 Ulm, Germany
[k]Agency for Preventive and Social Medicine, Rheinstraße 61, 6900 Bregenz, Austria
gabriele.nagel@uni-ulm.de

Bente Oftedal[gg]
[gg]Section of Air Pollution and Noise, Norwegian Institute of Public Health, P.O. Box 222, Skøyen, N-0213 Oslo, Norway
bente.oftedal@fhi.no

Göran Pershagen[f,g]
[f]Institute of Environmental Medicine, Karolinska Institutet, Box 210, SE-171 77 Stockholm, Sweden
[g]Center for Occupational and Environmental Medicine, Region Stockholm, Solnavägen 4, plan 10 SE-113 65, Stockholm, Sweden
goran.pershagen@ki.se

Raphael Simon Peter[ff]
[ff]Institute of Epidemiology and Medical Biometry, Ulm University, Helmholtzstr. 22, 89081 Ulm, Germany
Raphael.Peter@uni-ulm.de

Annette Peters[gg,hh]
[gg]Institute of Epidemiology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany
[hh]Chair of Epidemiology, Ludwig Maximilians Universität München, Munich, Germany
peters@helmholtz-muenchen.de

Matteo Renzi[j]
[j]Department of Epidemiology, Lazio Region Health Service / ASL Roma 1, via Cristoforo Colombo 112, 00147 Rome, Italy
m.renzi@deplazio.it

Debora Rizzuto[ii,jj]
[ii]Aging Research Center, Department of Neurobiology, Care Sciences, and Society, Karolinska Institutet and Stockholm University, Stockholm 17165, Sweden
[jj]Stockholm Gerontology Research Center, Stockholm 11346, Sweden
debora.rizzuto@ki.se

Sophia Rodopoulou[x]
[x]Dept. of Hygiene, Epidemiology and Medical Statistics, National and Kapodistrian University of Athens, Medical School 75, Mikras Asias street 115 27 Athens, Greece
srodopoyl@med.uoa.gr

Evangelia Samoli[x]
[x]Dept. of Hygiene, Epidemiology and Medical Statistics, National and Kapodistrian University of Athens, Medical School 75, Mikras Asias street 115 27 Athens, Greece
esamoli@med.uoa.gr

Per Everhard Schwarze[kk]
[kk]Infection Control and Environmental Health, Norwegian Institute of Public Health, Oslo, Norway
per.schwarze@fhi.no

Gianluca Severi[ll, mm]
[ll]CESP, UMR 1018, Université Paris-Saclay, Inserm, Gustave Roussy, Villejuif, France
[mm]Department of Statistics, Computer Science and Applications “G. Parenti“ (DISIA), University of Florence, Italy
gianluca.severi@inserm.fr

Torben Sigsgaard[nn]
[nn]Department of Public Health, Environment Occupation and Health, Danish Ramazzini Centre, Aarhus University, Bartholins Allé 2, 8000 Aarhus, Denmark
ts@ph.au.dk

Massimo Stafoggia[j,f]
[j]Department of Epidemiology, Lazio Region Health Service / ASL Roma 1, via Cristoforo Colombo 112, 00147 Rome, Italy
[f]Institute of Environmental Medicine, Karolinska Institutet, Box 210, SE-171 77 Stockholm, Sweden
m.stafoggia@deplazio.it

Maciej Strak[b,u]
[b]Institute of Risk Assessment Sciences (IRAS), University of Utrecht, P.O. Box 80177, NL 3508 TD Utrecht, The Netherlands
[u]National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands
maciek.strak@rivm.nl

Danielle Vienneau[s,t]
[s]Swiss Tropical and Public Health Institute, Socinstrasse 57, 4051 Basel, Switzerland
[t]University of Basel, Petersplatz 1, Postfach, 4001 Basel, Switzerland
danielle.vienneau@unibas.ch

Gudrun Weinmayr[ff]
[ff]Institute of Epidemiology and Medical Biometry, Ulm University, Helmholtzstr. 22, 89081 Ulm, Germany
Gudrun.Weinmayr@uni-ulm.de

Kathrin Wolf[gg]
[gg]Institute of Epidemiology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany
kathrin.wolf@helmholtz-muenchen.de

Ole Raaschou-Nielsen[a,h]
[a]Danish Cancer Society Research Centre, Strandboulevarden 49, 2100 Copenhagen, Denmark
[h]Department of Environmental Science, Aarhus University, Frederiksborgvej 399, P.O.Box 358, 4000 Roskilde, Denmark
ole@cancer.dk

Declarations of interest: none

Corresponding author:
Ulla Arthur Hvidtfeldt,
Danish Cancer Society Research Centre,
Strandboulevarden 49, 2100 Copenhagen Ø, Denmark
email: ullah@cancer.dk

Abstract

Background

An association between long-term exposure to fine particulate matter ($PM_{2.5}$) and lung cancer has been established in previous studies. $PM_{2.5}$ is a complex mixture of chemical components from various sources and little is known about whether certain components contribute specifically to the associated lung cancer risk. The present study builds on recent findings from the "Effects of Low-level Air Pollution: A Study in Europe" (ELAPSE) collaboration and addresses the potential association between specific elemental components of $PM_{2.5}$ and lung cancer incidence.

Methods

We pooled seven cohorts from across Europe and assigned exposure estimates for eight components of $PM_{2.5}$ representing non-tail pipe emissions (copper (Cu), iron (Fe), and zinc (Zn)), long-range transport (sulfur (S)), oil burning/industry emissions (nickel (Ni), vanadium (V)), crustal material (silicon (Si)), and biomass burning (potassium (K)) to cohort participants' baseline residential address based on 100 m by 100 m grids from newly developed hybrid models combining air pollution monitoring, land use data, satellite observations, and dispersion model estimates. We applied stratified Cox proportional hazards models, adjusting for potential confounders (age, sex, calendar year, marital status, smoking, body mass index, employment status, and neighborhood-level socio-economic status).

Results

The pooled study population comprised 306,550 individuals with 3,916 incident lung cancer events during 5,541,672 person-years of follow-up. We observed a positive association between exposure to all eight components and lung cancer incidence, with adjusted HRs of 1.10 (95% CI 1.05, 1.16) per 50 ng/m^3 $PM_{2.5}$ K, 1.09 (95% CI 1.02, 1.15) per 1 ng/m^3 $PM_{2.5}$ Ni, 1.22 (95% CI 1.11, 1.35) per 200 ng/m^3 $PM_{2.5}$ S, and 1.07 (95% CI 1.02, 1.12) per 200 ng/m^3 $PM_{2.5}$ V. Effect estimates were largely unaffected by adjustment for nitrogen dioxide (NO_2). After adjustment for $PM_{2.5}$ mass, effect estimates of K, Ni, S, and V were slightly attenuated, whereas effect estimates of Cu, Si, Fe, and Zn became null or negative.

Conclusions

Our results point towards an increased risk of lung cancer in connection with sources of combustion particles from oil and biomass burning and secondary inorganic aerosols rather than non-exhaust traffic emissions. Specific limit values or guidelines targeting these specific $PM_{2.5}$ components may prove helpful in future lung cancer prevention strategies.

Keywords: air pollution, fine particulate matter, elemental components, lung cancer incidence, pooled cohort.

Funding

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.

The Swedish Twin Registry is managed by Karolinska Institutet and receives funding through the Swedish Research Council under the grant no 2017-00641.

1. Background

Previous studies have reported a positive relationship between long-term exposure to fine particulate matter ($PM_{2.5}$) and lung cancer incidence.[1,2] Recently, we published* a study within the Effects of Low-level Air Pollution: A Study in Europe (ELAPSE) collaboration based on a large pooled cohort of more than 300,000 participants across Europe and close to 4,000 incident cases of lung cancer. The findings of this study indicated a positive association with $PM_{2.5}$ even at levels lower than the EU limit value of 25 μg/m^3 and possibly even below the WHO Air Quality Guideline value of 10 μg/m^3.[3]

$PM_{2.5}$ originates from various different sources and contains a complex mixture of chemical components. Primary particles are emitted from a variety of sources such as wood stove burning, brake and tyre wear, gasoline and diesel engines, and industry while secondary particles are formed from gaseous pollutants converted in chemical reactions in the atmosphere. Limited knowledge exists about which $PM_{2.5}$ components contribute to the associated lung cancer risk. Only four previous studies have addressed the influence of specific PM components on the incidence or mortality of lung cancer. Three American cohort studies have reported positive associations between long-term exposure to sulfate (SO_4) or sulfur and lung cancer mortality.[4–6] The exposure data in these studies were derived from centrally located air-monitoring stations. The large European Study of Cohorts for Air Pollution Effects (ESCAPE) applied a land use regression model for exposure assessment and investigated eight PM elements with sulfur (S) representing long-range transport of particles from combustion of sulfur-containing fossil fuels (e.g. in power plants), copper (Cu), iron (Fe), and zinc (Zn) representing non-tail pipe emissions, nickel (Ni) and vanadium (V) representing mixed oil-burning and industry, silicon (Si) representing crustal material, and potassium (K) representing biomass burning. The study included 14 cohorts from across eight European countries with a total of 1,878 incident lung cancer cases and reported positive associations for all PM components except V, however, with a large degree of statistical uncertainty.[7] Two-pollutant analyses including also total $PM_{2.5}$ mass in the model, suggested S to be the most important component in terms of lung cancer risk. Differences in effect estimates from epidemiological studies of $PM_{2.5}$ could partly be explained by variations in the chemical composition of $PM_{2.5}$.[8]

At present, no EU limit values or guidelines exist for PM components. Prevention strategies could be improved based on knowledge about the relative harm of the specific components. The objective of this study was to explore further our recently published* findings on $PM_{2.5}$ and lung cancer incidence within the ELAPSE collaboration[3] by addressing the potential association between specific elemental components of fine PM and lung cancer incidence. The present study builds on the data from the ESCAPE project and adds a longer follow-up with more than 2,000 additional incident lung cancer cases and a newly developed Europe-wide spatial land-use regression model for assessing long-term exposure to elemental particle composition.

*The paper is currently in revision

2. Methods

The methods including study population, outcome definition, and statistical analyses followed our earlier ELAPSE study.[3]

2.1 Study population

The ELAPSE collaboration includes nine cohorts with the following inclusion criteria: low-level air pollution data availability, relatively recent recruitment date, and ability to share data for pooling. Of these nine cohorts, seven included information on lung cancer incidence and the most important potential confounders. The cohorts originated in Sweden (*Cardiovascular Effects of Air Pollution and Noise in Stockholm* [CEANS], which is the collective name of the following four sub-cohorts: Swedish National Study on Aging and Care in Kungsholmen [SNAC-K],[9] Stockholm Screening Across the Lifespan Twin study [SALT],[10] Stockholm 60 years old study [Sixty],[11] and Stockholm Diabetes Prevention Program [SDPP]),[12] Denmark (*Diet, Cancer and Health cohort* [DCH][13] and *Danish Nurse Cohort* [DNC][14]), the Netherlands (*Dutch European Investigation into Cancer and Nutrition* [EPIC-NL] consisting of EPIC-Monitoring Project on Risk Factors and Chronic Diseases in the Netherlands [EPIC-MORGEN] and [EPIC-Prospect]),[15] France (*Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale* [E3N or EPIC-France]),[16] Germany (*Heinz Nixdorf Recall study* [HNR]),[17] and Austria (*Vorarlberg Health Monitoring and Prevention Programme* [VHM&PP]).[18] The French E3N and the Danish DNC covered large regions of the countries, whereas the study areas of the remaining cohorts represented a large city and its surrounding areas. We harmonized the variables across the individual cohorts according to a joint codebook. All seven cohorts had information available at baseline on age, sex, smoking status, amount and duration of smoking in current smokers, body mass index (BMI), employment status and area-level socio-economic status (SES). For the E3N and VHM&PP cohorts, which only had smoking intensity and duration in classes, we created uniform distributions within the bins. Each cohort is described in more detail in the first section of online supplemental material.

2.2 Exposure assessment

In line with the ESCAPE study, we selected eight components to represent major air pollution sources: Cu, Fe and Zn representing non-tailpipe traffic emissions such as brake and tyre wear, S representing secondary inorganic aerosols from long-range transported sulfur containing fuel combustion, Ni and V representing mixed oil burning/industry emissions, Si representing crustal material, and K representing biomass burning.[19]

We applied Europe-wide hybrid land use regression (LUR) models, which incorporated satellite observations, dispersion model estimates, land use, traffic variables, industrial point sources, and ESCAPE air pollution monitoring data for 2010 (436 sites). The exposure modelling and validation has been described in detail previously.[19] We developed PM composition models for the year 2010 using two algorithms: Supervised linear regression (SLR) and random forest. We assigned pollution surfaces (100 m x 100 m grids) from both algorithms to the baseline residential address of each of the cohort members. The models explained a moderate to large fraction of the measured concentration variation at the European scale, ranging from 41% to 91% across components. Random forest outperformed SLR in modelling between cohort variability, but the model performances were similar for the within-area

concentration variability in five-fold cross-validation.[19] In the current analysis, we primarily exploit within-cohort contrasts (section 2.4). $PM_{2.5}$ mass and NO_2 estimates were derived from hybrid LUR models which applied 2010 AirBase routine monitoring data maintained by the European Environmental Agency (EEA), predictors of satellite observations, dispersion model estimates, land use, and traffic variables, as described previously.[20]

We truncated negative predictions to zero and a few unrealistically high predictions at close distance to industrial sources.[19] We performed truncation (mainly for predictions below zero) in the main model population for SLR-modeled exposure: 11.3% for Cu, 0.5% for Fe, 11.6% for Ni, 14.3% for V and 2.6% for Zn. No truncation was needed for exposure modeled by random forest.

2.3 Outcome

We identified cases of lung cancer in cancer registries, death certificates or medical records with the exception of the E3N cohort in which self-reports from biannual questionnaires or death certificates were applied. The self-reported cases were confirmed through pathological reports and reviewed by a lung oncologist. We excluded persons registered with a cancer before baseline (except non-melanoma skin cancer). We included primary cancers located in the bronchus and the lung (ICD9 codes 162.2-162.9 and ICD10 code C34).

2.4 Statistical analyses

We applied Cox proportional hazards models and calculated hazard ratios (HR) with 95% confidence intervals (CI) using age as the underlying time scale. Each $PM_{2.5}$ component was included as a linear function with increments of 5 ng/m^3 Cu, 100 ng/m^3 Fe, 50 ng/m^3 K, 1 ng/m^3 Ni, 200 ng/m^3 S, 100 ng/m^3 Si, 2 ng/m^3 V, 10 ng/m^3 Zn, following the increments selected in previous publications from ESCAPE and ELAPSE.[7] The online appendix Table A.1 provides the interquartile ranges (IQR) for each elemental component for the pooled cohort. Our presented HRs therefore reflect a larger exposure contrast than the IQR for most elements. We censored each cohort member at time of first occurrence of any cancer other than lung cancer, date of death, emigration, loss to follow-up or at the end of follow-up, with the exception of the HNR cohort, for which we only had follow-up for lung cancer specifically and not for other cancers. We included strata per individual (sub) cohort to account for baseline hazard heterogeneity across the cohorts and to relax the proportional hazards assumption. The strata option had a superior model performance compared to alternative approaches (e.g. indicator per sub-cohort or a frailty term).[3,21] As a consequence of applying the strata option to account for between cohort heterogeneity, we primarily evaluate within (sub) cohort exposure contrasts.

We modelled the association between each PM component and lung cancer incidence in three á priori specified models: *1)* accounting for age (applied as the underlying time-scale), (sub) cohort ID (included as strata), sex (included as strata), and adjustment for year of enrolment in order to account for time-trends in exposure and outcome; *2)* further adjusted for individual-level factors marital status (married/cohabiting, divorced, single, widowed), smoking status (never, former, current), smoking duration (years of smoking) for current smokers, smoking intensity (cigarettes/day) for current smokers, square of smoking intensity, BMI (<18.5, 18.5–24, 25–29, and 30+ kg/m^2), and employment status (yes vs. no); *3)* (main model) further adjusted for neighborhood-level socio-economic status (SES) defined as mean income in 2001, which was the most consistently available variable and year

across cohorts. The spatial scale of a 'neighborhood' varied from smaller neighborhoods and city districts (CEANS, EPIC-NL, E3N, HNR) to municipalities (DNS, DCH, and VHM&PP). We excluded participants with incomplete information on model 3 variables from all analyses. Our previous study included a comprehensive analysis testing the sensitivity of missing confounders, for example educational level, smoking intensity in former smokers, and occupational class.[3]

We performed analyses with exposures estimated by the SLR and the random forest algorithm. We have no prior as to which exposure model is the primary model, as both models explain within-cohort exposure contrast with similar performance. We present most analyses with the SLR exposure model results, as our previous ELAPSE paper on $PM_{2.5}$ and lung cancer only used SLR and the current paper is a further exploration of these PM findings. The SLR model is also more comparable to the LUR models used in the ESCAPE study.

Sensitivity analyses included: 1) Fitting natural spline functions with 3 degrees of freedom to assess the shape of the association between air pollution and lung cancer; and 2) Two-pollutant models with particle components and either $PM_{2.5}$ mass or NO_2 as the second pollutant, with NO_2 representing traffic exhaust emission which is of special relevance for the analyses of associations with the traffic non-exhaust components Cu, Fe and Zn. The $PM_{2.5}$ mass and NO_2 estimates were developed with the SLR algorithm.[20]

We evaluated violation of the proportional hazards assumption of the Cox Models for all covariates by test of a non-zero slope in a generalized linear regression of the scaled Schoenfeld residuals on time. We performed all analyses in R version 3.4.0 using packages: *survival, coxme, Matrix, foreach, glmnet, multcomp, survey, splines, Hmisc, mfp, VIM, ggplot2, frailtySurv, survsim, eha, stamod*.

3. Results

The pooled study population comprised 306,104 individuals and 3,916 incident lung cancer events during 5,541,672 person-years of follow-up (Table 1). The recruitment period of participants ranged over the period 1985–2005 and the mean age at baseline ranged from 41.7 to 72.5 years across the individual (sub) cohorts with a pooled mean of 48.3 years. Four sub cohorts included women only and the pooled cohort comprised 34 % men. Current smokers at baseline ranged from 13 to 37 % across the individual (sub) cohorts with a pooled percentage of 24. Overweight or obese participants varied from 21 % in the French E3N cohort to 73 % in the German HNR cohort.

The exposure distribution of each $PM_{2.5}$ component according to (sub) cohort is provided in Figure 1. Overall, the exposure concentrations were lower in the North European cohorts compared to more Southern cohorts, with the exception of $PM_{2.5}$ Ni, Si, and V. We observed a substantial exposure contrast within each (sub) cohort for $PM_{2.5}$ Cu, Fe, Ni, Si, and V. Overall, exposure derived by SLR and random forest were similar, but large differences were found in individual cohorts between the two exposure algorithms (Figure 1). The correlations of the $PM_{2.5}$ components with $PM_{2.5}$ mass and NO_2 varied considerably from low to moderate across the (sub) cohorts (online Tables A.2 and A.3). The correlations with NO_2 were generally higher than with $PM_{2.5}$ mass with coefficients > 0.8 for Cu, Fe, and Si in some (sub) cohorts. The correlations between the specific $PM_{2.5}$ components varied substantially across individual (sub) cohorts (Table A.4).

The linear associations between $PM_{2.5}$ components and lung cancer incidence in the three models of increasing confounder adjustments are presented in Table 2 for SLR modelled exposure. We observed a positive association between exposure to all components considered and lung cancer incidence. Overall, adjustment for the individual-level confounders attenuated the HRs substantially (model 2). Further adjustment for area-level income resulted in small increases or no change in the HR (model 3). In the fully adjusted model 3, we found positive associations for all components, which were statistically significant for K, Ni, S, and V. The splines were generally linear to supra-linear and did not indicate a level below which no association was present (Figure 2). The model 3 estimates for analyses, in which we applied the random forest exposure, is presented in Figure 3. For most components, the HR point estimates were similar for random forest compared to exposures modelled by SLR, but random forest estimates generally had wider confidence intervals, because of the smaller predicted exposure contrast.

The estimates were generally unaffected by adjustment for NO_2 in the two-pollutant models, except for an increase in the HR for $PM_{2.5}$ S to 1.31 (95% CI 1.17, 1.48) (Table 3). The NO_2 estimate was generally attenuated by adjustment for each component (Table A.5). After adjustment for $PM_{2.5}$ mass, associations with $PM_{2.5}$ K, Ni, S and V remained positive and (borderline) statistically significant. The HR for $PM_{2.5}$ Si was attenuated towards the null. Associations for $PM_{2.5}$ Cu, Fe and Zn were null or somewhat negative after adjustment for $PM_{2.5}$ mass. The HR for $PM_{2.5}$ mass remained stable after adjustment for the $PM_{2.5}$ Si component and slightly reduced by adjustment for $PM_{2.5}$ K and V (Table A.5). Adjustment for $PM_{2.5}$ S reduced the $PM_{2.5}$ point estimate to unity whereas the estimate was slightly increased following adjustment for $PM_{2.5}$ Cu, Fe, and Zn.

4. Discussion

The results of this study point towards an elevated risk of lung cancer following exposure to several $PM_{2.5}$ components. The positive relationships between all components and lung cancer remained in analyses taking into account NO_2, which is consistent with the weak associations observed between NO_2 and lung cancer.[3] Adjustment for $PM_{2.5}$ mass attenuated the point estimates for some components, but a positive relationship remained for $PM_{2.5}$ K, Ni, V, and S. The positive associations with S, K and V were also found with exposures modelled with random forest. In contrast, the associations with Ni were essentially null with exposures by random forest. Thus, the results indicate an influence primarily of sources of biomass burning (K) and industrial and fuel-oil combustion particles (Ni, V) and long-range transported secondary inorganic aerosols from sulfur-containing fossil fuel combustion (S) in relation to lung cancer incidence.

Few previous studies have addressed the association between individual $PM_{2.5}$ components and lung cancer incidence. Our results are generally in line with those previously published from the ESCAPE collaboration,[7] although the ESCAPE estimates were less precise with wide confidence bounds and no evidence of an association between $PM_{2.5}$ V and lung cancer was found. With regards to $PM_{2.5}$ S, our point estimate of 1.22 (95% CI 1.11, 1.33) is somewhat lower than the corresponding result from ESCAPE of 1.34 (95% CI 0.74, 2.42) per 200 ng/m^3. In the current study we included six cohorts from the ESCAPE study and now performed a pooled analysis instead of cohort-specific analyses followed by meta-analysis. By increasing the follow-up time, we were able to include more than twice as many lung cancer cases compared to the ESCAPE study. We furthermore had a more harmonized Europe-wide exposure model and were able to use a larger part of some cohorts, such as DCH and E3N, because of the new exposure model. Results from the American Cancer Prevention Study (II) and the Harvard Six Cities study also suggested an elevated risk of lung cancer mortality in relation to sulfate exposure.[4,5] In analyses within the National Particle Component Toxicity (NPACT) initiative, a clear association between $PM_{2.5}$ S and lung cancer mortality was observed, but not for Fe, K, Ni, Si, Zn, or V.[6] In analyses according to source categories, the coal combustion category contributed most strongly to lung cancer mortality.[6]

Our results were most consistent for $PM_{2.5}$ S, Ni, V, and K, which could either reflect an impact of the component itself or of its dominant sources. $PM_{2.5}$ S is mostly present as sulfate in the particle phase, mostly formed in the atmosphere by oxidation of sulfur dioxide, which is emitted by combustion of sulfur-containing fuels. Sulfates are concentrated in fine particles and therefore able to be transported over long distances, resulting in a high background concentration with limited small-scale spatial variation. S may also represent other fine particle components formed simultaneously, such as polycyclic aromatic hydrocarbons (PAH) which have been linked to inflammation and lung cancer previously.[22,23] Ni and V both represent combustion of fuel oil/industrial emissions with shipping and oil refineries being the major sources. While these $PM_{2.5}$ components are highly correlated in many of the study regions, they are not identical and seem to jointly characterize relevant sources of particulate matter associated with lung cancer. Our new exposure models reflect the major sources for Ni and V better than in the ESCAPE study, because the models profit from observations in multiple study areas whereas the study-specific ESCAPE models did not always contain the known major sources or were based on few sites influenced by these sources. $PM_{2.5}$ K is used as a tracer of wood/biomass burning, but is also affected by soil. In the K model, no information on local wood burning is included because of lack of predictor data. We did not find evidence for an association of $PM_{2.5}$ Cu, Fe and Zn with lung cancer. Cu, Fe and Zn are predominantly markers of non-exhaust emissions of motorized road traffic,

related to road, brake and tyre wear. In our earlier ELAPSE paper, we also did not find evidence that NO_2 was associated with lung cancer in this population.[3]

Overall, $PM_{2.5}$ has been linked to lung cancer through mechanisms of oxidative stress and pulmonary inflammation leading to DNA damage, promotion of cell turnover and proliferation in the lung tissue. Also, epigenetic changes of the genome, and in particular promotor hypermethylation, are suspected of mediating the effects of air pollutants on lung cancer.[24] In support of our findings regarding specific PM components, ambient vanadium has previously been linked with an increase in certain biological markers for oxidative DNA damage,[25,26] and to *in vivo* lung tumor promotion in mice.[27] Likewise, nickel has been associated with carcinogenicity through oxidative stress and epigenetic mechanisms.[28] $PM_{2.5}$ S has been related to lung cancer risk through mechanisms of DNA methylation changes.[29,30]

Strengths of our study included the large sample size obtained by pooling seven cohorts combined with detailed information on individual lifestyle, and thus, the ability to include a broad range of potential confounders harmonized across cohorts for this specific project. The cohorts covered a large part of Europe and represented a broad range of exposure and the large sample size enabled multi-pollutant models to disentangle potential inter-dependencies between pollutants. The exposure models developed within the ELAPSE collaboration ensured comparable exposure estimates for the entire study population. However, the application of a model for exposure assignment inevitably imposes some misclassification due to uncertainties in input data and because exposure modelled at the residential address does not necessarily represent the true personal exposure. Our exposure was modelled for the year 2010 and applied to the baseline year of each cohort. The majority of (sub) cohorts had their baselines during the 1990's. In our previous paper on NO_2, $PM_{2.5}$, black carbon, and ozone in relation to lung cancer incidence, we applied exposure back-extrapolated to the baseline of each (sub) cohort.[3] The Spearman correlation coefficient between the 2010 exposure concentration and the exposure back-extrapolated to baseline was 0.76 for $PM_{2.5}$ and we observed lower though still statistically significant effect estimates for back-extrapolated $PM_{2.5}$ exposures compared to the main approach of 2010-exposures. We were not able to back-extrapolate individual $PM_{2.5}$ component exposure to the baseline because of insufficient information on concentrations of $PM_{2.5}$ components in Europe over time. Previous studies from Europe have found the spatial distribution of NO_2 and traffic intensities to be stable over several years,[31–33] which suggests that the spatial contrast for traffic-related components such as Cu and Fe may be relatively constant over time. However, we are not able to draw conclusions about the temporal and spatial pattern of the remaining components. An issue in two-pollutant models is that associations are more readily identified with the more precisely modelled component compared to a less precisely modelled component.[34] In addition, we lacked information on personal activity patterns (work place address, time spent indoors/outside etc.) as well as moving patterns from baseline until end of follow-up. Again, the results on $PM_{2.5}$ mass and lung cancer incidence from the previous study did not indicate major differences between estimates based on exposure applied to the address history compared to the main 2010 exposure.[3]

In conclusion, the results of this study point towards an increased risk of lung cancer in connection with sources of combustion particles rather than non-exhaust traffic emissions. The observed association with $PM_{2.5}$ S indicates that combustion of sulfur-containing fossil fuels may contribute to the lung cancer incidence also far away from the source via long-range transported secondary inorganic aerosols. Specific limit values or guidelines targeting these $PM_{2.5}$ components may prove helpful in future lung cancer prevention strategies.

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.

Acknowledgments

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

References

1 Raaschou-Nielsen O, Andersen ZJ, Beelen R, *et al.* Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). *Lancet Oncol* 2013; **14**: 813–22.

2 Hamra GB, Guha N, Cohen A, *et al.* Outdoor Particulate Matter Exposure and Lung Cancer: A Systematic Review and Meta-Analysis. *Environ Health Perspect* 2014; **122**: 906–11.

3 Hvidtfeldt UA. Long-term low-level ambient air pollution exposure and risk of lung cancer – a pooled analysis of 7 European cohorts. *Environ Int* 2020; : (in revision).

4 Pope CA, Burnett RT, Thun MJ, *et al.* Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* 2002; **287**: 1132–41.

5 Krewski D, Burnett RT, Goldberg MS, *et al.* Reanalysis of the Harvard Six Cities study and the American Cancer Society study of particulate air pollution and mortality: a special report of the Institute's Particle Epidemiology Reanalysis Project. *Cambridge MA Heal Eff Inst* 2000; : 97.

6 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. In: National Particle Component Toxicity (NPACT) Initiative: Integrated Epidemiologic and Toxicologic Studies of the Health Effects of Particulate Matter Components. Research report 177. Boston, MA: Health Effects Institute, 2013: 332.

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

8 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.

9 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 Clin Exp Res* 2004; **16**: 158–68.

10 Magnusson PKE, Almqvist C, Rahman I, *et al.* The Swedish Twin Registry: establishment of a biobank and other recent developments. *Twin Res Hum Genet* 2013; **16**: 317–29.

11 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 Metab* 2007; **33**: 30–6.

12 Eriksson A-K, Ekbom A, Granath F, Hilding A, Efendic S, Ostenson C-G. Psychological distress and risk of pre-diabetes and Type 2 diabetes in a prospective study of Swedish middle-aged men and women. *Diabet Med* 2008; **25**: 834–42.

13 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. *Scand J Public Health* 2007; **35**: 432–41.

14 Hundrup YA, Simonsen MK, Jørgensen T, Obel EB. Cohort profile: The danish nurse cohort. *Int J Epidemiol* 2012; **41**: 1241–7.

15 Beulens JWJ, Monninkhof EM, Monique Verschuren WM, *et al.* Cohort profile: The EPIC-NL study. *Int J Epidemiol* 2010; **39**: 1170–8.

16 Clavel-Chapelon F, E3N Study Group for the ES. Cohort Profile: The French E3N Cohort Study. *Int J Epidemiol* 2015; **44**: 801–9.

17 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. *Am Heart J* 2002; **144**: 212–8.

18 Ulmer H, Kelleher CC, 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. *J Intern Med* 2007; **261**: 566–76.

19 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; : (in revision).

20 de Hoogh K, Chen J, Gulliver J, *et al.* Spatial PM2.5, NO2, O3 and BC models for Western Europe – Evaluation of spatiotemporal stability. *Environ Int* 2018; **120**: 81–92.

21 Samoli E, Rodopoulou S, Hvidtfeldt UA, *et al.* Modeling multi-level survival data in multi-center epidemiological cohort studies: applications from the ELAPSE project. *Submitted* 2020.

22 O'Driscoll CA, Gallo ME, Hoffmann EJ, *et al.* Polycyclic aromatic hydrocarbons (PAHs) present in ambient urban dust drive proinflammatory T cell and dendritic cell responses via the aryl hydrocarbon receptor (AHR) in vitro. *PLoS One* 2018; **13**. DOI:10.1371/journal.pone.0209690.

23 Yuan JM, Butler LM, Gao YT, *et al.* Urinary metabolites of a polycyclic aromatic hydrocarbon and volatile organic compounds in relation to lung cancer development in lifelong never smokers in the Shanghai Cohort Study. *Carcinogenesis* 2014; **35**: 339–45.

24 Straif K, Cohen A, Samet J. IARC SCIENTIFIC PUBLICATION NO. 161: AIR POLLUTION AND CANCER. Lyon: International Agency for Research on Cancer, 2013.

25 Sørensen M, Schins RPF, Hertel O, Loft S. Transition metals in personal samples of PM2.5 and oxidative stress in human volunteers. *Cancer Epidemiol Biomarkers Prev* 2005; **14**: 1340–3.

26 Ehrlich VA, Nersesyan AK, Atefie K, *et al.* Inhalative exposure to vanadium pentoxide causes DNA damage in workers: Results of a multiple end point study. *Environ Health Perspect* 2008; **116**: 1689–93.

27 Rondini EA, Walters DM, Bauer AK. Vanadium pentoxide induces pulmonary inflammation and tumor promotion in a strain-dependent manner. *Part Fibre Toxicol* 2010; **7**: 9.

28 Genchi G, Carocci A, Lauria G, Sinicropi MS, Catalano A. Nickel: Human health and environmental toxicology. Int. J. Environ. Res. Public Health. 2020; **17**. DOI:10.3390/ijerph17030679.

29 Hou L, Zhang X, Zheng Y, *et al.* Altered methylation in tandem repeat element and elemental component levels in inhalable air particles. *Environ Mol Mutagen* 2014; **55**: 256–65.

30 Yun Y, Gao R, Yue H, Guo L, Li G, Sang N. Sulfate Aerosols Promote Lung Cancer Metastasis by Epigenetically Regulating the Epithelial-to-Mesenchymal Transition (EMT). *Environ Sci Technol* 2017; **51**: 11401–11.

31 Beelen R, Hoek G, Fischer P, Brandt PA va. den, Brunekreef B. Estimated long-term outdoor air pollution concentrations in a cohort study. *Atmos Environ* 2007; **41**: 1343–58.

32 Cesaroni G, Porta D, Badaloni C, *et al.* Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. *Environ Heal* 2012; **11**: 48.

33 Gulliver J, Morris C, Lee K, Vienneau D, Briggs D, Hansell A. Land use regression modeling to estimate historic (1962-1991) concentrations of black smoke and sulfur dioxide for Great

Britain. *Environ Sci Technol* 2011; **45**: 3526–32.

34 Zeger SL, Thomas D, Dominici F, *et al.* Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect* 2000; **108**: 419–26.

Table 1. Description of the included (sub) cohort studies

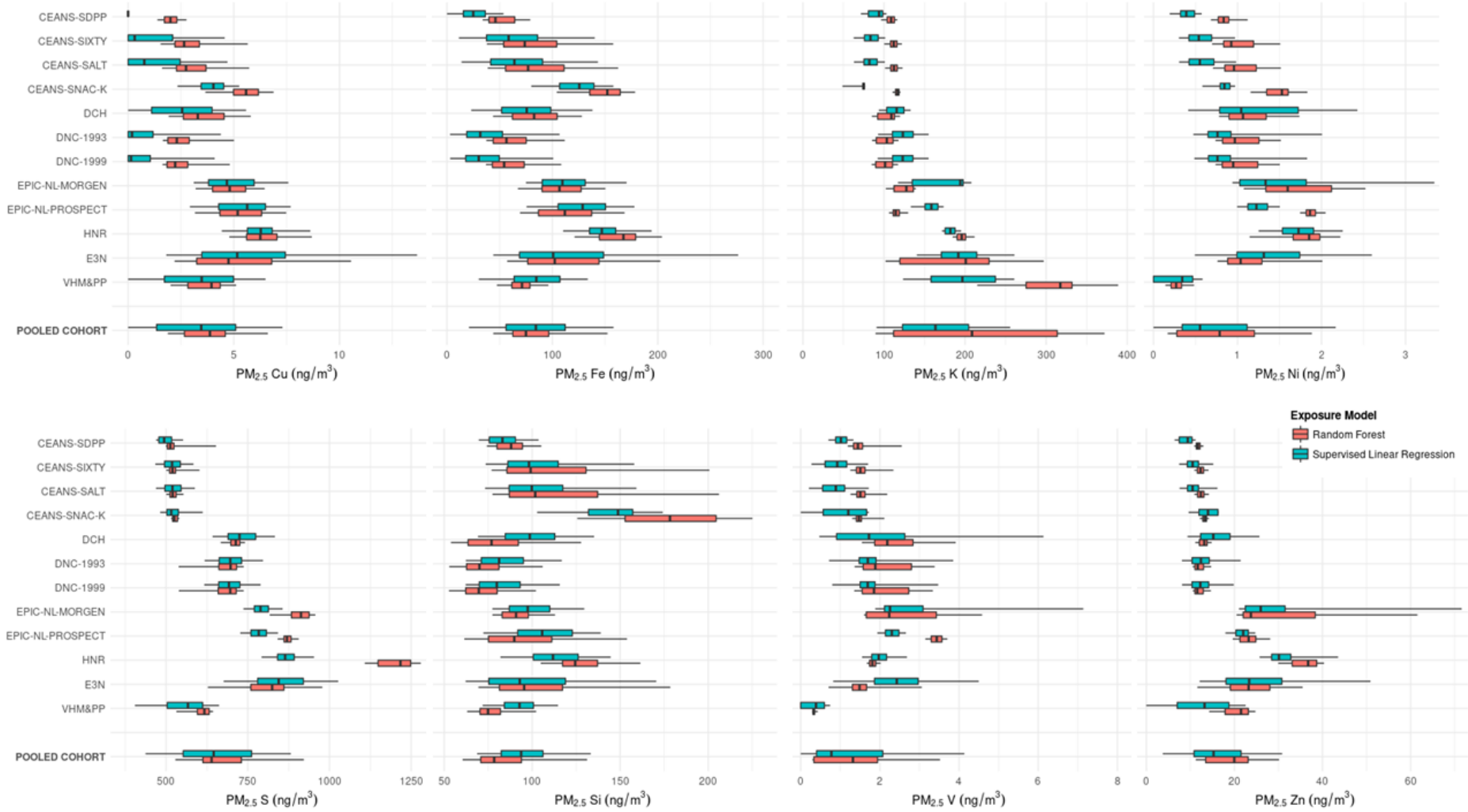
	Total participants	Baseline period	End of follow-up	Baseline age Mean (SD)	Lung cancers	Males %	Current smokers %	Cigarettes /day[a] Mean (SD)	Years of smoking[a] Mean (SD)	BMI ≥ 25 kg/m^2 %	Not em-ployed %	Married/ cohabiting %	Neighborhood income[b] Mean (SD)
CEANS Stockholm, Sweden													
SDPP	7,305	1992–1998	31-12-2011	47.0 (4.9)	42	41	26	13.5 (7.4)	27.8 (8.6)	51	9	84	24.3 (4.2)
SIXTY	3,660	1997–1999	31-12-2011	60 (0)	38	50	21	13.3 (7.7)	27.8 (8.6)	65	32	74	24.7 (6.8)
SALT	5,625	1998–2003	31-12-2011	57.3 (10.4)	43	47	21	12.7 (8.0)	37.6 (9.1)	41	33	68	25.4 (6.6)
SNAC-K	2,359	2001–2004	31-12-2011	72.5 (10.4)	21	38	15	11.7 (8.3)	43.2 (13.5)	53	76	46	28.7 (2.2)
DCH, Copenhagen /Aarhus, Denmark	52,779	1993–1997	31-12-2015	56.7 (4.4)	1,474	47	36	16.5 (9.0)	36.3 (7.7)	56	22	71	20.1 (3.4)
DNC, Denmark													
DNC-1993	15,556	1993	31-12-2012	56.0 (8.3)	299	0	37	13.8 (8.1)	31.4 (9.9)	28	29	68	19.2 (2.6)
DNC-1999	7,430	1999	31-12-2012	47.9 (4.1)	25	0	33	13.2 (7.4)	27.1 (7.1)	30	5	76	19.0 (2.4)
EPIC-NL, Netherlands													
MORGEN	17,792	1993–1997	31-12-2012	42.7 (11.2)	170	46	35	15.7 (8.6)	24.5 (10.6)	49	31	65	12.2 (1.6)
Prospect	13,640	1993–1997	31-12-2012	57.6 (6.0)	191	0	23	13.6 (8.7)	36.7 (7.6)	55	49	77	13.1 (1.4)
HNR, Ruhr area, Germany	3,611	2000–2003	26-04-2017	59.1 (7.7)	69	50	25	19.1 (12.5)	33.9 (9.2)	73	57	75	25.1 (8.1)
E3N, France	36,258	1989–1991	08-12-2014	52.8 (6.7)	157	0	13	11.3 (9.1)	28.5 (7.6)	21	31	84	11.2 (3.0)
VHM&PP, Vorarlberg, Austria	140,089	1985–2005	31-12-2014	41.7 (14.9)	1,387	44	20	15.6 (8.9)	13.4 (8.2)	42	29	69	22.9 (1.7)
Pooled cohort	**306,104**	**1985-2005**	**2011–2017**	**48.3 (13.4)**	**3,916**	**34**	**24**	**15.2 (8.9)**	**25.3 (13.1)**	**43**	**29**	**72**	**19,8 (5,3)**

CEANS: Cardiovascular Effects of Air Pollution and Noise in Stockholm; SDPP: The Stockholm Diabetes Preventive Program; SIXTY: The Stockholm cohort of 60-year-olds; SALT: Screening Across the Lifespan Twin Study; SNAC-K: The Swedish National Study of Aging and Care in Kungsholmen; DCH: Diet, Cancer and Health; DNC: Danish Nurses Cohort; EPIC-NL: European Prospective Investigation into Cancer and Nutrition, the Netherlands; MORGEN: Monitoring Project on Risk Factors and chronic diseases in the Netherlands; HNR: Heinz Nixdorf Recall study; E3N (EPIC-France): Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale; VHM&PP: Vorarlberg Health Monitoring and Prevention Programme.

[a]Among current smokers at baseline.

[b]Euros x 1,000, year 2001.

Figure 1. Distribution of $PM_{2.5}$ components for the year 2010 at baseline addresses estimated from SLR and random forest models.



The box boundaries indicate P25 and P75; the bold line in the middle of the box – P50; whiskers indicate P5 and P95.

Table 2. Pooled analyses of $PM_{2.5}$ components (SLR) exposure and risk of lung cancer

		Model 1[a] N=306,104			**Model 2[b]** N=306,104			**Model 3[c]** N=306,104		
$PM_{2.5}$ component	*Increment*	**HR**	**95% CI**		**HR**	**95% CI**		**HR**	**95% CI**	
$PM_{2.5}$ mass	5 μg/m^3	1.19	1.11	1.28	1.12	1.04	1.22	1.14	1.05	1.23
$PM_{2.5}$ Cu	5 ng/m^3	1.21	1.12	1.32	1.02	0.94	1.11	1.04	0.96	1.13
$PM_{2.5}$ Fe	100 ng/m^3	1.26	1.16	1.38	1.02	0.93	1.12	1.04	0.95	1.14
$PM_{2.5}$ K	50 ng/m^3	1.12	1.06	1.18	1.09	1.04	1.15	1.10	1.05	1.16
$PM_{2.5}$ Ni	1 ng/m^3	1.23	1.17	1.30	1.10	1.04	1.16	1.09	1.02	1.15
$PM_{2.5}$ S	200 ng/m^3	1.45	1.31	1.59	1.23	1.12	1.35	1.22	1.11	1.35
$PM_{2.5}$ Si	100 ng/m^3	1.84	1.56	2.18	1.17	0.98	1.39	1.14	0.96	1.35
$PM_{2.5}$ V	2 ng/m^3	1.18	1.14	1.24	1.08	1.03	1.13	1.07	1.02	1.12
$PM_{2.5}$ Zn	10 ng/m^3	1.09	1.05	1.13	1.02	0.98	1.07	1.03	0.98	1.07

HR, hazard ratio; CI, confidence interval

[a]Adjusted for study (strata), age (time-scale), sex (strata), year of baseline visit

[b]Further adjusted for smoking status, duration, intensity, intensity^2, BMI, marital status, and employment status

[c]Further adjusted for 2001 mean income at the neighborhood level

Table 3. Pooled two-pollutant analyses of $PM_{2.5}$ components (SLR) and co-pollutants and risk of lung cancer

		Single pollutant[a] N=306,104			**Adjusted for $PM_{2.5}$ mass** N=306,104			**Adjusted for NO_2** N=306,104		
$PM_{2.5}$ component	*Increment*	**HR**	**95% CI**		**HR**	**95% CI**		**HR**	**95% CI**	
$PM_{2.5}$ Cu	5 ng/m^3	1.04	0.96	1.13	0.92	0.82	1.03	1.04	0.90	1.20
$PM_{2.5}$ Fe	100 ng/m^3	1.04	0.95	1.14	0.92	0.83	1.03	1.02	0.86	1.21
$PM_{2.5}$ K	50 ng/m^3	1.10	1.05	1.16	1.08	1.01	1.15	1.10	1.04	1.16
$PM_{2.5}$ Ni	1 ng/m^3	1.09	1.02	1.15	1.05	0.99	1.12	1.10	1.03	1.18
$PM_{2.5}$ S	200 ng/m^3	1.22	1.11	1.35	1.21	1.06	1.39	1.31	1.17	1.48
$PM_{2.5}$ Si	100 ng/m^3	1.14	0.96	1.35	1.01	0.83	1.22	1.17	0.92	1.49
$PM_{2.5}$ V	2 ng/m^3	1.07	1.02	1.12	1.05	0.99	1.10	1.07	1.02	1.13
$PM_{2.5}$ Zn	10 ng/m^3	1.03	0.98	1.07	0.98	0.93	1.27	1.02	0.98	1.07

HR, hazard ratio; CI, confidence interval

[a]Adjusted for study (strata), age (time-scale), sex (strata), year of baseline visit, smoking status, duration, intensity, intensity^2, BMI, marital status, employment status, and 2001 mean income at the neighborhood level

Figure 2. Natural spline functions (3 df) of $PM_{2.5}$ components (SLR) and lung cancer incidence

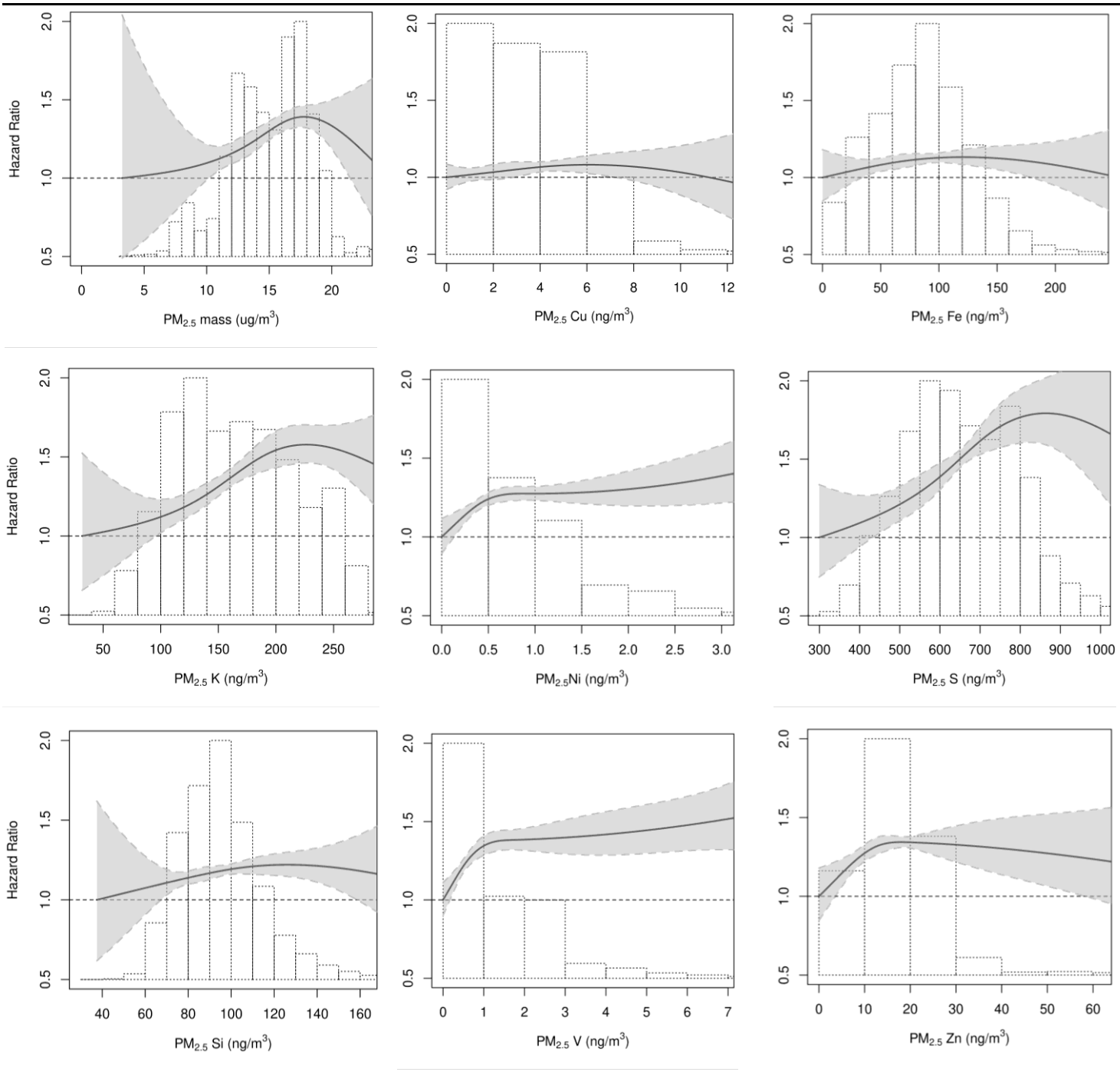
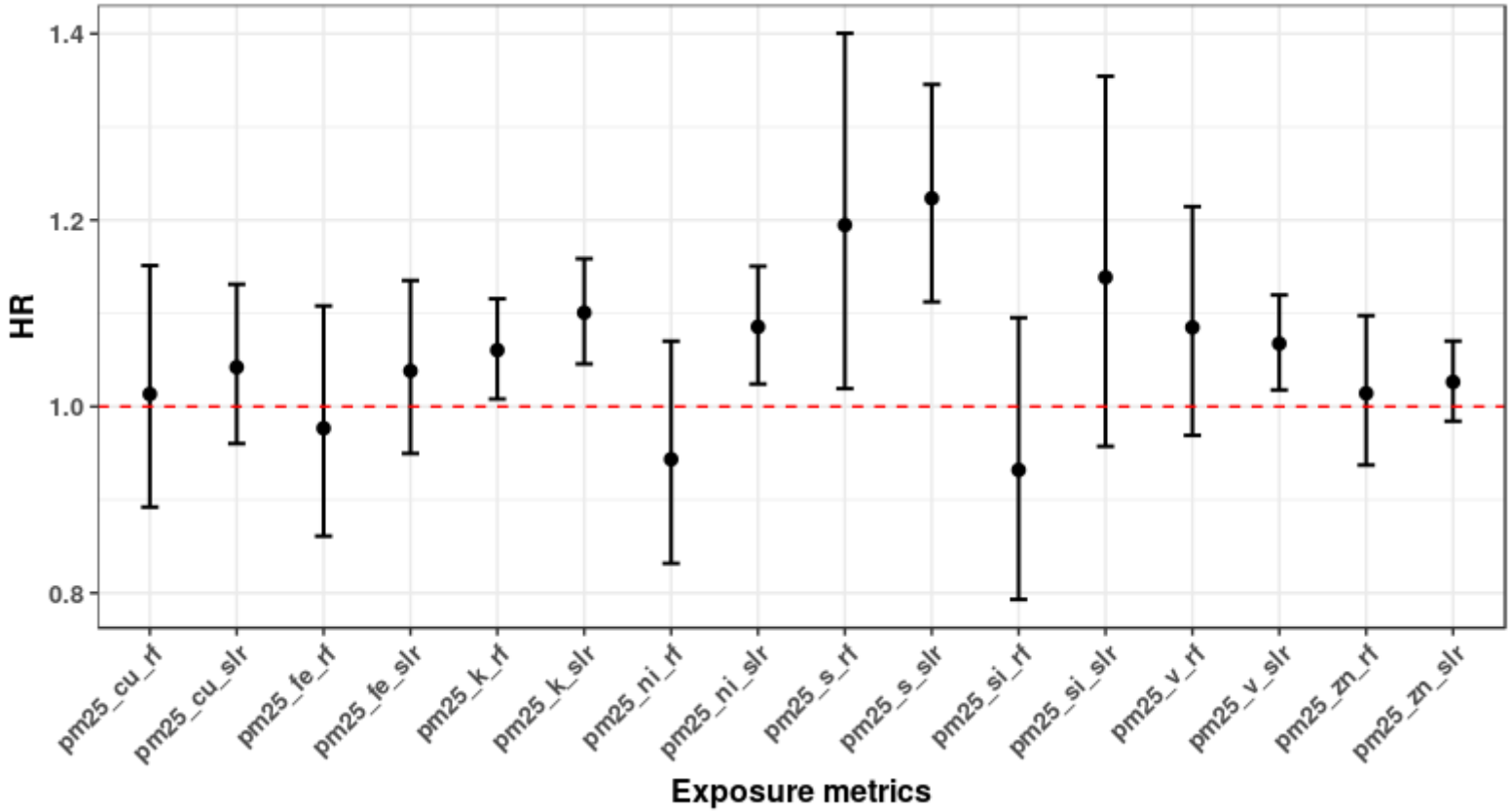


Figure 3. Associations between $PM_{2.5}$ components and lung cancer based on SLR and random forest exposure algorithms (N=306,104)



SLR, Supervised Linear regression; RF, Random Forest

Model 3 estimates adjusted for study (strata), age (time-scale), sex (strata), year of baseline visit, smoking status, duration, intensity, intensity^2, BMI, marital status, employment status, and 2001 mean income at the neighborhood level

CRediT authorship contribution statement

Ulla Arthur Hvidtfeldt: *Formal analysis*, *Methodology, Software, Visualization, Writing – original draft*; **Jie Chen:** *Data curation, Methodology, Project administration, Software, Writing - review & editing*; **Zorana Jovanovic Andersen:** *Writing - review & editing*; **Richard Atkinson:** *Writing - review & editing*; **Mariska Bauwelinck:** *Data curation, Writing - review & editing*; **Tom Bellander:** *Writing - review & editing*; **Jørgen Brandt:** *Data curation, Methodology, Writing - review & editing*; **Bert Brunekreef:** *Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing - review & editing*; **Giulia Cesaroni:** *Methodology, Writing - review & editing*; **Hans Concin:** *Project administration, Resources*; **Daniela Fecht:** *Conceptualization, Data curation, Methodology, Supervision, Writing - review & editing*; **Francesco Forastiere:** *Conceptualization, Funding acquisition, Writing - review & editing*; **Carla H. van Gils:** *Writing - review & editing*; **John Gulliver:** *Conceptualization, Funding acquisition, Methodology, Validation, Visualization, Writing - review & editing*; **Ole Hertel:** *Data curation, Methodology, Writing - review & editing*; **Gerard Hoek:** *Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing - review & editing*; **Barbara Hoffmann:** *Conceptualization, Data curation, Funding acquisition, Resources, Writing - review & editing*; **Kees de Hoogh:** *Data curation, Methodology, Writing - review & editing*; **Nicole Janssen:** *Writing - review & editing*; **Karl-Heinz Jöckel:** *Data curation, Investigation, Resources, Writing - review & editing*; **Jeanette Therming Jørgensen:** *Data curation, Writing - review & editing*; **Klea Katsouyanni:** *Methodology, Writing - review & editing*; **Matthias Ketzel:** *Data curation, Methodology, Writing - review & editing*; **Jochem O. Klompmaker:** *Writing - review & editing*; **Alois Lang:** *Data curation, Writing - review & editing*; **Karin Leander:** *Resources, Writing - review & editing*; **Shuo Liu:** *Writing - review & editing*; **Petter L.S. Ljungman:** *Conceptualization, Data curation, Writing - review & editing*; **Patrik K.E. Magnusson:** *Investigation, Project administration, Writing - review & editing*; **Amar Jayant Mehta;** *Writing - review & editing*; **Gabriele Nagel:** *Data curation, Writing - review & editing*; **Bente Oftedal:** *Writing - review & editing*; **Göran Pershagen:** *Data curation, Writing - review & editing*; **Raphael Simon Peter:** *Data curation, Writing - review & editing*; **Annette Peters:** *Data curation, Funding acquisition, Project administration, Resources, Writing - review & editing*; **Matteo Renzi:** *Writing - review & editing*; **Debora Rizzuto:** *Data curation, Writing - review & editing*; **Sophia Rodopoulou:** *Data curation, Methodology, Software, Writing - review & editing*; **Evangelia Samoli:** *Methodology, Software, Writing - review & editing*; **Per Everhard Schwarze:** *Writing - review & editing*; **Gianluca Severi:** *Data curation, Resources, Writing - review & editing*; **Torben Sigsgaard:** *Methodology, Writing - review & editing*; **Massimo Stafoggia:** *Methodology, Writing - review & editing*; **Maciek Strak:** *Writing - review & editing*; **Danielle Vienneau:** *Writing - review & editing*; **Gudrun Weinmayr:** *Methodology, Writing - review & editing*; **Kathrin Wolf:**

Data curation, Software, Writing - review & editing; **Ole Raaschou-Nielsen:** *Conceptualization, Data curation, Methodology, Supervision, Writing - review & editing*.

conflict of Interest Statement

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.

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.

The Swedish Twin Registry is managed by Karolinska Institutet and receives funding through the Swedish Research Council under the grant no 2017-00641.

Declaration of interests

☒ 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.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: