



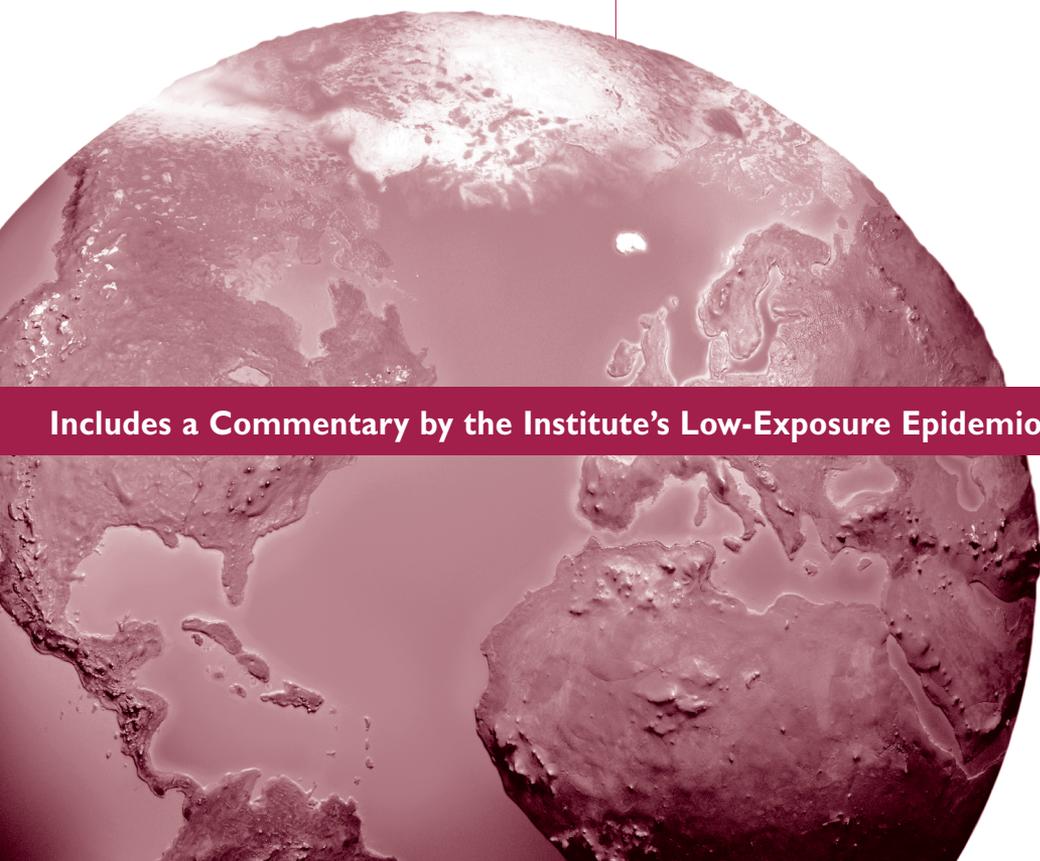
## RESEARCH REPORT

HEALTH  
EFFECTS  
INSTITUTE

Number 208  
September 2021

# Mortality and Morbidity Effects of Long-Term Exposure to Low-Level PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub>: An Analysis of European Cohorts in the ELAPSE Project

Bert Brunekreef, Maciej Strak, Jie Chen, Zorana J. Andersen, Richard Atkinson, Mariska Bauwelinck, Tom Bellander, Marie-Christine Boutron, Jørgen Brandt, Iain Carey, Giulia Cesaroni, Francesco Forastiere, Daniela Fecht, John Gulliver, Ole Hertel, Barbara Hoffmann, Kees de Hoogh, Danny Houthuijs, Ulla Hvidtfeldt, Nicole Janssen, Jeanette Jørgensen, Klea Katsouyanni, Matthias Ketzel, Jochem Klompmaker, Norun Hjertager Krog, Shuo Liu, Petter Ljungman, Amar Mehta, Gabriele Nagel, Bente Oftedal, Göran Pershagen, Annette Peters, Ole Raaschou-Nielsen, Matteo Renzi, Sophia Rodopoulou, Evi Samoli, Per Schwarze, Torben Sigsgaard, Massimo Stafoggia, Danielle Vienneau, Gudrun Weinmayr, Kathrin Wolf, and Gerard Hoek



Includes a Commentary by the Institute's Low-Exposure Epidemiology Studies Review Panel

ISSN 1041-5505 (print)  
ISSN 2688-6855 (online)

**Mortality and Morbidity Effects of Long-Term Exposure to  
Low-Level PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub>:  
An Analysis of European Cohorts in the ELAPSE Project**

Bert Brunekreef, Maciej Strak, Jie Chen, Zorana J. Andersen, Richard Atkinson,  
Mariska Bauwelinck, Tom Bellander, Marie-Christine Boutron, Jørgen Brandt,  
Iain Carey, Giulia Cesaroni, Francesco Forastiere, Daniela Fecht, John Gulliver,  
Ole Hertel, Barbara Hoffmann, Kees de Hoogh, Danny Houthuijs, Ulla Hvidtfeldt,  
Nicole Janssen, Jeanette Jørgensen, Klea Katsouyanni, Matthias Ketzel,  
Jochem Klompmaker, Norun Hjertager Krog, Shuo Liu, Petter Ljungman, Amar Mehta,  
Gabriele Nagel, Bente Oftedal, Göran Pershagen, Annette Peters, Ole Raaschou-Nielsen,  
Matteo Renzi, Sophia Rodopoulou, Evi Samoli, Per Schwarze, Torben Sigsgaard,  
Massimo Stafoggia, Danielle Vienneau, Gudrun Weinmayr, Kathrin Wolf, and Gerard Hoek

---

with a Commentary by the  
HEI Low-Exposure Epidemiology Studies Review Panel

Research Report 208  
Health Effects Institute  
Boston, Massachusetts

*Trusted Science • Cleaner Air • Better Health*

Publishing history: This document was posted at [www.healtheffects.org](http://www.healtheffects.org) in September 2021.

Citation for document:

Brunekreef B, Strak M, Chen J, Andersen ZJ, Atkinson R, Bauwelinck M, et al. 2021. Mortality and Morbidity Effects of Long-Term Exposure to Low-Level PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub>: An Analysis of European Cohorts in the ELAPSE Project. Research Report 208. Boston, MA:Health Effects Institute.

© 2021 Health Effects Institute, Boston, Mass., U.S.A. David Wade, Virginia Beach, Va., Compositor. Printed by Recycled Paper Printing, Boston, Mass. Library of Congress Catalog Number for the HEI Report Series: WA 754 R432.

---

 Cover paper: made with at least 55% recycled content, of which at least 30% is post-consumer waste; free of acid and elemental chlorine. Text paper: made with 100% post-consumer waste recycled content; acid free; no chlorine used in processing. The book is of permanent archival quality.

# CONTENTS

About HEI	vii
About This Report	ix
Contributors	xi
Preface	xiii
<b>HEI STATEMENT</b>	<b>1</b>
<b>INVESTIGATORS' REPORT</b> <i>by Brunekreef et al.</i>	<b>5</b>
<b>ABSTRACT</b>	<b>5</b>
<b>INTRODUCTION</b>	<b>8</b>
<b>OBJECTIVES</b>	<b>10</b>
<b>METHODS AND STUDY DESIGN</b>	<b>10</b>
<b>Study Design</b>	<b>10</b>
<b>Study Populations</b>	<b>11</b>
Pooled Cohort	11
Administrative Cohorts	13
<b>Exposure Assessment</b>	<b>13</b>
ELAPSE Europewide Hybrid Land Use Regression Models	14
Exposure Assignment	16
Back-Extrapolation	16
MAPLE PM <sub>2.5</sub> Exposure Model	17
Health Outcome Definition	17
<b>STATISTICAL METHODS AND DATA ANALYSIS</b>	<b>18</b>
<b>Model Specification for Pooled Cohort Analysis</b>	<b>20</b>
<b>Model Specification for Analysis of Administrative     Cohorts</b>	<b>22</b>
<b>Meta-analysis of Effect Estimates from Administrative     Cohorts</b>	<b>24</b>
<b>Combination of Pooled Cohort and Administrative     Cohorts' Effect Estimates</b>	<b>24</b>
<b>RESULTS</b>	<b>24</b>
<b>Pooled Cohort</b>	<b>24</b>
Exposure	24
Population Characteristics	26
Natural- and Cause-Specific Mortality Analyses	26
Additional Analyses	32
Summary of the Mortality Findings in the Pooled Cohort	37
Incidence of Acute Coronary Heart Disease, Stroke, Asthma, COPD, and Lung Cancer	38
Main Analysis of Association with Incidence of Disease	38
Subset Analysis	46
Additional Analyses	46
Summary of the Associations with Incidence of Disease in the Pooled Cohort	50

## Research Report 208

<b>Administrative Cohorts</b>	50
Exposure	50
Descriptive Data on Population and Mortality Outcomes	52
Main Analyses of Associations with Natural-Cause and Cause-Specific Mortality	52
Additional Analyses of the Association with Mortality	66
Summary of Mortality Findings in Administrative Cohorts	78
<b>Overview of Evidence from the Pooled and Administrative Cohorts</b>	80
<b>DISCUSSION AND CONCLUSIONS</b>	82
<b>Main Findings</b>	82
<b>Associations at Low Pollution Levels: Concentration–Response Functions</b>	82
<b>Mortality Associations in the Pooled Versus Administrative Cohorts</b>	84
<b>Comparison of the Shape of the Concentration–Response Function with Previous Studies</b>	85
<b>Associations with PM<sub>2.5</sub> from ELAPSE and MAPLE Models</b>	86
<b>Associations with Ozone</b>	86
<b>Two-Pollutant Models</b>	87
<b>Further Methodological Issues</b>	87
Exposure Assessment	87
Impact of Measurement Error in Exposure	89
Health Data	89
Confounding and Effect Modification	90
Noise Adjustment	91
Multiple Testing	91
<b>Comparison with the Canadian and U.S. HEI-Funded Studies, ESCAPE, and Recent Systematic Reviews</b>	91
<b>Strengths and Limitations</b>	92
<b>IMPLICATIONS OF FINDINGS</b>	94
<b>ACKNOWLEDGMENTS</b>	94
<b>REFERENCES</b>	94
<b>HEI QUALITY ASSURANCE STATEMENT</b>	99
<b>MATERIALS AVAILABLE ON THE HEI WEBSITE</b>	100
<b>ABOUT THE AUTHORS</b>	100
<b>OTHER PUBLICATIONS RESULTING FROM THIS RESEARCH</b>	103

## Research Report 208

<b>COMMENTARY</b> <i>by HEI's Low-Exposure Epidemiology Studies Review Panel</i>	105
<b>INTRODUCTION</b>	105
<b>SCIENTIFIC AND REGULATORY BACKGROUND</b>	105
Setting Air Quality Standards in the United States	106
Setting Air Quality Standards in Europe	106
The Advent of HEI Studies Observing Associations Below Current Air Quality Standards and Guidelines	107
<b>STUDY SUMMARY</b>	107
Study Objectives	107
Methods and Study Design	108
Exposure Modeling	108
Study Populations	110
Exposure Assignment	112
Main Epidemiological Analyses	112
Additional Epidemiological Analyses	113
<b>SUMMARY OF KEY RESULTS</b>	113
Modeling and Exposure Estimation Results	113
Pooled Cohort Health Analysis	114
Administrative Cohort Health Analysis	116
<b>EVALUATION BY THE HEI LOW-EXPOSURE EPIDEMIOLOGY STUDIES REVIEW PANEL</b>	116
Evaluation of Study Design and Approach	117
Air Pollution Models and Exposure Estimation	117
Evaluation of Epidemiological Analysis	118
Discussion of the Findings and Interpretation	120
Conclusions	121
<b>ACKNOWLEDGMENTS</b>	122
<b>REFERENCES</b>	122
<b>Related HEI Publications</b>	125
<b>Abbreviations and Other Terms</b>	127
<b>HEI Board, Committees, and Staff</b>	129



# ABOUT HEI

The Health Effects Institute is a nonprofit corporation chartered in 1980 as an independent research organization to provide high-quality, impartial, and relevant science on the effects of air pollution on health. To accomplish its mission, the institute

- Identifies the highest-priority areas for health effects research;
- Competitively funds and oversees research projects;
- Provides intensive independent review of HEI-supported studies and related research;
- Integrates HEI's research results with those of other institutions into broader evaluations; and
- Communicates the results of HEI's research and analyses to public and private decision makers.

HEI typically receives balanced funding from the U.S. Environmental Protection Agency and the worldwide motor vehicle industry. Frequently, other public and private organizations in the United States and around the world also support major projects or research programs. HEI has funded more than 340 research projects in North America, Europe, Asia, and Latin America, the results of which have informed decisions regarding carbon monoxide, air toxics, nitrogen oxides, diesel exhaust, ozone, particulate matter, and other pollutants. These results have appeared in more than 260 comprehensive reports published by HEI, as well as in more than 1,000 articles in the peer-reviewed literature.

HEI's independent Board of Directors consists of leaders in science and policy who are committed to fostering the public-private partnership that is central to the organization. The Research Committee solicits input from HEI sponsors and other stakeholders and works with scientific staff to develop a Five-Year Strategic Plan, select research projects for funding, and oversee their conduct. For this study, a special panel — HEI's Low-Exposure Epidemiology Studies Oversight Panel — has worked with the Research Committee in project selection and oversight. The Review Committee, which has no role in selecting or overseeing studies, works with staff to evaluate and interpret the results of funded studies and related research. For this study, a special review panel — HEI's Low-Exposure Epidemiology Studies Review Panel — is fulfilling this role.

All project results and accompanying comments by the Review Committee (or, in this case, the Low-Exposure Epidemiology Studies Review Panel) are widely disseminated through HEI's website ([www.healtheffects.org](http://www.healtheffects.org)), printed reports, newsletters and other publications, annual conferences, and presentations to legislative bodies and public agencies.



# ABOUT THIS REPORT

Research Report 208, *Mortality and Morbidity Effects of Long-Term Exposure to Low-Level PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub>: An Analysis of European Cohorts in the ELAPSE Project*, presents a research project funded by the Health Effects Institute and conducted by Dr. Bert Brunekreef of the Institute for Risk Assessment Sciences, Utrecht University, the Netherlands, and his colleagues. The report contains three main sections.

**The HEI Statement**, prepared by staff at HEI, is a brief, nontechnical summary of the study and its findings; it also briefly describes the Low-Exposure Epidemiology Studies Review Panel's comments on the study.

**The Investigators' Report**, prepared by Brunekreef and colleagues, describes the scientific background, aims, methods, results, and conclusions of the study.

**The Commentary**, prepared by members of the Low-Exposure Epidemiology Studies Review Panel with the assistance of HEI staff, places the study in a broader scientific context, points out its strengths and limitations, and discusses remaining uncertainties and implications of the study's findings for public health and future research.

This report has gone through HEI's rigorous review process. When an HEI-funded study is completed, the investigators submit a draft final report presenting the background and results of the study. This draft report was first examined by outside technical reviewers and a biostatistician. The report and the reviewers' comments were then evaluated by members of the Low-Exposure Epidemiology Studies Review Panel, an independent panel of distinguished scientists who have no involvement in selecting or overseeing HEI studies. During the review process, the investigators have an opportunity to exchange comments with the Review Panel and, as necessary, to revise their report. The Commentary reflects the information provided in the final version of the report.



# CONTRIBUTORS

## LOW-EXPOSURE EPIDEMIOLOGY STUDIES OVERSIGHT PANEL

**Jonathan M. Samet (Chair)** *Dean of the Colorado School of Public Health, University of Colorado–Denver*

**Amy H. Herring** *Sara & Charles Ayres Professor of Statistical Science and Global Health, Duke University, and member of the HEI Research Committee*

**Jay H. Lubin** *Senior Research Scientist (retired), National Cancer Institute, Division of Cancer Epidemiology & Genetics, Biostatistics Branch*

**Fred W. Lurmann** *Chairman / Manager of Exposure Assessment Studies, Sonoma Technology, Inc.*

## LOW-EXPOSURE EPIDEMIOLOGY STUDIES REVIEW PANEL

**Sverre Vedal (Chair)** *Professor, Environmental and Occupational Health Sciences, University of Washington*

**Sara D. Adar** *Associate Professor, Epidemiology, School of Public Health, University of Michigan*

**Benjamin Barratt** *Senior Lecturer in Chinese Environment, King's College, London, United Kingdom*

**Kiros T. Berhane** *Professor of Biostatistics and Director of Graduate Programs in Biostatistics and Epidemiology, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, and member of the HEI Review Committee*

**Christopher J. Paciorek** *Adjunct Professor, Statistical Computing Consultant, Department of Statistics, University of California–Berkeley*

**Jennifer L. Peel** *Professor of Epidemiology, Colorado School of Public Health and Department of Environmental and Radiological Health Sciences, Colorado State University, and member of the HEI Review Committee*

**Gavin Shaddick** *Chair of Data Science and Statistics, Department of Mathematics, University of Exeter, United Kingdom*

## HEI PROJECT STAFF

**Hanna Boogaard** *Consulting Principal Scientist (Study Oversight)*

**Dan Crouse** *Senior Scientist (Report Review and Commentary)*

**Eleanne van Vliet** *Staff Scientist (Report Review)*

**Joanna Keel** *Research Assistant*

**Mary Brennan** *Consulting Editor*

**Kristin Eckles** *Senior Editorial Manager*

**Hope Green** *Editorial Project Manager*



# PREFACE

## HEI's Program to Assess Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution

### INTRODUCTION

Levels of ambient air pollution have declined significantly over the last few decades in North America, Europe, and in other developed regions. Despite the decreasing levels of air pollution, several large epidemiologic studies published in the early 2010s reported associations between adverse health effects and exposure to air pollution. These studies found associations between exposure to fine particulate matter (PM<sub>2.5</sub>\*) and mortality at levels below the then-current ambient air quality standards (e.g., Beelen et al. 2014a, b; Crouse et al. 2012; Hales et al. 2012; Preface Figure 1). In order to inform future risk assessment and regulation, it is important to confirm whether associations with adverse health effects continue to be observed as levels of air pollution decline still further. It is also important to better understand the shape of the exposure–response function at those low levels. Both issues remain major uncertainties in setting air quality standards.

The growing scientific evidence for effects at pollution levels below current air quality standards, the large overall estimates of the burden of disease attributable to air pollution, as well as the interest in reducing greenhouse gases, suggest that more stringent air quality standards and guidelines may be considered in the future. For these reasons, there was a need for additional investigation to improve our understanding of exposure–response function(s) for mortality and morbidity at low levels of PM<sub>2.5</sub>, ozone (O<sub>3</sub>), and other ambient air pollutants. Such studies would inform risk

assessors and policy makers regarding exposure–response functions at levels of ambient air pollution currently prevalent in North America, Western Europe, and other high-income regions of the world.

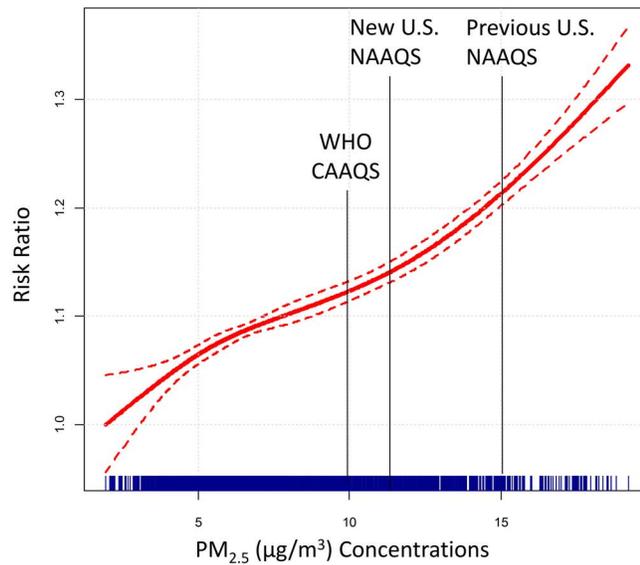
In 2014, HEI issued RFA 14-3, *Assessing Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution*, to solicit studies to address these important questions. The main goals of the RFA were to (1) fund studies to assess health effects of long-term exposure to low levels of ambient air pollution, including all-cause and cause-specific mortality and morbidity. Such studies should analyze and evaluate exposure–response function(s) for PM<sub>2.5</sub> and other pollutants at levels currently prevalent in North America, Western Europe, and other high-income regions and may also address related questions about health effects at low levels of ambient air pollution; and (2) develop statistical and other methodology required for, and specifically suited to, conducting such research including, but not limited to, evaluation and correction of exposure measurement error.

Investigators were asked to pay particular attention to having sufficiently large cohorts and statistical power to detect associations should they exist; having the ability to test various potential confounders of any associations; and to developing exposure assessment approaches and statistical methodology to enable a robust examination of the associations.

Specifically, investigators were asked to propose studies to:

1. Compare and contrast alternative analytic models and their uncertainty. For example, compare threshold against nonthreshold models, linear against nonlinear models, and parametric against nonparametric models, to characterize

\* A list of abbreviations and other terms appears at the end of this volume.



**Preface Figure 1. Shape of concentration–response function for mortality associated with fine particulate matter in a Canadian cohort.** (Adapted from Crouse et al. 2012, courtesy R. Burnett.)

the exposure–response function(s) at low levels of ambient air pollution.

2. Explore possible variability in estimates of risk at low levels among populations and identify possible contributing factors. Such factors could include age, smoking, socio-economic position, health status, and access to medical care, as well as differences in air pollution sources and time–activity patterns.
3. Develop and evaluate exposure assessment methods suitable to estimate exposure to low levels of air pollution at various spatial and temporal scales in large study populations, including people who reside in areas not covered by routine ground-level monitoring.
4. Develop, evaluate, and apply statistical methods to quantify and correct for exposure measurement error in risk estimates and in characterization of exposure–response relationships.
5. Develop and validate approaches to assess the effects of co-occurring pollutants on any health effect associations at low ambient concentrations.
6. Develop and validate indirect approaches to correct risk estimates for the effects of important

potential confounding variables, such as smoking, in the absence of such data at the individual level.

7. Improve techniques for record linkage and methods for disclosure protection for optimal use of large administrative databases in air pollution and health research.

---

## STUDY SELECTION

---

HEI established an independent Low Exposure Epidemiology Oversight Panel — consisting of outside experts and HEI Research Committee members — to prepare RFA 14-3 and review all applications submitted in response (see Contributors' page). Members of HEI's Research Committees with any conflict of interest were recused from all discussions and from the decision-making process. The HEI Research Committee reviewed the Panel's recommendations and recommended three studies for funding to HEI's Board of Directors, which approved funding in 2015.

This Preface summarizes the three studies, HEI's oversight process, and the review process for the Phase I reports.

**OVERVIEW OF THE LOW EXPOSURE  
EPIDEMIOLOGY STUDIES**

After a rigorous selection process, HEI funded three teams — led by Michael Brauer at the University of British Columbia, Canada; Francesca Dominici at the Harvard T.H. Chan School of Public Health, United States; and Bert Brunekreef at the University of Utrecht, the Netherlands — to investigate the health effects of exposure to low levels of air pollution in very large populations in Canada, the United States, and Europe, respectively (see Preface Table and Preface Figure 1). The studies included large population cohorts (with detailed individual information about potential confounders on all or a subset of the cohort) as well as large administrative databases with greater statistical power (albeit with less individual covariate information). Additionally, the three teams employed satellite data and ground-level exposure measurements, used high-quality exposure assessment models at high spatial resolutions, and set out to develop and apply novel statistical methods.

The three studies were expected to inform the scientific community and risk assessors and policy makers regarding exposure–response functions at levels of

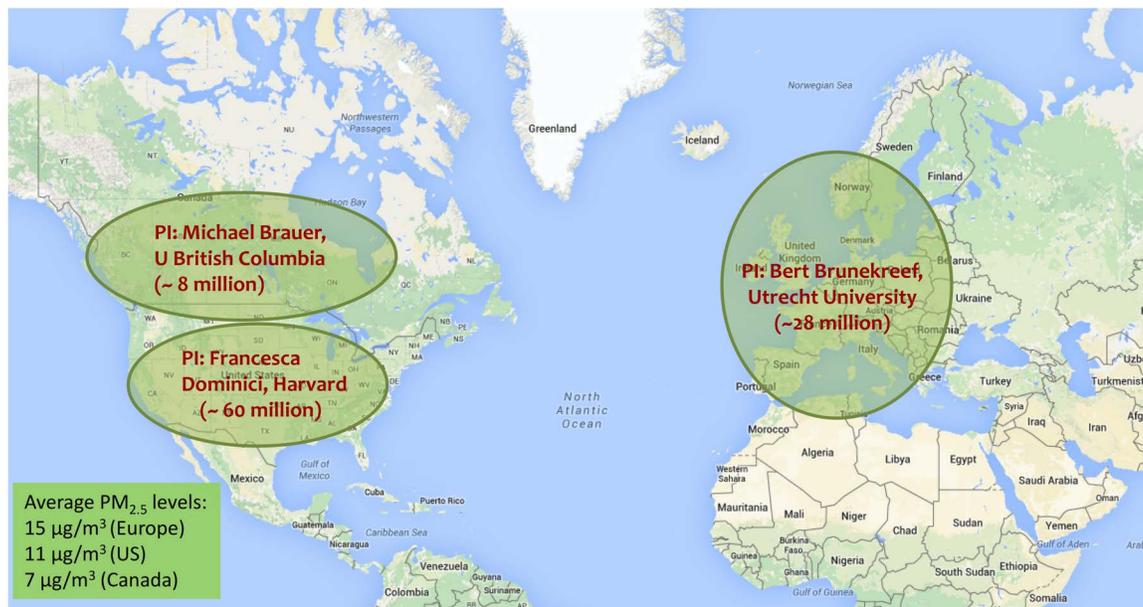
ambient air pollution currently prevalent in North America, Western Europe, and other developed regions. The full sets of analyses were expected to be completed in 2021 (see below).

**CANADIAN STUDY (MICHAEL BRAUER ET AL.)**

Brauer and colleagues proposed to assess the relationship between nonaccidental mortality and long-term exposure to low concentrations of PM<sub>2.5</sub> in four large population-based cohorts, including a careful characterization of the shape of the exposure–response function. The investigators used Canadian census data and had access to a nationally representative population of approximately 8.5 million Canadians (ages 25–90 yr) (Preface Figure 2). The Canadian team proposed developing hybrid models using primarily satellite data, as well as chemical transport models, land use variables, and routinely collected monitoring data for PM<sub>2.5</sub>, as well as estimating exposures for NO<sub>2</sub> and O<sub>3</sub> for Canada and the United States during the period 1981–2016. Additionally, they planned to validate satellite data against ground-based monitors in Canada as part of the SPARTAN network (Snider et al. 2015).

**Preface Table.** HEI's Research Program to Assess Adverse Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution

Investigator (institution)	Study Title	Phase 1 Report	Final Report Published
Brauer, Michael (University of British Columbia, Canada)	Mortality–Air Pollution Associations in Low Exposure Environments (MAPLE)	HEI Research Report 213 (2019)	Early 2022 (in review)
Brunekreef, Bert (Utrecht University, the Netherlands)	Mortality and Morbidity Effects of Long-Term Exposure to Low-Level PM <sub>2.5</sub> , BC, NO <sub>2</sub> , and O <sub>3</sub> : An Analysis of European Cohorts in the ELAPSE Project	None	September 2021 (HEI Research Report 208)
Dominici, Francesca (Harvard T.H. Chan School of Public Health, USA)	Assessing Adverse Health Effects of Long- Term Exposure to Low Levels of Ambient Pollution	HEI Research Report 200 (2019)	Late 2021 (in review)



Preface Figure 2. Geographical areas and populations covered by HEI’s Research Program to assess adverse effects of long-term exposure to low levels of ambient air pollution.

The exposure models were to be applied to estimate effects of air pollution exposure on all-cause and cause-specific mortality in four Canadian cohorts:

1. About 2.6 million subjects who completed the 1991 census long-form of the Canadian Census Health & Environment Cohorts (CanCHEC),
2. About 3.5 million subjects who completed the 1996 CanCHEC census long-form,
3. About 3.5 million subjects who completed the 2001 CanCHEC census long-form, and
4. About 540,000 subjects who participated in the Canadian Community Health Survey (CCHS) between 2001 and 2012, and reported individual-level risk factors, including smoking.

**EUROPEAN STUDY (BERT BRUNEKREEF ET AL.)**

Brunekreef and colleagues based their study on the European Study of Cohorts for Air Pollution Effects (ESCAPE), which started about a decade ago; its results have been published widely (e.g., Beelen et al. 2014a, b; Cesaroni et al. 2014; Eeftens et al. 2012a, b). In the

current HEI-funded study, the investigators proposed to analyze pooled data from 10 ESCAPE cohorts (instead of the cohort-specific approach they used previously). In addition, they planned to use data from six large administrative cohorts to yield a total study population of approximately 28 million Europeans (Preface Figure 2). They proposed developing hybrid Europe-wide and location-specific exposure models that would utilize land use information, dispersion modeling, satellite data, ESCAPE monitoring data, and routinely collected monitoring data for PM<sub>2.5</sub>, NO<sub>2</sub>, O<sub>3</sub>, and black carbon at high spatial resolution (residential address level; such detailed information is very difficult to obtain in the United States).

Brunekreef and colleagues proposed to investigate the following health outcomes: all-cause and cause-specific mortality, incidence of coronary and cerebrovascular events, and lung cancer incidence. The incorporation of ESCAPE cohorts with individual covariate information as well as very large administrative cohorts (albeit with less detailed information) will provide new insights into the merits of both approaches.

### UNITED STATES STUDY (FRANCESCA DOMINICI ET AL.)

Dominici and colleagues proposed to evaluate Medicare and Medicaid data for a study population of approximately 60 million Americans (Preface Figure 2). They planned to develop hybrid exposure models that incorporate satellite data, chemical transport models, land use, and weather variables, and routinely collected monitoring data for PM<sub>2.5</sub> and its components, NO<sub>2</sub>, and O<sub>3</sub>, at high spatial resolution (1-km<sup>2</sup> grid) for the continental United States during the period 2000–2014. Exposure models were to be applied to estimate adverse health effects of air pollution in three cohorts:

1. Medicare enrollees (28.6 million elderly enrollees per year, 2000–2014);
2. Medicaid enrollees (28 million enrollees per year, 2010–2014); and
3. Medicare Current Beneficiary Survey enrollees (nationally representative sample of approximately 15,000 enrollees per year with rich individual-level risk factor information, including smoking).

Dominici and colleagues planned to analyze the following health outcomes: time to hospitalization by cause, disease progression (time to rehospitalization), and time to death. They proposed developing and applying new causal inference methods to estimate exposure–response functions to adjust for confounding and exposure measurement error. Additionally, they proposed developing tools for reproducible research including approaches for data sharing, record linkage, and statistical software.

---

### STUDY OVERSIGHT

HEI's independent Low Exposure Epidemiology Oversight Panel provided advice and feedback on the study designs, analytical plans, and study progress throughout the duration of the research program (see Contributors' page).

Given the substantial challenges in conducting a systematic analysis to assess health effects of long-term exposure to low levels of ambient air pollution, HEI worked actively with the study teams to coordinate their efforts and ensure the maximum degree of

comparable epidemiologic results at the end of this research effort. To this end, HEI regularly held investigator workshops and site visits, among other activities. In addition, the studies were subject to HEI's special quality assurance procedures that included an audit by an independent audit team (see [www.healtheffects.org/research/quality-assurance](http://www.healtheffects.org/research/quality-assurance)).

---

### REVIEW OF PHASE I AND FINAL (PHASE 2) REPORTS

To inform the ongoing review of the U.S. National Ambient Air Quality Standards (NAAQS) for PM<sub>2.5</sub> and O<sub>3</sub> during 2019–2020, HEI requested Phase I reports based on the research completed during the first two years of the Canadian and U.S. studies. Thus, the Phase I reports by Drs. Brauer and Dominici provided summaries of results to date, including those published in journal articles, with accompanying Commentaries by an independent Special Review Panel. These Phase I reports provided an opportunity to review the results to date and evaluate their strengths and weaknesses, a process normally performed after a study has been completed.

As is common for major research programs, HEI convened a Special Review Panel to independently review the Phase I reports by Drs. Brauer and Dominici. The Panel consists of seven experts in epidemiology, exposure assessment, and biostatistics (see Contributors' page). The Panel also reviewed the final (Phase 2) reports of the three studies.

The three studies commenced in Spring 2016 and were completed at different times in 2020, with final reports published during 2021. In addition, further analyses, for example to compare approaches among the three teams, are ongoing and are expected to be completed at the end of 2021.

---

### REFERENCES

Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B, et al. 2014a. Effects of long-term exposure to air pollution on natural-cause mortality: An analysis of 22 European cohorts within the multicentre ESCAPE project. *Lancet* 383:785–795.

## Preface

---

Beelen R, Stafoggia M, Raaschou-Nielsen O, Andersen ZJ, Xun WW, Katsouyanni K, et al. 2014b. Long-term exposure to air pollution and cardiovascular mortality: an analysis of 22 European cohorts. *Epidemiology* 25:368–378; doi: 10.1097/EDE.0000000000000076.

Cesaroni G, Forastiere F, Stafoggia M, Andersen ZJ, Badaloni C, Beelen R, et al. 2014. Long term exposure to ambient air pollution and incidence of acute coronary events: Prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE project. *BMJ* 348:f7412; doi: 10.1136/bmj.f7412.

Crouse DL, Peters PA, van Donkelaar A, Goldberg MS, Villeneuve PJ, Brion O, et al. 2012. Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: A Canadian national-level cohort study. *Environ Health Perspectives* 120:708–714.

Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. 2012a. Development of land use regression models for  $PM_{2.5}$ ,  $PM_{2.5}$  absorbance,  $PM_{10}$

and  $PM_{coarse}$  in 20 European study areas: Results of the ESCAPE project. *Environ Sci Technol* 46:11195–11205; doi: 10.1021/es301948k.

Eeftens M, Tsai MY, Ampe C, Anwander B, Beelen R, Bellander T, et al. 2012b. Spatial variation of  $PM_{2.5}$ ,  $PM_{10}$ ,  $PM_{2.5}$  absorbance and  $PM_{coarse}$  concentrations between and within 20 European study areas and the relationship with  $NO_2$ : Results of the ESCAPE project. *Atmos Environ* 62:303–317.

Hales S, Blakely T, Woodward A. 2012. Air pollution and mortality in New Zealand: Cohort study. *J Epidemiol Community Health* 66:468–473; doi: 10.1136/jech.2010.112490.

Snider G, Weagle CL, Martin RV, van Donkelaar A, Conrad K, Cunningham D, et al. 2015. SPARTAN: A global network to evaluate and enhance satellite-based estimates of ground-level particulate matter for global health applications. *Atmos Meas Tech* 8:505–521; doi.org/10.5194/amt-8-505-2015.

# HEI STATEMENT

## Synopsis of Research Report 208

### Effects of Low-Level Air Pollution: A Study in Europe (ELAPSE)

#### BACKGROUND

The growing scientific evidence for effects of air pollution on health at concentrations below current air quality standards and the large burden of disease attributed to air pollution suggest that more stringent air quality standards and guidelines might be warranted in the future. To improve our understanding of exposure–response functions for mortality and morbidity at low concentrations of PM<sub>2.5</sub>, NO<sub>2</sub>, O<sub>3</sub>, and other ambient air pollutants, HEI issued RFA 14-3, *Assessing Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution*. Three studies based in the United States, Canada, and Europe were funded that used state-of-the-art exposure methods and large cohorts in high-income countries where ambient concentrations are generally low, that is, lower than current air quality guidelines and standards for Europe and the United States for PM<sub>2.5</sub>, NO<sub>2</sub>, and O<sub>3</sub>. HEI convened an independent Special Review Panel to evaluate the studies' strengths and weaknesses. This Statement highlights results from the European study.

#### APPROACH

The ELAPSE study by Brunekreef and colleagues examined whether exposure to low concentrations of ambient air pollution is associated with adverse effects on human health in 22 European cohorts. The study focused on four pollutants — PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub> — and developed new exposure models that combined monitoring data, land use data, satellite observations, and dispersion models to estimate exposures for cohort participants throughout Europe. The study had the following objectives:

1. To estimate long-term average exposure to air pollution in seven large administrative cohorts and in a pooled cohort consisting of participants

#### What This Study Adds

- The ELAPSE study evaluated health outcomes related to variation in exposures to low ambient air pollution concentrations, below current international guidelines.
- Brunekreef and colleagues developed new exposure models for all of Europe at a spatial resolution of 100 m × 100 m for four pollutants, namely fine particulate matter (PM<sub>2.5</sub>), black carbon (BC), nitrogen dioxide (NO<sub>2</sub>), and ozone (O<sub>3</sub>), as well as PM<sub>2.5</sub> particle composition.
- They used data from 11 European countries to analyze (1) a pooled cohort of 15 well-characterized cohorts and (2) seven large administrative cohorts individually, followed by a meta-analysis.
- For both approaches, they reported that exposure to PM<sub>2.5</sub>, BC, and NO<sub>2</sub> was significantly associated with natural-cause, cardiovascular, respiratory, and lung cancer mortality. They also reported inverse associations between O<sub>3</sub> and all causes of death examined.
- The shape of the associations with natural-cause mortality showed steeper slopes at lower exposures, with no evidence of concentrations below which no associations were found for PM<sub>2.5</sub>, BC, and NO<sub>2</sub>.
- This study contributes important evidence of associations between long-term exposures to relatively low concentrations of ambient air pollution and several important health endpoints.

This Statement, prepared by the Health Effects Institute, summarizes a research project funded by HEI and conducted by Dr. Bert Brunekreef at the Institute for Risk Assessment Sciences, Utrecht University, the Netherlands, and colleagues. Research Report 208 contains both the detailed Investigators' Report and a Commentary on the study prepared by the Institute's Low-Exposure Epidemiology Studies Review Panel.  
BRUNEKREEF 208

from 15 existing cohorts from the European Study of Cohorts for Air Pollution Effects (ESCAPE).

2. To investigate the shape of the concentration–response relationship between long-term exposure to these pollutants and four broad categories of health outcomes: (1) natural and cause-specific mortality, (2) coronary and cerebrovascular events, (3) lung cancer incidence, and (4) asthma and chronic obstructive pulmonary disease incidence.
3. To investigate variability of the concentration–response function across populations and different exposure assessment methods, the effect of different methods for addressing exposure measurement error, the role of co-occurring pollutants, and the effect of indirect approaches for confounder control.
4. To compare epidemiological effect estimates for those obtained using the ELAPSE PM<sub>2.5</sub> exposure model with those obtained using an exposure model developed through the Mortality–Air Pollution Associations in Low-Exposure Environments (MAPLE) study in Canada that was funded under the same RFA.

The ELAPSE study consists of two parallel sets of epidemiologic analyses. First, Brunekreef and colleagues created Europe-wide exposure models for all pollutants of interest with monitoring data from 2010 at a spatial resolution of 100 m × 100 m. They then assigned estimates of exposure to participants in two sets of cohorts:

- They analyzed a pooled cohort that included 15 conventional research cohorts (i.e., those for which individuals were invited to participate and to respond to questionnaires). Most of those cohorts are located in a region that included at least one large city and a surrounding smaller town. The key strength of the pooled cohort approach is the rich amount of individual-level information available for about 325,000 participants.
- They analyzed seven large administrative cohorts individually and then conducted a meta-analysis to produce overall results. The administrative cohorts were formed by linking census data, population registries, and death registries. The key strength of the administrative cohorts is their large sample size (about 28 million total) and national representativeness.

The investigators applied standard Cox proportional hazard models to describe associations between exposures to the pollutants and the health outcomes of interest. Briefly, all models included age, sex, calendar year of enrollment, and selected individual and area-level (i.e., neighborhood or community-level) information, with slight variations in model specification for the two approaches. The investigators conducted many sensitivity analyses to evaluate exposure specification, confounder control, and various concentration–response functions.

### KEY RESULTS

The Europe-wide exposure models explained 51% to 66% of the variability in concentrations of PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub>, with good spatial stability (across the full study area) and good temporal stability (throughout the study period). In 2010, almost all participants had annual average exposures below the European Union limit values for PM<sub>2.5</sub> (25 µg/m<sup>3</sup>) and NO<sub>2</sub> (40 µg/m<sup>3</sup>), and about 14% had exposures below the U.S. National Ambient Air Quality Standards for PM<sub>2.5</sub> (12 µg/m<sup>3</sup>). Participants in the pooled cohort were exposed on average to 15 µg/m<sup>3</sup> PM<sub>2.5</sub>, 1.5 × 10<sup>-5</sup>/m BC, 25 µg/m<sup>3</sup> (13 ppb) NO<sub>2</sub>, and 67 µg/m<sup>3</sup> (34 ppb) O<sub>3</sub>. Among the administrative cohorts, mean concentrations of PM<sub>2.5</sub> ranged from 12 to 19 µg/m<sup>3</sup>, except for the Norwegian cohort (8 µg/m<sup>3</sup>).

In analyses with both sets of cohorts, Brunekreef and colleagues reported significant positive associations between PM<sub>2.5</sub>, BC, and NO<sub>2</sub> and natural-cause (Statement Figure) cardiovascular, respiratory, and lung cancer mortality. They also reported inverse associations between O<sub>3</sub> and all causes of death examined. The estimated risks associated with exposure were generally greater in the pooled cohort than in the administrative cohorts.

The hazard ratios for natural-cause mortality remained elevated and significant for PM<sub>2.5</sub> even when the analyses were restricted to observations below 12 µg/m<sup>3</sup>. For NO<sub>2</sub>, hazard ratios remained elevated and significant when analyses were restricted to observations below 20 µg/m<sup>3</sup>.

In the pooled cohort, the investigators found significant positive associations between PM<sub>2.5</sub>, BC, and NO<sub>2</sub> and incidence of stroke, asthma, and COPD hospital admissions. Additionally, they reported significant associations between NO<sub>2</sub> and

acute coronary heart disease and between  $PM_{2.5}$  and lung cancer incidence.

In spline plots examining the shape of associations between exposure and natural-cause mortality, the investigators observed generally supralinear patterns (i.e., steeper slopes at lower exposures) with no evidence of concentrations below which no associations were found for  $PM_{2.5}$ , BC, and  $NO_2$ . That is, they found increased risks for mortality at even the lowest observed concentrations.

Lastly, in both sets of cohorts, the investigators reported comparable associations between mortality and exposures to  $PM_{2.5}$  using the exposure estimates from the MAPLE study.

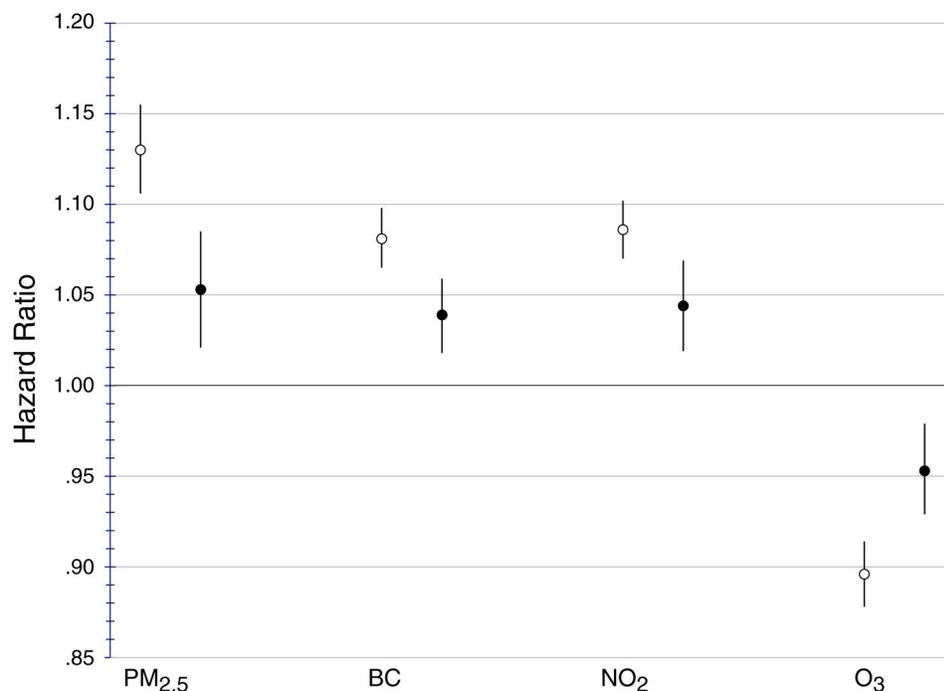
### INTERPRETATION AND CONCLUSIONS

A key highlight of this study is the presentation of results from 11 European countries (pooled and individually), including heterogeneity in associations across countries. The creation of exposure

models at  $100\text{ m} \times 100\text{ m}$  spatial resolution across Europe and analyses of associations between four pollutants and many important health endpoints, using two parallel cohort approaches, are all impressive achievements that provide important evidence for effects on health at low concentrations of exposures.

The analyses restricted to participants with the very lowest exposures provide further support that exposure to air pollution at low concentrations is associated with adverse health outcomes. However, it is important to acknowledge that these analyses were based on smaller numbers of cohorts that were less heterogeneous. The analysis at the lowest concentrations (below  $10\text{ }\mu\text{g}/\text{m}^3$ ) in particular included data only from Norway and Stockholm. Thus, those findings might not be generalizable to the broader population.

Generally, the investigators have carefully explored several approaches to modeling concentration-response functions. However, the heterogeneity in



**Statement Figure.** Associations between air pollution and all-natural causes of death among participants in the pooled (white) and administrative (black) ELAPSE cohorts. Note that hazard ratios for the four pollutants were calculated over different concentration increments (see Tables 5 and 19 in the Investigators' Report).

the shapes of those functions for the administrative cohorts was not explained well (beyond acknowledging that the cohorts differed in mean exposures). On the other hand, the fact that the associations observed with the exposure model developed for the MAPLE study were similar to those observed with the exposure model developed for ELAPSE corroborated the robustness of the associations. The near-consistent inverse associations between  $O_3$  and the risk of the various health outcomes were unexpected and remained largely unexplained.

In summary, this study has provided important evidence of associations between long-term exposures to low concentrations of  $PM_{2.5}$ , BC, and  $NO_2$  and various health outcomes, including mortality. Evidence for associations at the lowest concentrations remains limited because those analyses were based primarily on data from Norway and Stockholm. Continuing research on the effects of low concentrations of air pollutants in North America and Europe is expected to further inform the process of setting air quality standards in those and other global regions.

## Mortality and Morbidity Effects of Long-Term Exposure to Low-Level PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub>: An Analysis of European Cohorts in the ELAPSE Project

Bert Brunekreef<sup>1</sup>, Maciej Strak<sup>1,2</sup>, Jie Chen<sup>1</sup>, Zorana J. Andersen<sup>3</sup>, Richard Atkinson<sup>4</sup>, Mariska Bauwelinck<sup>5</sup>, Tom Bellander<sup>6</sup>, Marie-Christine Boutron<sup>7</sup>, Jørgen Brandt<sup>8</sup>, Iain Carey<sup>4</sup>, Giulia Cesaroni<sup>9</sup>, Francesco Forastiere<sup>9,10</sup>, Daniela Fecht<sup>11</sup>, John Gulliver<sup>11,12</sup>, Ole Hertel<sup>8</sup>, Barbara Hoffmann<sup>13</sup>, Kees de Hoogh<sup>14</sup>, Danny Houthuijs<sup>2</sup>, Ulla Hvidtfeldt<sup>15</sup>, Nicole Janssen<sup>2</sup>, Jeanette Jørgensen<sup>15</sup>, Klea Katsouyanni<sup>10,16</sup>, Matthias Ketzel<sup>8</sup>, Jochem Klompmaker<sup>2,17</sup>, Norun Hjertager Krog<sup>18</sup>, Shuo Liu<sup>3</sup>, Petter Ljungman<sup>6</sup>, Amar Mehta<sup>3,19</sup>, Gabriele Nagel<sup>20,21</sup>, Bente Oftedal<sup>18</sup>, Göran Pershagen<sup>6</sup>, Annette Peters<sup>22,23</sup>, Ole Raaschou-Nielsen<sup>15</sup>, Matteo Renzi<sup>9</sup>, Sophia Rodopoulou<sup>16</sup>, Evi Samoli<sup>16</sup>, Per Schwarze<sup>18</sup>, Torben Sigsgaard<sup>8</sup>, Massimo Stafoggia<sup>9</sup>, Danielle Vienneau<sup>14</sup>, Gudrun Weinmayr<sup>19</sup>, Kathrin Wolf<sup>22</sup>, and Gerard Hoek<sup>1</sup>

<sup>1</sup>Institute for Risk Assessment Sciences, Utrecht University, the Netherlands; <sup>2</sup>National Institute for Public Health and the Environment, Bilthoven, the Netherlands; <sup>3</sup>Department of Public Health, University of Copenhagen, Copenhagen, Denmark; <sup>4</sup>Population Health Research, Institute St George's, University of London, London, UK; <sup>5</sup>Interface Demography – Department of Sociology, Vrije Universiteit Brussel, Brussels, Belgium; <sup>6</sup>Institute of Environmental Medicine, Karolinska Institutet, and Centre for Occupational and Environmental Medicine, Region Stockholm, Stockholm, Sweden; <sup>7</sup>Centre for Research in Epidemiology and Population Health, INSERM, Villejuif, France; <sup>8</sup>Department of Environmental Science, Aarhus University, Roskilde, Denmark; <sup>9</sup>Department of Epidemiology (DEP) Lazio Regional Health Service, Rome, Italy; <sup>10</sup>Science Policy & Epidemiology Environmental Research Group King's College London, London UK; <sup>11</sup>MRC Centre for Environment and Health, School of Public Health, Imperial College London, London, UK; <sup>12</sup>Centre for Environmental Health and Sustainability, University of Leicester, Leicester, UK; <sup>13</sup>Institute for Occupational, Social and Environmental Medicine, University of Duesseldorf, Duesseldorf, Germany; <sup>14</sup>Swiss Tropical and Public Health Institute, Basel, Switzerland; <sup>15</sup>Danish Cancer Society Research Center, Copenhagen, Denmark; <sup>16</sup>Department of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, Athens, Greece; <sup>17</sup>Harvard T.H. Chan School of Public Health, Boston, MA, USA; <sup>18</sup>Department of Environmental Health, Norwegian Institute of Public Health, Oslo, Norway; <sup>19</sup>Methods and Analysis, Statistics Denmark, Copenhagen, Denmark; <sup>20</sup>Institute for Epidemiology and Medical Biometry, Ulm University, Ulm, Germany; <sup>21</sup>Agency for Preventive and Social Medicine, Bregenz, Austria; <sup>22</sup>Institute of Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany; <sup>23</sup>Department of Epidemiology, Ludwig Maximilians Universität München, Munich, Germany

### ABSTRACT

**Introduction** Epidemiological cohort studies have consistently found associations between long-term exposure to outdoor air pollution and a range of morbidity and

mortality endpoints. Recent evaluations by the World Health Organization and the Global Burden of Disease study have suggested that these associations may be nonlinear and may persist at very low concentrations. Studies conducted in North America in particular have suggested that associations with mortality persisted at concentrations of particulate matter with an aerodynamic diameter of less than 2.5 µm (PM<sub>2.5</sub>\*) well below current air quality standards and guidelines. The uncertainty about the shape of the concentration–response function at the low end of the concentration distribution, related to the scarcity of observations in the lowest range, was the basis of the current project. Previous studies have focused on PM<sub>2.5</sub>, but increasingly associations with nitrogen dioxide (NO<sub>2</sub>) are being reported, particularly in studies that accounted for the fine spatial

This Investigators' Report is one part of Health Effects Institute Research Report 208, which also includes a Commentary by the Low-Exposure Epidemiology Studies Review Panel and an HEI Statement about the research project. Correspondence concerning the Investigators' Report may be addressed to Dr. Bert Brunekreef, Address Utrecht University, PO Box 80176 3508 TD Utrecht, the Netherlands; e-mail: [b.brunekreef@uu.nl](mailto:b.brunekreef@uu.nl). No potential conflict of interest was reported by the authors.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83467701 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

\* A list of abbreviations and other terms appears at the end of this volume.

scale variation of NO<sub>2</sub>. Very few studies have evaluated the effects of long-term exposure to low concentrations of ozone (O<sub>3</sub>). Health effects of black carbon (BC), representing primary combustion particles, have not been studied in most large cohort studies of PM<sub>2.5</sub>. Cohort studies assessing health effects of particle composition, including elements from nontailpipe traffic emissions (iron, copper, and zinc) and secondary aerosol (sulfur) have been few in number and reported inconsistent results.

The overall objective of our study was to investigate the shape of the relationship between long-term exposure to four pollutants (PM<sub>2.5</sub>, NO<sub>2</sub>, BC, and O<sub>3</sub>) and four broad health effect categories using a number of different methods to characterize the concentration–response function (i.e., linear, nonlinear, or threshold). The four health effect categories were (1) natural- and cause-specific mortality including cardiovascular and nonmalignant as well as malignant respiratory and diabetes mortality; and morbidity measured as (2) coronary and cerebrovascular events; (3) lung cancer incidence; and (4) asthma and chronic obstructive pulmonary disease (COPD) incidence. We additionally assessed health effects of PM<sub>2.5</sub> composition, specifically the copper, iron, zinc, and sulfur content of PM<sub>2.5</sub>.

**Methods** We focused on analyses of health effects of air pollutants at low concentrations, defined as less than current European Union (EU) Limit Values, U.S. Environmental Protection Agency (U.S. EPA), National Ambient Air Quality Standards (NAAQS), and/or World Health Organization (WHO) Air Quality Guideline values for PM<sub>2.5</sub>, NO<sub>2</sub>, and O<sub>3</sub>.

We address the health effects at low air pollution levels by performing new analyses within selected cohorts of the ESCAPE study (European Study of Cohorts for Air Pollution Effects; Beelen et al. 2014a) and within seven very large European administrative cohorts. By combining well-characterized ESCAPE cohorts and large administrative cohorts in one study the strengths and weaknesses of each approach can be addressed. The large administrative cohorts are more representative of national or citywide populations, have higher statistical power, and can efficiently control for area-level confounders, but have fewer possibilities to control for individual-level confounders. The ESCAPE cohorts have detailed information on individual confounders, as well as country-specific information on area-level confounding.

The data from the seven included ESCAPE cohorts and one additional non-ESCAPE cohort have been pooled and analyzed centrally. More than 300,000 adults were included in the pooled cohort from existing cohorts in

Sweden, Denmark, Germany, the Netherlands, Austria, France, and Italy. Data from the administrative cohorts have been analyzed locally, without transfer to a central database. Privacy regulations prevented transfer of data from administrative cohorts to a central database. More than 28 million adults were included from national administrative cohorts in Belgium, Denmark, England, the Netherlands, Norway, and Switzerland as well as an administrative cohort in Rome, Italy.

We developed central exposure assessment using Europe-wide hybrid land use regression (LUR) models, which incorporated European routine monitoring data for PM<sub>2.5</sub>, NO<sub>2</sub>, and O<sub>3</sub>, and ESCAPE monitoring data for BC and PM<sub>2.5</sub> composition, land use, and traffic data supplemented with satellite observations and chemical transport model estimates. For all pollutants, we assessed exposure at a fine spatial scale, 100 × 100 m grids. These models have been applied to individual addresses of all cohorts including the administrative cohorts. In sensitivity analyses, we applied the PM<sub>2.5</sub> models developed within the companion HEI-funded Canadian MAPLE study (Brauer et al. 2019) and O<sub>3</sub> exposures on a larger spatial scale for comparison with previous studies.

Identification of outcomes included linkage with mortality, cancer incidence, hospital discharge registries, and physician-based adjudication of cases. We analyzed natural-cause, cardiovascular, ischemic heart disease, stroke, diabetes, cardiometabolic, respiratory, and COPD mortality. We also analyzed lung cancer incidence, incidence of coronary and cerebrovascular events, and incidence of asthma and COPD (pooled cohort only).

We applied the Cox proportional hazard model with increasing control for individual- and area-level covariates to analyze the associations between air pollution and mortality and/or morbidity for both the pooled cohort and the individual administrative cohorts. Age was used as the timescale because of evidence that this results in better adjustment for potential confounding by age. 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. A priori we specified three confounder models, following the modeling methods of the ESCAPE study. Model 1 included only age (time axis), sex (as strata), and calendar year of enrollment. Model 2 added individual-level variables that were consistently available in the cohorts contributing to the pooled cohort or all variables available in the administrative cohorts, respectively. Model 3 further added area-level socioeconomic status (SES) variables. A priori model 3 was selected as the main model. All analyses in the pooled cohort were stratified by sub-

cohort. All analyses in the administrative cohorts accounted for clustering of the data in neighborhoods by adjusting the variance of the effect estimates. The main exposure variable we analyzed was derived from the Europe-wide hybrid models based on 2010 monitoring data. Sensitivity analyses were conducted using earlier time periods, time-varying exposure analyses, local exposure models, and the PM<sub>2.5</sub> models from the Canadian MAPLE project.

We first specified linear single-pollutant models. Two-pollutant models were specified for all combinations of the four main pollutants. Two-pollutant models for particle composition were analyzed with PM<sub>2.5</sub> and NO<sub>2</sub> as the second pollutant. We then investigated the shape of the concentration–response function using natural splines with two, three, and four degrees of freedom; penalized splines with the degrees of freedom determined by the algorithm and shape-constrained health impact functions (SCHIF) using confounder model 3. Additionally, we specified linear models in subsets of the concentration range, defined by removing concentrations above a certain value from the analysis, such as for PM<sub>2.5</sub> 25 µg/m<sup>3</sup> (EU limit value), 20, 15, 12 µg/m<sup>3</sup> (U.S. EPA National Ambient Air Quality Standard), and 10 µg/m<sup>3</sup> (WHO Air Quality Guideline value). Finally, threshold models were evaluated to investigate whether the associations persisted below specific concentration values. For PM<sub>2.5</sub>, we evaluated 10, 7.5, and 5 µg/m<sup>3</sup> as potential thresholds. Performance of threshold models versus the corresponding no-threshold linear model were evaluated using the Akaike information criterion (AIC).

**Results** In the pooled cohort, virtually all subjects in 2010 had PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below the EU limit values (25 µg/m<sup>3</sup> and 40 µg/m<sup>3</sup>, respectively). More than 50,000 had a residential PM<sub>2.5</sub> exposure below the U.S. EPA NAAQS (12 µg/m<sup>3</sup>). More than 25,000 subjects had a residential PM<sub>2.5</sub> exposure below the WHO guideline (10 µg/m<sup>3</sup>). We found significant positive associations between PM<sub>2.5</sub>, NO<sub>2</sub>, and BC and natural-cause, respiratory, cardiovascular, and diabetes mortality. In our main model, the hazard ratios (HRs) (95% [confidence interval] CI) were 1.13 (CI = 1.11, 1.16) for an increase of 5 µg/m<sup>3</sup> PM<sub>2.5</sub>, 1.09 (CI = 1.07, 1.10) for an increase of 10 µg/m<sup>3</sup> NO<sub>2</sub>, and 1.08 (CI = 1.06, 1.10) for an increase of 0.5 × 10<sup>-5</sup>/m BC for natural-cause mortality. The highest HRs were found for diabetes mortality. Associations with O<sub>3</sub> were negative, both in the fine spatial scale of the main ELAPSE model and in large spatial scale exposure models. For PM<sub>2.5</sub>, NO<sub>2</sub>, and BC, we generally observed a supralinear association with steeper slopes at low exposures and no evidence of a concentration below which no

association was found. Subset analyses further confirmed that these associations remained at low levels: below 10 µg/m<sup>3</sup> for PM<sub>2.5</sub> and 20 µg/m<sup>3</sup> for NO<sub>2</sub>. HRs were similar to the full cohort HRs for subjects with exposures below the EU limit values for PM<sub>2.5</sub> and NO<sub>2</sub>, the U.S. NAAQS values for PM<sub>2.5</sub>, and the WHO guidelines for PM<sub>2.5</sub> and NO<sub>2</sub>. The mortality associations were robust to alternative specifications of exposure, including different time periods, PM<sub>2.5</sub> from the MAPLE project, and estimates from the local ESCAPE model. Time-varying exposure natural spline analyses confirmed associations at low pollution levels. HRs in two-pollutant models were attenuated but remained elevated and statistically significant for PM<sub>2.5</sub> and NO<sub>2</sub>. In two-pollutant models of PM<sub>2.5</sub> and NO<sub>2</sub> HRs for natural-cause mortality were 1.08 (CI = 1.05, 1.11) for PM<sub>2.5</sub> and 1.05 (CI = 1.03, 1.07) for NO<sub>2</sub>. Associations with O<sub>3</sub> were attenuated but remained negative in two-pollutant models with NO<sub>2</sub>, BC, and PM<sub>2.5</sub>.

We found significant positive associations between PM<sub>2.5</sub>, NO<sub>2</sub>, and BC and incidence of stroke and asthma and COPD hospital admissions. Furthermore, NO<sub>2</sub> was significantly related to acute coronary heart disease and PM<sub>2.5</sub> was significantly related to lung cancer incidence. We generally observed linear to supralinear associations with no evidence of a threshold, with the exception of the association between NO<sub>2</sub> and acute coronary heart disease, which was sublinear. Subset analyses documented that associations remained even with PM<sub>2.5</sub> below 20 µg/m<sup>3</sup> and possibly 12 µg/m<sup>3</sup>. Associations remained even when NO<sub>2</sub> was below 30 µg/m<sup>3</sup> and in some cases 20 µg/m<sup>3</sup>. In two-pollutant models, NO<sub>2</sub> was most consistently associated with acute coronary heart disease, stroke, asthma, and COPD hospital admissions. PM<sub>2.5</sub> was not associated with these outcomes in two-pollutant models with NO<sub>2</sub>. PM<sub>2.5</sub> was the only pollutant that was associated with lung cancer incidence in two-pollutant models. Associations with O<sub>3</sub> were negative though generally not statistically significant.

In the administrative cohorts, virtually all subjects in 2010 had PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below the EU limit values. More than 3.9 million subjects had a residential PM<sub>2.5</sub> exposure below the U.S. EPA NAAQS (12 µg/m<sup>3</sup>) and more than 1.9 million had residential PM<sub>2.5</sub> exposures below the WHO guideline (10 µg/m<sup>3</sup>). We found significant positive associations between PM<sub>2.5</sub>, NO<sub>2</sub>, and BC and natural-cause, respiratory, cardiovascular, and lung cancer mortality, with moderate to high heterogeneity between cohorts. We found positive but statistically nonsignificant associations with diabetes mortality. In our main model meta-analysis, the HRs (95% CI) for natural-cause mortality were 1.05 (CI = 1.02, 1.09) for an increase of 5 µg/m<sup>3</sup> PM<sub>2.5</sub>, 1.04 (CI = 1.02, 1.07) for an

increase of 10 µg/m<sup>3</sup> NO<sub>2</sub>, and 1.04 (CI = 1.02, 1.06) for an increase of 0.5 × 10<sup>-5</sup>/m BC, and 0.95 (CI = 0.93, 0.98) for an increase of 10 µg/m<sup>3</sup> O<sub>3</sub>.

The shape of the concentration–response functions differed between cohorts, though the associations were generally linear to supralinear, with no indication of a level below which no associations were found. Subset analyses documented that these associations remained at low levels: below 10 µg/m<sup>3</sup> for PM<sub>2.5</sub> and 20 µg/m<sup>3</sup> for NO<sub>2</sub>. BC and NO<sub>2</sub> remained significantly associated with mortality in two-pollutant models with PM<sub>2.5</sub> and O<sub>3</sub>. The PM<sub>2.5</sub> HR attenuated to unity in a two-pollutant model with NO<sub>2</sub>. The negative O<sub>3</sub> association was attenuated to unity and became nonsignificant. The mortality associations were robust to alternative specifications of exposure, including time-varying exposure analyses. Time-varying exposure natural spline analyses confirmed associations at low pollution levels. Effect estimates in the youngest participants (<65 years at baseline) were much larger than in the elderly (>65 years at baseline). Effect estimates obtained with the ELAPSE PM<sub>2.5</sub> model did not differ from the MAPLE PM<sub>2.5</sub> model on average, but in individual cohorts, substantial differences were found.

**Conclusions** Long-term exposure to PM<sub>2.5</sub>, NO<sub>2</sub>, and BC was positively associated with natural-cause and cause-specific mortality in the pooled cohort and the administrative cohorts. Associations were found well below current limit values and guidelines for PM<sub>2.5</sub> and NO<sub>2</sub>. Associations tended to be supralinear, with steeper slopes at low exposures with no indication of a threshold. Two-pollutant models documented the importance of characterizing the ambient mixture with both NO<sub>2</sub> and PM<sub>2.5</sub>. We mostly found negative associations with O<sub>3</sub>. In two-pollutant models with NO<sub>2</sub>, the negative associations with O<sub>3</sub> were attenuated to essentially unity in the mortality analysis of the administrative cohorts and the incidence analyses in the pooled cohort. In the mortality analysis of the pooled cohort, significant negative associations with O<sub>3</sub> remained in two-pollutant models.

Long-term exposure to PM<sub>2.5</sub>, NO<sub>2</sub>, and BC was also positively associated with morbidity outcomes in the pooled cohort. For stroke, asthma, and COPD, positive associations were found for PM<sub>2.5</sub>, NO<sub>2</sub>, and BC. For acute coronary heart disease, an increased HR was observed for NO<sub>2</sub>. For lung cancer, an increased HR was found only for PM<sub>2.5</sub>. Associations mostly showed steeper slopes at low exposures with no indication of a threshold.

---

## INTRODUCTION

---

Outdoor air pollution is a mixture of generally correlated pollutants, from a large number of sources including fossil fuel combustion in industry, power generation, motorized traffic, and (residential) heating as well as noncombustion sources such as emissions from livestock farming and the resuspension of soil and dust. Major health-relevant outdoor air pollutants include PM, NO<sub>2</sub>, and O<sub>3</sub> (Brook et al. 2010; Brunekreef and Holgate 2002). Both reviews provide further general background on outdoor air pollution exposure and studies of health effects.

Before we started this study, epidemiological cohort studies had consistently found associations between long-term exposure to outdoor air pollution and a range of morbidity and mortality endpoints. This was especially so for studies of the health effects of PM<sub>2.5</sub>, and to a lesser extent for studies of the health effects of NO<sub>2</sub> and O<sub>3</sub>. Landmark studies include the Harvard Six Cities study (Dockery et al. 1993), which documented an association between PM<sub>2.5</sub> and cardiorespiratory mortality in a cohort of subjects followed for 14 years, who lived in six U.S. cities. This association was observed over a range of about 10–26 µg/m<sup>3</sup> long-term PM<sub>2.5</sub> concentrations. Another landmark study was the American Cancer Society study, which was conducted in a very large cohort of about 500,000 subjects across the United States (Pope et al. 2002). This study also documented a strong association between long-term PM<sub>2.5</sub> and cardiorespiratory mortality as well as lung cancer mortality (Pope et al. 2002). Later work with this cohort also suggested an association between long-term O<sub>3</sub> concentrations and respiratory mortality (Jerrett et al. 2009). Cohort studies of associations between long-term exposure to NO<sub>2</sub> and mortality were, until recently, primarily conducted in Europe. Examples of large cohort studies demonstrating positive associations include the Netherlands Longitudinal Cohort Study (NLCS) (Beelen et al. 2008) and the Rome Longitudinal Study (ROLS) (Cesaroni et al. 2012). A review of multipollutant studies suggested that these associations were largely independent from those with PM<sub>2.5</sub> (Faustini et al. 2014).

The early cohort studies assessed air pollution exposure at the city level. More recent studies have also modeled within-urban exposures using LUR and/or dispersion models (Hoek 2017).

More recent evaluations by the WHO and the Global Burden of Disease study have suggested that these associations may be nonlinear and may persist at very low concentrations (Burnett et al. 2014; Cohen et al. 2017; WHO 2013). Studies conducted in Canada in particular had suggested that associations with mortality persisted at very

low PM<sub>2.5</sub> and NO<sub>2</sub> concentrations (Crouse et al. 2012, 2015). WHO is currently updating the Air Quality Guidelines, which were last revised in 2005 (WHO 2005). Systematic reviews of long-term effects of PM<sub>2.5</sub>, PM<sub>10</sub>, NO<sub>2</sub>, and O<sub>3</sub> were recently published (Chen and Hoek 2020; Huangfu and Atkinson 2020) that continue to support the findings of effects of long-term exposure to PM<sub>2.5</sub> and PM<sub>10</sub> at low to very low concentrations on mortality from natural causes, cardiovascular disease, diabetes, non-malignant respiratory disease, and lung cancer. These reviews also support effects of long-term exposure to NO<sub>2</sub> on natural-cause and respiratory mortality and to O<sub>3</sub> on respiratory mortality.

The uncertainty about the shape of the concentration–response function for the low end of the concentration distribution, which was related to the scarcity of observations in the lowest range, was the basis of the current project.

In this project we focused on analyses of health effects of spatially resolved air pollution concentrations at low concentrations, defined as less than current EU Limit Values, U.S. EPA NAAQS, and/or WHO guideline values for PM<sub>2.5</sub>, NO<sub>2</sub>, and O<sub>3</sub>. Most previous studies have focused on PM<sub>2.5</sub>, but increasingly associations with low-level concentrations of NO<sub>2</sub> are being reported, particularly in studies that accounted for the fine spatial scale variation of NO<sub>2</sub> (Crouse et al. 2015; Jerrett et al. 2013). Very few studies have evaluated long-term morbidity and mortality effects of long-term exposure to very low concentrations of O<sub>3</sub> (Atkinson et al. 2016).

We addressed the issue of health effects at low air pollution levels by performing targeted analyses of all-cause and cause-specific mortality and morbidity endpoints within selected cohorts of the ESCAPE study and seven large European administrative cohorts.

Within ESCAPE, exposure assessment was harmonized, based upon standardized measurements of PM<sub>2.5</sub>, PM<sub>10</sub>, BC, and NO<sub>2</sub> across Europe (Cyrys et al. 2012; Eeftens et al. 2012b). Cohort air pollution exposures were assessed with LUR models that were developed for each area to optimally predict outdoor concentrations in each study area (Beelen et al. 2013; de Hoogh et al. 2013; Eeftens et al. 2012a). In ESCAPE, we analyzed associations per cohort and then calculated combined effect estimates from the cohort-specific individual HRs using meta-analysis techniques. In the analysis, we thus exploited only within-cohort exposure contrasts. Consistent associations between air pollution and all-cause mortality, lung cancer incidence, and cardiovascular disease incidence (Beelen et al. 2014a; Cesaroni et al. 2014; Raaschou-Nielsen et al. 2013; Stafoggia et al. 2014) have been reported, whereas no associations were found for cardiovascular and respiratory

disease mortality for the ESCAPE study (Beelen et al. 2014b; Dimakopoulou et al. 2014). Subset analyses suggested that associations remained well below the EU limit values for PM<sub>2.5</sub> and NO<sub>2</sub>.

In ESCAPE, exposure to eight PM components was also quantified (de Hoogh et al. 2013; Tsai et al. 2015). A number of papers described associations between PM components and various health effects, such as nonaccidental mortality and lung cancer (Beelen et al. 2015; Raaschou-Nielsen et al. 2016; Wang et al. 2014; Wolf et al. 2015). We exploited these exposure data in the current project for additional analyses of health effects of PM elemental composition. These analyses were conducted because of significant uncertainty about the components in fine and coarse PM that affect PM health effects (Lippmann et al. 2013; Vedal et al. 2013; WHO 2013).

In some of the seven large administrative cohorts, dispersion models and LUR models based upon routine monitoring data or previous intense NO<sub>2</sub> monitoring campaigns have been applied in the past (Atkinson et al. 2013; Carey et al. 2013; Cesaroni et al. 2013; Fischer et al. 2015). Associations between air pollution and cause-specific mortality have been published for three of these large administrative cohorts: the Rome longitudinal study, the England national cohort, and the Netherlands national cohort (Atkinson et al. 2013; Carey et al. 2013; Cesaroni et al. 2013; Fischer et al. 2015). In the other four cohorts, associations between air pollution and mortality have not yet been studied.

The project was carried out in two stages. In the main phase, we focused on health effects of low levels of PM<sub>2.5</sub>, NO<sub>2</sub>, O<sub>3</sub>, and BC. In the extension phase, we performed analyses of the health effects of particle composition, PM<sub>2.5</sub> estimated using the Canadian MAPLE project methods, and added diabetes and cardiometabolic mortality as endpoints. The project started May 1, 2016, and ended April 30, 2020.

The report consists of a core report with key tables and figures and four appendices (found in Additional Materials on the HEI website). In Appendix 1 we provide more detail on the methods applied in this research. In Appendix 2 we provide more detailed results from analyses of the pooled cohort. In Appendix 3, we provide more results from the analysis of the administrative cohorts. In Appendix 4, we provide the methods and results of the particle composition analyses.

---

**OBJECTIVES**

---

The main hypothesis we investigated was that long-term exposure to low concentrations of outdoor air pollution is related to adverse health effects. We defined “low” using various cut-points defined by current EU limit values, U.S. air quality standards, and WHO guidelines. We assessed the shape of the concentration–response functions with a focus on the shape at low exposures. We exploited selected well-characterized cohorts from the ESCAPE study and seven large European administrative cohorts to study the health effects of low-level air pollution.

The specific objectives were as follows:

- To estimate long-term average exposure to PM<sub>2.5</sub>, PM<sub>2.5</sub> composition, NO<sub>2</sub>, O<sub>3</sub>, and BC by developing new hybrid models that combine monitoring data, land use, satellite observations, and dispersion models of the pooled ESCAPE cohort and seven large administrative cohorts.
- To investigate the shape of the relationship between long-term exposure to PM<sub>2.5</sub>, NO<sub>2</sub>, BC, and O<sub>3</sub> and four broad health effect categories — (1) natural- and cause-specific mortality including cardiovascular and non-malignant as well as malignant respiratory and diabetes mortality; and morbidity measured as, (2) coronary and cerebrovascular events; (3) lung cancer incidence; and (4) asthma and COPD incidence — using a number of different methods to characterize the concentration–response function (linear, nonlinear, or threshold).
- To investigate, in the context of the second objective, the variability of the exposure–response function across populations and different exposure assessment methods (as well as the impact of different methods for addressing exposure measurement error); the role of co-occurring pollutants; and the effect of indirect approaches for confounder control in administrative cohorts.
- To compare epidemiological effect estimates of those obtained using the ELAPSE and MAPLE exposure modeling frameworks (MAPLE is the Canadian companion study funded by HEI within RFA14-3 [Brauer et al. 2019]).

By performing targeted analyses of all-cause and cause-specific mortality and morbidity endpoints within selected cohorts of the ESCAPE study and seven large European administrative cohorts, the project was designed to significantly advance our knowledge of the health effects of low-level air pollution on the mortality and morbidity endpoints that commonly drive air pollution health

impact assessments and air quality standards and guidelines. The project was designed to provide novel insight into the shape of concentration–response functions at low concentrations and novel insights into the sources of heterogeneity of risk estimates between populations by investigating between- and within-area exposure contrasts on a fine spatial scale. By selecting key criteria pollutants and BC in large, pooled datasets of well-characterized cohorts, we anticipated being able to advance the understanding of the role of single pollutants in a multi-pollutant atmosphere. This knowledge may contribute to future assessments of global burden of disease and revisions of air quality guidelines and standards. The project focused on multiple key pollutants: NO<sub>2</sub>, BC, and O<sub>3</sub> in addition to PM<sub>2.5</sub>, the most often studied pollutant. We also assessed PM<sub>2.5</sub> composition, specifically the copper, iron, zinc, and sulfur content of PM<sub>2.5</sub>. All methods, results, and discussion of particle composition analyses are included in Appendix 4. The incorporation of both the ESCAPE cohorts (with detailed individual covariate information) and very large administrative cohorts (with limited individual covariate data) in one study provided new insights into the merits of both approaches.

---

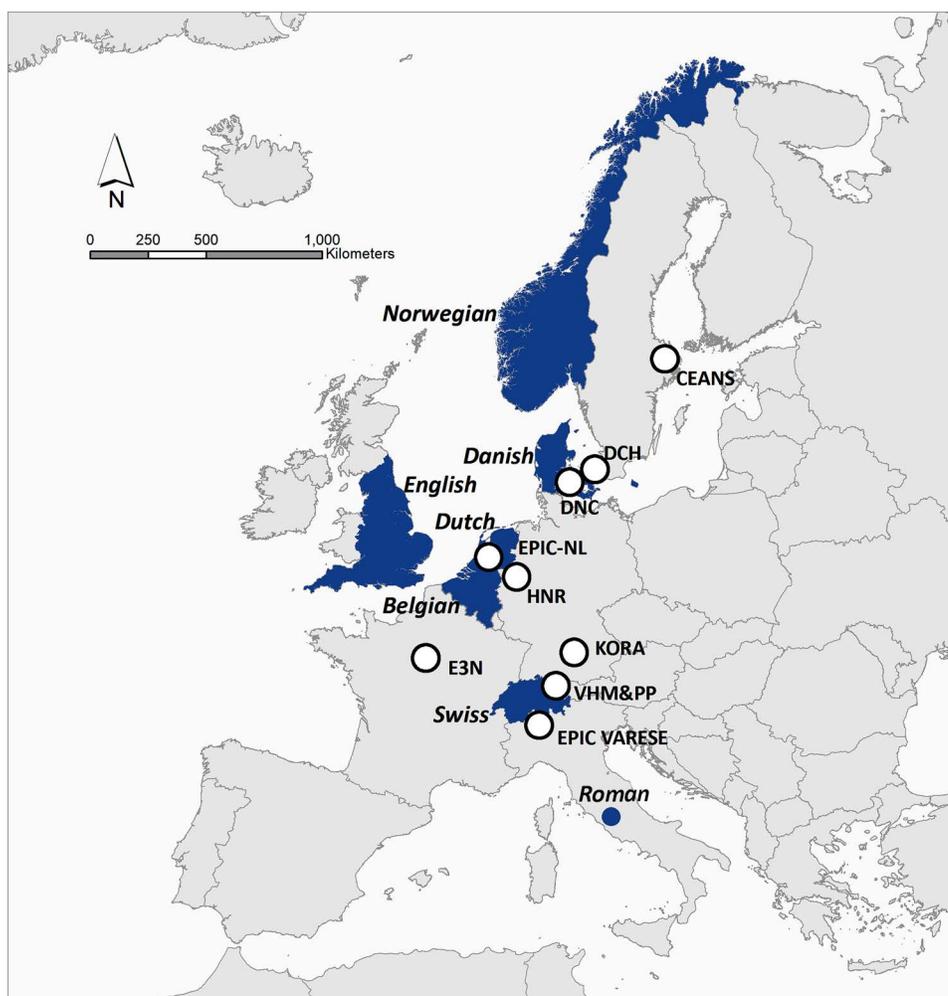
**METHODS AND STUDY DESIGN**

---

**STUDY DESIGN**

We addressed the health effects at low air pollution levels by performing new analyses within selected cohorts of the ESCAPE study and in seven very large European administrative cohorts (Figure 1, Table 1). By combining well-characterized ESCAPE cohorts and large administrative cohorts in one study the strengths and weaknesses of each approach can be adequately addressed. The large administrative cohorts have more statistical power and can efficiently control for area-level confounders, but have fewer possibilities to control for individual-level confounders. The ESCAPE cohorts have detailed information on individual confounders, as well as country-specific information on area-level confounding.

The data from the included ESCAPE cohorts and one additional non-ESCAPE cohort (the Danish Nurse Cohort) have been pooled and analyzed centrally. Data have been stored on a secure server in Utrecht and were made available for remote analyses. Data from the administrative cohorts have been analyzed locally, without transfer to a central database, because privacy regulations prevented transfer of data from administrative cohorts to a central database.



**Figure 1. Study cohorts.** Shaded regions indicate the areas included in the administrative cohorts; open circles indicate the cohorts included in the pooled cohort.

## STUDY POPULATIONS

### Pooled Cohort

From the cohorts included in ESCAPE, we selected for the current study those cohorts that carried the most weight in the ESCAPE analyses of the mortality and morbidity endpoints for the current study, that were willing to pool data, and that were recruited relatively recently. The main criteria for excluding a cohort were: not contributing to analysis of low-level air pollution (i.e., Turin), inability to participate in pooling data (i.e., EPIC Athens, Rome SIDRIA, and FINRISK), and relatively early recruitment date of the full cohort (i.e., SAPALDIA and SALIA). Included cohorts all had validated morbidity outcome data and, collectively, sufficient geographical spread to effectively make use of exposure contrasts between cohorts. We did not select

cohorts based on whether they demonstrated increased relative risks in analyses carried out in ESCAPE previously. The Danish Nurse Cohort, which was not a part of ESCAPE, was included in ELAPSE by the University of Copenhagen and provided their own funding.

Table 1 shows the selected cohorts and key characteristics. Most of the cohorts consist of a population of a region including one or multiple large cities and surrounding smaller towns (Table 1). The DCH and E3N cohorts were nationwide cohorts. Ten of the cohorts were ESCAPE cohorts. The Stockholm CEANS cohort was constructed from four individual ESCAPE cohorts that were analyzed separately within ESCAPE. For DCH, only the Copenhagen part of the cohort was included in ESCAPE, while both cities of Copenhagen and Aarhus (entire cohort) participated in ELAPSE. EPIC-NL comprised two cohorts

**Table 1.** Study Populations

Cohort	Study Area	Population Size <sup>a</sup>	Recruitment	Follow-up	Endpoints
<b>Pooled Cohorts</b>					
CEANS-SDPP	Stockholm county, Sweden	7,835	1992–1998	2011	M, C, LC, AC
CEANS-SIXTY		4,180	1997–1999	2014	M, C, LC, AC
CEANS-SALT		6,724	1998–2002	2011	M, C, LC, AC
CEANS-SNACK		3,248	2001–2004	2011	M, C, LC, AC
DCH	Cities of Copenhagen and Aarhus, Denmark	56,308	1993–1997	2015	M, C, LC, AC
DNC-1993	Denmarkwide	19,664	1993	2013	M, C, LC, AC
DNC-1999		8,769	1999	2013	M, C, LC, AC
E3N	Francewide	53,521	1993–1996	2011	M, LC
EPIC-NL-MORGEN	Four cities, the Netherlands	20,711	1993–1997	2013	M, C, LC
EPIC-NL-PROSPECT		16,194	1993–1997	2013	M, C, LC
EPIC-VARESE	City of Varese, Italy	12,028	1993–1997	2014	M, C, LC
HNR	Ruhr area, Germany	4,809	2000–2003	2015	M, C, LC
KORA-S3	Augsburg area, Germany	4,566	1994–1995	2011	M, C
KORA-S4		4,257	1999–2001	2014	M, C
VHM&PP	Vorarlberg region, Austria	170,250	1985–2005	2014	M, LC
<b>Administrative Cohorts</b>					
Belgian	National	6,491,801	2001	2011	M
Danish	National	3,409,517	2000	2015	M, LC
Dutch	National	10,532,360	2008	2012	M
English	National	1,491,124	2011	2017	M, LC
Norwegian	National	2,516,192	2001	2016	M, LC
Roman	City of Rome, Italy	1,263,712	2001	2015	M, LC
Swiss	National	4,293,521	2000	2014	M

AC = asthma and COPD incidence; C = cardiovascular incidence; LC = lung cancer incidence (pooled cohort) or lung cancer mortality (administrative cohorts); M = mortality.

<sup>a</sup> Population size is the number of subjects for which information was transferred to Utrecht University for construction of the pooled cohort. It may differ from what is shown in Appendix 2, section P1, for example, because of missing exposure data.

(EPIC-Morgen and EPIC-Prospect), which were analyzed separately in ESCAPE. In the project proposal, we also included the Oslo-based HUBRO study and the English EPIC-Oxford study but both studies were unable to transfer data because internal procedures took much longer owing to unforeseen changes in privacy regulations. At the deadline of November 2018, both cohorts were unable to provide data and were therefore excluded from the analysis.

Compared with the published epidemiological analyses of the ESCAPE cohorts, our study included the following major innovations:

- We conducted pooled analysis instead of cohort-specific analysis. Because of the successful collaboration within ESCAPE, pooling of adult cohort data was now possible. The rationale for pooling was a gain in statistical power for the epidemiological analyses and the ability to exploit between-cohort exposure contrasts as well. Especially for  $PM_{2.5}$ , between-cohort exposure contrast is substantially larger than within-cohort contrast (Eeftens et al. 2012b). The pooled data set importantly allowed more straightforward investigation of the shape of the concentration–response function using a range of nonparametric methods.
- We strengthened exposure assessment. We developed Europe-wide hybrid models, incorporating satellite data and chemical transport models in a LUR framework (de Hoogh et al. 2018). Because the focus of the current study is on pooled analyses and because of new insights in the robustness of LUR models in relation to number of monitoring sites, we developed models for combined study areas. This also allowed better use of those cohorts within ESCAPE residing in large study areas of which typically only a fraction was covered by dedicated monitoring campaigns in ESCAPE (e.g., the French E3N cohort and the Danish DCH). Residential history data are also incorporated in the exposure assessment.
- We added  $O_3$  to the evaluated components, as previous work from the United States suggested associations of  $O_3$  with respiratory mortality (Jerrett et al. 2009). Recently published work also suggested an association between long-term  $O_3$  exposure and cardiovascular mortality (Turner et al. 2016)
- We incorporated updated mortality and morbidity follow-up data (from typically up to 2008 in ESCAPE to up to 2011–2017 in the current study), which substantially increased the number of cases and hence study power.

### Administrative Cohorts

We also analyzed seven large administrative cohorts from seven countries with a total of more than 28 million participants (Table 1). A Belgian nationwide administrative cohort study group participated in the study and provided its own funding, in addition to the six cohorts mentioned in the project proposal.

The administrative cohorts contributed primarily to the mortality analyses. Some of the cohorts also included lung cancer incidence. The cohorts were based on linking different administrative databases including a census, a population registry, and death registries under strict national data protection regulations. The cohorts were much larger than the pooled cohort but lacked information on individual lifestyle, with the exception of the English cohort. In the English cohort, information on smoking and body mass index (BMI) was available from general practitioner records. The administrative cohorts were recruited more recently than the pooled cohort. More details can be found in Appendix 1, section M1.

In the previous publications of three of the administrative cohorts, exposure assessment and analytical methods differed between the cohorts, complicating comparison of the reported associations. In ELAPSE we analyzed all administrative cohorts using the same exposure assessment method and analytical methods. We further assessed  $PM_{2.5}$  at a finer spatial scale ( $100 \times 100$  m) than in earlier publications ( $1 \times 1$  km).

### EXPOSURE ASSESSMENT

We focused on  $PM_{2.5}$ ,  $NO_2$ ,  $O_3$ , and BC.  $PM_{2.5}$  composition was added at a later stage. Air pollution exposure has been assessed for the study population using central and local exposure assessments (Table 2):

1. The central exposure assessment used Europe-wide hybrid LUR models, which incorporate AirBase routine monitoring data for  $PM_{2.5}$ ,  $NO_2$ , and  $O_3$ , and ESCAPE monitoring data for BC and  $PM_{2.5}$  composition, land use, and traffic data supplemented with satellite observations and dispersion model estimates. Airbase is the database of routine monitoring data maintained by the European Environmental Agency. These models have been applied to all cohorts including the administrative cohorts. In addition, we applied the  $PM_{2.5}$  models developed within the companion HEI-funded Canadian MAPLE study (Brauer et al. 2019) for comparison.
2. Local exposure models were used that were already available for the study population concerned (Table 2). These models differed per cohort but used data

**Table 2.** Exposure Assessment for the Cohorts

Cohort	PM <sub>2.5</sub>	NO <sub>2</sub>	O <sub>3</sub>	BC	PM Composition
<b>Central Exposure Assessment</b>					
Pooled cohort	European LUR AirBase routine monitoring data	European LUR AirBase routine monitoring data	European LUR AirBase routine monitoring data	European LUR ESCAPE monitoring data	European LUR ESCAPE monitoring data
All administrative cohorts					
<b>Local Exposure Assessment</b>					
Pooled cohort	Local ESCAPE <sup>a</sup>	Local ESCAPE <sup>a</sup>	None	Local ESCAPE <sup>a</sup>	NA
Belgian	RIO model	RIO model	RIO model	RIO model	NA
Danish	AirGIS DM	AirGIS DM	AirGIS DM	AirGIS DM	NA
Dutch	Local ESCAPE <sup>a</sup> and DM	Local ESCAPE <sup>a</sup> and DM	National LUR and DM	Local ESCAPE <sup>a</sup> and DM	NA
English	NA	DM	NA	NA	NA
Norwegian	NILU DM	NILU DM	NA	NA	NA
Roman	Local ESCAPE <sup>a</sup> and DM	Local ESCAPE <sup>a</sup> and Italian LUR	NA	Local ESCAPE <sup>a</sup>	NA
Swiss	Pollumap DM	Pollumap DM	Pollumap DM	NA	NA

DM = dispersion model.

<sup>a</sup> Local ESCAPE is study area-specific LUR (Beelen et al., 2013; Eeftens et al., 2012a).

that are only available for that cohort. The main focus of exposure modeling work for this project was on Europe-wide hybrid models. In all countries where local models were available, we made a comparison between the Europe-wide model estimates and the estimates by the local models. For the pooled cohort, we evaluated the cohort-specific ESCAPE LUR models. For the administrative cohorts, local models were national dispersion models for five cohorts, an interpolation model for one cohort, and the ESCAPE model for two cohorts. Appendix 1, section M2 describes the local models in more detail.

Consistency of exposure estimation favored the use of the central exposure model as the main exposure model.

### ELAPSE Europe-wide Hybrid Land Use Regression Models

We developed Europe-wide models for annual average PM<sub>2.5</sub>, NO<sub>2</sub>, O<sub>3</sub>, BC, and particle composition (copper, iron, zinc, and sulfur in PM<sub>2.5</sub>) and warm and cold season O<sub>3</sub>. The monitoring data, GIS data, and modeling methods for ELAPSE closely followed our previous European modeling paper (de Hoogh et al. 2016). In this paper, models were developed based upon ESCAPE and AirBase routine monitoring data of the year 2010. An important finding was that models based on one monitoring database performed well when validated with the other database. We used models based on AirBase monitoring as the main

exposure variable because this modeling approach can be performed for multiple years. The models we developed have now been published (de Hoogh et al. 2018). Here we summarize the key elements of the method and its performance. We selected 2010 as the primary year of modeling because this was the earliest year of a sufficiently wide coverage of PM<sub>2.5</sub> monitoring across Europe. For BC, 2009–2010 was the period of ESCAPE monitoring used to develop BC models. For reasons of consistency, we used the 2010 period for NO<sub>2</sub> and O<sub>3</sub> for our main models as well.

Annual average concentration data for PM<sub>2.5</sub>, NO<sub>2</sub>, and O<sub>3</sub> (warm season average) for 2010 were derived from AirBase routine air pollution monitoring data for 543 sites (PM<sub>2.5</sub>), 2,399 sites (NO<sub>2</sub>), and 1,730 sites (O<sub>3</sub>) spread across Europe. Models for BC and particle composition were developed based upon ESCAPE monitoring data for 2010 (436 sites). In ESCAPE, we measured the reflectance of PM<sub>2.5</sub> filters and transformed these into PM<sub>2.5</sub> absorbance (described as units of 10<sup>-5</sup>/m). For brevity, we use the term BC for PM<sub>2.5</sub> absorbance. A measure of 1 × 10<sup>-5</sup>/m PM<sub>2.5</sub> absorbance is approximately equivalent to 1.1 µg/m<sup>3</sup> BC (Eeftens et al. 2012b).

Figure M1 in Appendix 1 shows the maps of monitoring sites. As predictor variables we used road and land use data supplemented with satellite data and dispersion model estimates. The satellite data were the satellite-derived (SAT) PM<sub>2.5</sub> product (V3.01) at a 0.1° × 0.1° (~10 km) resolution and two newly available products from the global dataset reported in van Donkelaar and colleagues (2015). The pregeographically weighted regression dataset used here includes aerosol optical density (AOD) from multiple satellite products (MISR, MODIS Dark Target, MODIS and SeaWiFS Deep Blue, and MODIS MAIAC) together with simulation-based sources, with information content below ~10 km provided by the MAIAC AOD retrieval. NO<sub>2</sub> SAT estimates for 2010 were derived from the tropospheric NO<sub>2</sub> columns measured with the OMI O3 monitoring instrument on board the Aura satellite. Satellite column-integrated retrievals for NO<sub>2</sub> were related to ground-level concentrations using the global GEOS-Chem model, producing an annual gridded NO<sub>2</sub> surface at a 10-km resolution (Bechle et al. 2015). Pollutant estimates for 2010 from two long-range chemical transport models were obtained as potential predictor variables in the models. Annual PM<sub>2.5</sub>, NO<sub>2</sub>, and O<sub>3</sub> estimates were derived from the MACC-II ENSEMBLE model at a 0.1° × 0.1° (~10 km) resolution (Inness et al. 2013). We additionally used a dataset from the Danish Eulerian Hemispheric Model (DEHM) for PM<sub>2.5</sub>, NO<sub>2</sub>, O<sub>3</sub>, and BC at a monthly (temporal) 26 × 26 km (spatial) resolution (downscaled from an original 50 × 50 km resolution using bi-linear

interpolation) (Brandt et al. 2012). Annual DEHM averages were calculated for NO<sub>2</sub>, PM<sub>2.5</sub>, BC, and O<sub>3</sub>. Road data, classified as “all” and “major” roads, were extracted from the 1:10,000 EuroStreets digital road network. Traffic intensity data were not used as the data were not available consistently across Europe. Land cover data were extracted from European Corine Land Cover 2006 data. Elevation was extracted from the SRTM Digital Elevation Database version 4.1, which has a resolution of approximately 90 m. Population data (1 km resolution) for 2011 was obtained from Eurostat.

LUR models were developed using the supervised linear regression approach used within ESCAPE (Eeftens et al. 2012a). Models were validated using five-fold cross-validation for PM<sub>2.5</sub> and NO<sub>2</sub> on ESCAPE external data. Five models were developed, each built on 80% of the monitoring sites with the remaining 20% used for validation (sites selected at random, stratified by site type and country). We explored universal kriging or, if not feasible, X and Y coordinates to further explain spatial variation in the residuals. Kriging significantly improved the PM<sub>2.5</sub> and O<sub>3</sub> models. Overall our models that included kriging explained 66%, 58%, 60%, and 51% of the variability in measured concentrations in five-fold cross-validation for PM<sub>2.5</sub>, NO<sub>2</sub>, O<sub>3</sub>, and BC, respectively. Appendix 1, Table M1, provides more detail on model structure and performance. Performance of the models in specific study areas differed but was generally somewhat less than the overall performance (de Hoogh 2018). Table M2 provides the performance of the models in subsets of low concentrations. The performance of the model measured by the R<sup>2</sup> but not the RMSE was lower for PM<sub>2.5</sub> and BC when restricted to lower levels. Lower variability in subsets of low concentrations has likely contributed to the decrease in R<sup>2</sup>. Few sites remained to evaluate the model performance for PM<sub>2.5</sub> concentrations below 10 µg/m<sup>3</sup>. For NO<sub>2</sub> the model performance was only modestly worse at low levels. For O<sub>3</sub>, the model performed poorly when restricted to concentrations below 80 µg/m<sup>3</sup>. Our main epidemiological analyses used the model developed using all sites (100%). We applied the five cross-validation models in selected epidemiological analyses as a test of robustness of the findings.

To test for spatial stability of the pollution surfaces, we also developed NO<sub>2</sub> and O<sub>3</sub> models for the years 2000 and 2005 and a PM<sub>2.5</sub> model for the year 2013 (sufficient PM<sub>2.5</sub> monitoring sites for LUR modeling in Europe were only available from around 2008, with the highest number so far recorded in 2013). These additional models (2000 and 2005 NO<sub>2</sub> and O<sub>3</sub> models and 2013 PM<sub>2.5</sub> model) were not used in the epidemiological analysis but were compared with the 2010 surfaces to evaluate the stability of the

spatial structure of the 2010 models. Figure M2 in Appendix 1 presents the pollution maps from 2010 and other years. For the ELAPSE countries overall, squared correlations ( $R^2$ ) between concentrations predicted at random sites of different years were very high (>78%). Within countries, correlations were lower but still high: >49% for PM<sub>2.5</sub>, >80% for NO<sub>2</sub>, and >47% for O<sub>3</sub> (except Italy). Correlations between AirBase measurements in different years showed the same pattern. Appendix 1, section M3, provides more detailed information.

We evaluated the performance of 14 algorithms including supervised linear regression to develop LUR models for PM<sub>2.5</sub> and NO<sub>2</sub> (Chen et al. 2019). We evaluated machine learning methods (random forest, gradient boosting, and neural network) and data regularization methods (elastic net and lasso). We observed that the performance of most algorithms was similar, with little indication of better performance of more sophisticated algorithms compared with supervised linear regression.

### Exposure Assignment

Once the models were developed, 100 × 100 m grids of predicted concentrations of the four pollutants were produced and the relevant parts transferred to the partners for linking to the cohort addresses. In this method, no cohort addresses were transferred to the pooled database or administrative cohort. Careful procedures were applied to ensure that correct linking occurred, including clarification of the correct coordinate system. Checking involved exposure assignment to a set of randomly selected coordinates provided by de Hoogh and comparison with his exposure assignment. For some cohorts without the necessary geographic information system (GIS) capabilities, the local cohort sent the Swiss Tropical and Public Health Institute (Swiss TPH) an anonymous dataset with X and Y coordinates to which Swiss TPH linked the exposures. The local cohort data manager then linked the resulting exposures back into their database.

For approximately half of the subjects of the pooled cohort and all the administrative cohorts, data on residential history were available during follow-up and partly before recruitment. We assigned exposure to all available addresses using back-extrapolated concentrations.

### Back-Extrapolation

Our main model based on 2010 monitoring represents exposure toward the end of follow-up (Table 1). We estimated pollutant concentrations for each year from recruitment to end of follow-up for PM<sub>2.5</sub>, NO<sub>2</sub>, BC, and O<sub>3</sub> using back-extrapolation to 1990. We back-extrapolated by using estimated concentrations from the DEHM (Brandt et al.

2012). DEHM models monthly average concentrations across Europe at 26 × 26 km spatial resolution back to at least 1990. The rationale to perform back-extrapolation by modeled concentrations is the consistent availability of estimates across Europe for the full study period for all pollutants. In contrast, routine monitoring was less consistent, not available for BC, and only available from about 2008 for PM<sub>2.5</sub>. We used monitoring data to compare temporal patterns of modeled and measured concentrations for countries with measurements. For application to the cohorts, we calculated population-weighted average concentrations at the NUTS-1 (Nomenclature of Territorial Units for Statistics) spatial scale, allowing different spatial trends within Europe. The NUTS classification is a hierarchical system for dividing up the economic territory of the EU and the UK for the purpose of the collection, development, and harmonization of European regional statistics.

NUTS-1 reflects major socioeconomic regions. NUTS-1 may be an entire (small) country or parts of a country (e.g., four regions in the Netherlands and 14 regions in France). We back-extrapolated concentrations, using both a difference and a ratio method with 2010 as the baseline. With the difference method, the concentration difference between a year and 2010 from the DEHM model is added to all cohort exposures for that year in the same NUTS-1 area. With the ratio method, the concentration ratio between a year and 2010 from the DEHM model is used to multiply all cohort exposure for that year in the same NUTS-1 area. In the case of higher concentrations in the past, the ratio method therefore increases the contrast in cohort exposures. Appendix 1, section M4, presents details of back-extrapolation, including a comparison between modeled and measured concentration trends (Appendix 1, Figures M8–M11). For NO<sub>2</sub> and O<sub>3</sub>, time trends generally agreed well between modeled and measured concentrations: a fairly large downward trend for NO<sub>2</sub> and a small downward trend for O<sub>3</sub>. For PM<sub>2.5</sub>, the trends of measured and modeled concentrations agreed well in some countries (the Netherlands and Switzerland) but not in other countries (France and Belgium). The number of sites (two in France and one in Belgium) is too small to draw meaningful conclusions. Given the regional nature of PM<sub>2.5</sub>, very different trends across neighboring countries (e.g., Netherlands and Belgium) are unlikely. To further judge measured time trends in measured PM<sub>2.5</sub>, Figure M12 shows time trends of measured PM<sub>10</sub> concentrations across Europe. PM<sub>10</sub> has been measured at a fairly large number of monitoring sites since about 2000, whereas PM<sub>2.5</sub> monitoring started about 10 years later on a large scale (Figure M6). Downward trends in PM<sub>10</sub> concentrations have been observed across Europe and in most but not all countries (i.e., France and the first years in Sweden). The largest decrease in concentrations occurred between 1990 and 2000, with more

modest decreases between 2000 and 2010.  $PM_{10}$  trends do not necessarily fully reflect  $PM_{2.5}$  trends, as the coarse particle fraction may also exhibit separate time trends. However,  $PM_{10}$  probably provides a useful approximation of the trends in  $PM_{2.5}$ , as air pollution policies have primarily addressed PM precursors ( $SO_2$  and  $NO_2$ ) and primary fine particles. Comparison of the  $PM_{2.5}/PM_{10}$  ratio in two monitoring studies using identical equipment and analysis conducted in the Netherlands in 2000 and 2009 showed ratios of 0.67 in both years (Eeftens et al. 2012b).

### MAPLE $PM_{2.5}$ Exposure Model

The MAPLE method assesses  $PM_{2.5}$  at a  $1 \times 1$  km spatial resolution primarily based upon satellite AOD (Brauer et al. 2019). The method uses a global chemical transport model (GEOS-Chem) to calculate surface concentrations. A geographically weighted regression approach is used to calibrate the surface concentrations to ground monitoring data, incorporating data on land use (van Donkelaar et al. 2015, 2016). The MAPLE project has further refined the

method by long-term AOD measurements from the ground at selected Canadian stations in the framework of the HEI-funded low-level project (Brauer et al. 2019). We used the refined  $PM_{2.5}$  estimates at  $1 \times 1$  km resolution. The ELAPSE model also includes satellite  $PM_{2.5}$  data but does not include the refinements added by geographically weighted regression and the ground-based AOD monitoring, which have been shown to improve  $PM_{2.5}$  exposure estimation significantly in the Canadian setting.

### Health Outcome Definition

Identification of outcomes includes linkage with registries for mortality, cancer incidence, hospital discharge, and physician-based adjudication of cases. We analyzed natural-cause, cardiovascular, diabetes, cardiometabolic, and respiratory mortality (Table 3). We also analyzed lung cancer incidence, incidence of coronary and cerebrovascular events, and incidence of asthma and COPD. Follow-up for mortality, cancer, and cardiovascular and cerebrovascular

**Table 3.** Outcome Definitions

Cause	ICD-9 Codes	ICD-10 Codes	Reference
<b>Mortality</b>			
Natural cause	001–779	A00–R99	Beelen et al. 2014a
All cardiovascular	400–440	I10–I70	Beelen et al. 2014b
Ischemic heart disease	410–414	I20–I25	Beelen et al. 2014b
Cerebrovascular disease	430–438	I60–I69	Beelen et al. 2014b
Respiratory	460–519	J00–J99	Dimakopoulou et al. 2014
COPD (without asthma)	490–492, 494, 496	J40–J44, J47	
Diabetes	249, 250	E10–E14	
Cardiometabolic	400–440, 249, 250	I10–I70, E10–E14	
<b>Incidence</b>			
Lung cancer	162.2–162.9	C34.0–C34.9	Raaschou-Nielsen et al. 2013
Acute coronary heart disease	410, 411, 427.5 (hospital) and 410–414, 427.5 (mortality)	I20.0, I21, I22, I23, I46 (hospital) and I20–I25, I46 (mortality)	Cesaroni et al. 2014; Luepker et al. 2003
Cerebrovascular events	431, 433.x1, 434, 436 (hospital) and 431–436 (mortality)	I61, I63, I64 (hospital) and I61–I64 (mortality)	Stafoggia et al. 2014
COPD, asthma events	490–492, 494–496 (COPD), 493 (asthma)	J40–44 (COPD), J45–46 (asthma)	Andersen et al. 2011; Fisher et al. 2016

COPD = chronic obstructive pulmonary disease; ICD = International Classification of Diseases.

events was extended until 2013 or later, which included an additional 5 years compared with the original average 13 years of follow-up in ESCAPE. It is important that we had one harmonized definition (based on codes from the International Classification of Diseases [ICD]) for specific diseases and outcomes for both administrative and ESCAPE cohorts.

Mortality outcomes were defined on the basis of the underlying cause of death recorded on the death certificates. Cause of death was coded with ICD-9 or ICD-10. As air pollution is related to more and more different diseases in recent studies, we decided to add diabetes and cardio-metabolic to the list of causes of death.

In the lung cancer incidence analyses, we included only primary cancers. Prevalent cancer cases were excluded with the exception of nonmelanoma skin cancer. We further evaluated histological subtypes as in the ESCAPE paper (Raaschou-Nielsen et al. 2013). In the administrative cohorts, we analyzed lung cancer mortality, as few cohorts had access to lung cancer incidence data. Because of the high fatality rate of lung cancer, we assumed we could compare the lung cancer findings in the pooled and administrative cohorts.

Coronary and cerebrovascular events ascertainment is based upon hospital discharge records and death certificates or physician-based adjudication of cases (Cesaroni et al. 2014; Stafoggia et al. 2014). To identify incident cases, we exclude subjects who had an acute coronary event or cerebrovascular event in the last 3 years preceding the enrollment. In order to harmonize the ICD-9 (ICD-8 in Denmark until 1993) and ICD-10 codes, the definition indicated in Table 3 deviates slightly from the ESCAPE definition (i.e., I24 is left out).

Asthma and COPD ascertainment is based upon hospital discharge records or physician-based adjudication of cases (Andersen et al. 2011; Fisher et al. 2016). The best method available within a cohort is used. To identify incident cases, we exclude subjects who had an acute respiratory event in the last three years preceding the enrollment.

All cohorts contributed data on mortality. In the pooled cohort, all cohorts contributed to lung cancer incidence except KORA (Table 1). In the pooled cohort, all cohorts contributed to coronary or cerebrovascular events except VHM&PP and E3N. The two Danish and the Swedish cohorts contributed to the asthma and COPD incidence analyses. Three of the seven administrative cohorts (English, Norwegian, and Danish) had data on lung cancer incidence. The Roman cohort used hospital admission data to assess lung cancer incidence. Therefore, we decided to add lung cancer mortality analyses instead of lung cancer incidence for the administrative cohorts. Administrative cohorts did not contribute to coronary and cerebrovascular events analyses.

---

## STATISTICAL METHODS AND DATA ANALYSIS

---

For the analysis of the pooled cohort and the individual administrative cohorts, we applied the Cox proportional hazard model with increasing control for individual- and area-level covariates to analyze the associations between air pollution and mortality or morbidity. Age was used as the timescale because of evidence that this results in better adjustment for potential confounding by age (Thiebaut and Benichou 2004). 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. A priori we specified three confounder models, following the modeling methods of the ESCAPE study (Beelen et al. 2014a; Raaschou-Nielsen et al. 2013; Stafoggia et al. 2014). Model 1 included only age (time axis), sex (as strata), and calendar year(s) of enrollment. Model 2 added individual-level variables that were consistently available in the cohorts contributing to the pooled cohort or all variables available in the administrative cohorts. Model 3 added to the model 2 area-level SES variables. Appendix 1, section M5, provides the manual we prepared to harmonize the definition of area-level variables both with respect to spatial scale and domain (income and education). We aimed at including variables from multiple SES domains (income, occupation, and education) at both a neighborhood scale and regional scale. A priori model 3 was selected as the main model. Only subjects with complete information for model 3 variables were included in the analyses.

The main exposure variable we analyzed was derived from the Europe-wide hybrid models based upon 2010 monitoring data, assigned to the address at recruitment. Sensitivity analyses were conducted using earlier time periods, time-varying exposure analyses incorporating residential history, local exposure models, PM<sub>2.5</sub> models from the Canadian MAPLE project, and O<sub>3</sub> at a larger spatial scale.

We first specified linear single air pollutant models. Two-pollutant models were specified for all combinations of the four main pollutants. Two-pollutant models for particle composition were analyzed with PM<sub>2.5</sub> and NO<sub>2</sub> as the second pollutant to limit the number of analyses. We then investigated the shape of the concentration–response function using natural splines with two, three, and four degrees of freedom; penalized splines and SCHIF functions using confounder model 3; and subset and threshold models.

We used natural splines with two, three, and four degrees of freedom, following earlier analysis in the CanCHEC cohort (Crouse et al. 2012, 2015), as a flexible method allowing multiple shapes in different parts of the exposure distribution. Natural splines are cubic polynomial-like

functions in intervals of the variable distribution as these are defined by the knots (used in formulating basis functions) and impose a smoothness criterion to satisfy certain differentiability properties. The degrees of freedom of a natural spline correspond to the number of knots plus 1 (as implemented in the context of our models). The choice of this number is critical as increasing the number of knots may overfit the data and increase the variance, while decreasing the number of knots may result in a rigid and restrictive function that has more bias (Hastie 1992; Perperoglou et al. 2019). We used the AIC and the Bayesian information criterion (BIC) to compare the goodness of fit of the models with varying of degrees of freedom for the definition of the natural splines and the linear model. The R library “splines” and the function `ns` were used to fit the splines. We present the spline plots with harmonized exposure and HR axes. The exposure axis starts at zero and is truncated at the high end of the distribution, where there is sparse data. Given the goal of studying low-level pollution, we did not truncate at the low end of the distribution. HRs are presented in comparison with the lowest exposure level. We truncated the HR axis such that the full curves are shown, and the plots are still informative. We allowed different HR axes for the pooled and administrative cohorts and for the different outcomes.

We further applied penalized splines as these may provide more flexibility compared with natural splines because knots, and correspondingly basis functions, are not predetermined but they rather use a very large number of base functions that is controlled by a roughness penalty (Perperoglou et al. 2019). We applied them as defined in the survival R statistical package using the AIC (default option in R) for optimal smoothing choice.

We furthermore applied the SCHIF function, which was used in the Canadian MAPLE study (Burnett et al. 2018; Nasari et al. 2016). We used version 2.10 (February 24, 2016) of the SCHIF code provided by Burnett, adapted to the Cox model 3 applied in our analysis. That is, for the pooled cohorts analysis we used a Cox model stratified by sex and individual cohort controlling for the selected covariates and for the administrative cohorts analysis we used a Cox model stratified by sex controlling for available covariates and for the clustering at neighborhood level. In contrast to the splines, the SCHIF methodology specifies a range of functions with a plausible shape, including sub-linear, linear, and supralinear functions (Nasari et al. 2016). Based upon statistical performance an “optimal model” is selected. An ensemble model combines the different sigmoidal models, weighted by the likelihood of the model (Nasari et al. 2016). A detailed description of the SCHIF method was included in the MAPLE study report (Brauer et al. 2019).

Additionally, we specified linear models in subsets of the concentration range, defined by removing concentrations above a certain value from the analysis, such as for  $\text{PM}_{2.5}$  below 25  $\mu\text{g}/\text{m}^3$  (EU limit value); 20, 15, and 12  $\mu\text{g}/\text{m}^3$  (U.S. EPA NAAQS); and 10  $\mu\text{g}/\text{m}^3$  (WHO Air Quality Guideline value). We specifically evaluated the associations:

- $\text{PM}_{2.5}$ : below 25, 20, 15, 12, and 10  $\mu\text{g}/\text{m}^3$
- $\text{NO}_2$ : 40, 30, and 20  $\mu\text{g}/\text{m}^3$
- $\text{O}_3$ : 120, 100, 80, 60, and 40  $\mu\text{g}/\text{m}^3$
- BC: 3.0, 2.5, 2.0, 1.5, and  $1.0 \times 10^{-5}/\text{m}$

Finally, threshold models were evaluated to investigate whether the associations persisted below specific concentration values. Specifically, we evaluated threshold values for each pollutant for pollutant-specific increments in the low range of the observed concentrations to investigate the “no association” hypothesis in these ranges. For each threshold value  $h$ , we added in the model a term  $x_+ = (\text{pollutant concentration} - \text{threshold value } h)_+$  where  $x_+ = x$  if  $x \geq 0$  and 0 otherwise. Performance of threshold models was evaluated by the comparison of the AIC and the BIC with the corresponding linear model. These models were applied in addition to the subsets analyses as it is possible that an association below, for example, 12  $\mu\text{g}/\text{m}^3$   $\text{PM}_{2.5}$  is found, but with a threshold at 7.5  $\mu\text{g}/\text{m}^3$ . We evaluated the following cut-offs for the threshold analysis:

- $\text{PM}_{2.5}$ : 5, 7.5, and 10  $\mu\text{g}/\text{m}^3$
- $\text{NO}_2$ : 10, 15, and 20  $\mu\text{g}/\text{m}^3$
- $\text{O}_3$ : 40, 60, 80, and 100  $\mu\text{g}/\text{m}^3$
- BC: 0.5, 1.0, and  $1.5 \times 10^{-5}/\text{m}$

The main difference in the methods is that subset and threshold analyses provide easy ways to interpret estimates (applicable also in health impact assessment) whereas spline approaches are more flexible and hence may be more appropriate for capturing departure from linearity in the low end of the concentration–response curves. Spline models also fit the subsets of the data. We compared population characteristics of the full population and the subsets with low exposure levels to assess potential selection effects.

The Cox proportional hazards model assumption was evaluated by assessing “log-log” plots, specifically plot  $-\ln[-\ln(\text{survival})]$  curves for each category of a nominal or ordinal covariate versus  $\ln(\text{analysis time})$ . We tested for proportional hazards assumption by test of nonzero slope in a generalized linear regression of the scaled Schoenfeld residuals on time. The null hypothesis in this test is a zero slope, which is the equivalent of testing that the log HR is constant over time.

To harmonize analyses, our working groups were set up to prepare a common codebook to define all variables based on the ESCAPE codebook, to define area-level socioeconomic status variables, and to develop statistical procedures. Common R scripts were developed by the University of Athens partner and were used by all analysts of the pooled and administrative cohorts. We used one version of R software throughout the project (version 3.4.0) and the R packages used in the analyses were “frozen” using the Packrat package to make sure these remained unchanged.

### **MODEL SPECIFICATION FOR POOLED COHORT ANALYSIS**

We first defined a maximum covariate model based upon the ESCAPE study (Beelen et al. 2014a; Cesaroni et al. 2014; Raaschou-Nielsen et al. 2013; Stafoggia et al. 2014) and then evaluated how many cohorts had information on these confounders. The final model was defined balancing the need to adjust for a specific confounder (determined by sensitivity analyses presented in Appendix 3) and the wish to include a large number of cohorts. We identified a list of individual-level confounder variables that each cohort should transfer to the central database. Exact definitions of variables were prepared, starting from the ESCAPE codebook, by the variable coding working group. Within ESCAPE, much effort had already gone into preparing a common codebook defining health outcome and potential confounding variables. Data transfer agreements had been set up with all partners. No individual identifiers were transferred. After data had been transferred, detailed checks were made by the Institute for Risk Assessment Sciences of the consistency of the transferred data. Potential problems were checked with the partners. Final ELAPSE pooled datasets were stored at the YODA, a secure server of Utrecht University. Data were accessible for statistical analyses by the partners. We used an R-Studio Server Pro environment running on a physical server of Utrecht University (16-core CPU, 192 GB RAM). All the analyses and output generation were done in this environment. Using a remote access environment prevented database distribution and ensured that the latest database version was used for analyses.

All cohorts had information on age, sex, smoking status, BMI, and individual- and area-level SES. Almost all cohorts had information available on amount and duration of smoking (E3N had smoking intensity only in classes) (Appendix 1, Tables M9 and M10). For the pooled analysis, problems included nonavailable variables in some cohorts and different definitions of the same variable. The latter applied to area-level SES and diet variables (i.e.,

alcohol and fruit intake), which have been defined with actual quantitative consumption (e.g., in grams/day) in some cohorts and with frequency of consumption (e.g., daily, weekly, and seldom) in other cohorts. Appendix 1, section M6, documents the transformations we applied to harmonize variables, mostly by classifying variables into low, medium, and high, starting from different variables. We performed detailed sensitivity analyses to determine the final confounder models 2 and 3, balancing the need to adjust for a comprehensive set of confounders and the availability in a larger number of cohorts. The final model 2 for mortality and lung cancer incidence included the following: age (time axis), sex (strata), calendar year(s) of enrollment, smoking status (never, former, or current), smoking intensity for current smokers (number of cigarettes/day), smoking intensity squared, smoking duration for current smokers (years of smoking), BMI (kg/m<sup>2</sup>) (BMI categories defined by WHO: <18.5, 18.5–24.9, 25–29.9, and >30), employment status (yes or no), and marital status (single, married, divorced, or widowed). The final model for mortality and lung cancer incidence does not include education and smoking intensity for former smokers because the large VHM&PP cohort does not have this information. Sensitivity analyses showed that these confounders did not change effect estimates in the cohorts that did have information on these confounders. For the coronary and cerebrovascular events and asthma and COPD analyses, education and smoking for former smokers were included in the main model 2, as the VHM&PP cohort did not contribute to these analyses. Additionally, the final model 3 included neighborhood-level income in 2001 as the area-level SES variable was available in all cohorts except the EPIC-VARESE cohort. The moderately small EPIC-VARESE cohort, which experiences relatively high air pollution concentrations, was thus excluded from the main analyses. We use the term “neighborhood income,” but the actual spatial scale differed across cohorts from neighborhoods of cities in study areas with major cities to municipality. Also, in the cohorts with data on the municipality scale, a large number of distinct values for the neighborhood SES variable was included (also in the Stockholm cohorts). The large Vorarlberg cohort consisted of a large number of smaller towns, with the largest town (Bregenz) having about 30,000 inhabitants. In this setting municipality is an adequate scale and similar to neighborhoods of major cities. We used the SES data for 2001 (early in the follow-up of most cohorts) because the requested data for 1991 (just prior to baseline) were very incomplete. In the main model, we added neighborhood income as a covariate in the standard Cox model. The years 1991, 2001, and 2011 were census years in several years. Appendix 2, Tables P14 and P34–P36 in the Results section show the results of the sensitivity analyses

documenting robust results after including smoking intensity and duration for former smokers, educational level (low, medium, or high), occupational class (white or blue collar classification), dietary variables, and further area-level SES variables (unemployment and education). with missing values in covariates that were chosen for models 2 and 3 were excluded from all analyses before fitting the models to ensure comparability between the model results. In a sensitivity analysis, we used multiple imputation to fill in missing values provided that a cohort had information for a variable for part of the cohort (Appendix 1, section M7). We did not estimate completely missing variables in cohorts based on information from other cohorts because we did not want to assume that relationships between variables were identical in cohorts from different countries, age, and sex distributions.

In our main model we used strata for the original (subcohort contributing to the pooled cohort to account for differences in baseline hazard between the cohorts. We decided this would be our main approach given the differences in population, analyses showing that the assumption of proportional hazards did not hold with respect to cohort, and a substantially better model performance compared with alternative specifications. We evaluated the sensitivity to alternative specifications including no adjustment for cohort, indicators for cohort, frailty models, a mixed effect for cohort, and correction of the standard error for clustering within cohort. The choice of the modeling approach affects the exposure contrast that is exploited in the analysis. In the case of strata and fixed cohort effects, the analysis is largely based upon exposure contrast within cohorts, similar to earlier analyses within ESCAPE. With random cohort and especially no cohort adjustment, we may also exploit between cohort contrasts, which especially for  $PM_{2.5}$  was more substantial. Based upon the small number of subjects and especially events per neighborhood, we did not include adjustment for neighborhood in our main analyses. We evaluated the sensitivity to correction of the standard error for clustering within cohort. Appendix 1, section M8, provides more detail about these approaches. The more detailed documentation and evaluation of these methods including a simulation study has been published separately (Samoli et al. 2021). Figure P22 in Appendix 2 documents the robustness of the approaches including adjustment for cohort for mortality.

In our main model we did not adjust the standard errors of our effect estimates for clustering inherent in our data due to the subjects' residence in the same neighborhood. Due to a small number of subjects and especially cases per neighborhood, correction was unstable. In the much larger administrative cohorts, we did apply this correction.

We performed additional analyses to assess the robustness of the findings in our main model:

1. We used time-varying exposure analyses and the concentration back-extrapolated to the year of recruitment as alternative exposure variables. Residential history was incorporated in the time-varying exposure analyses, such that both changes in air pollution spatial patterns and moving residential address were accounted for. The baseline exposure was assessed at the residential address at baseline, as was the 2010 exposure. Back-extrapolation was performed with both the difference and ratio method (see the paragraph describing back-extrapolation in the Exposure Assessment section below). The procedure scales back the 2010 exposure surface, with time trends represented at a large spatial scale, typically one or a few distinct values per cohort. The trend correction was allowed to differ across cohorts. In the time-varying analyses we assessed exposure at the address applicable to each year using the same back-extrapolation procedure. This analysis was only possible for cohorts with reliable residential history data (about 60% of the total pooled cohort). To allow comparison, we also specified the main exposure model for the subjects included in the time-varying analysis. For the analyses we created datasets for every year of follow-up. The same models were specified in the stacked dataset as in the static main exposure analysis. Both for subjects with and without an event, the annual average exposure of each year was used as the exposure variable. Because of concerns with bias due to time trends in air pollution and mortality in the time-varying analysis, we additionally specified strata for calendar time periods. We note that in our analysis we used age as the underlying time axis. We used 1-year and 5-year periods to define the strata. We used strata because we did not want to assume a linear time trend. The 1-year strata provide the most stringent control for secular time trend. We specified the 5-year strata period because of concerns with potential over-specifying time trends.
2. In addition to the linear time-varying exposure analyses, we also specified spline functions to test the sensitivity of our findings with respect to the level at which associations are observed. We used natural splines with three degrees of freedom, the ratio back-extrapolation, and 1-year strata to adjust for time trends. We used the time-varying exposure analysis for this sensitivity analysis because of the hypothesis that the time lag between exposure and (cardiovascular) mortality is relatively short.

3. We performed an additional analysis starting the follow-up in 2000, 2005, and 2008 such that the main exposure variables more closely represented exposure during most of the follow-up time period. In this analysis only those subjects who survived from recruitment until the new start of follow-up could be included. An analysis starting in 2010 was not interpretable because follow-up ended in 2011 for several cohorts.
4. We tested the sensitivity of our findings from the main model 3 by applying multiple imputation methods to address missing data in confounder variables for natural-cause mortality (Appendix 1, section M7). We additionally compared HRs for the minimally adjusted model 1 in the full population and the smaller complete case population.
5. We tested the sensitivity of our findings from the main model 3 by adding additional confounder variables: education, smoking intensity and duration for former smokers, alcohol, diet, and additional SES indicators.
6. For all endpoints we evaluated effect modification by age at baseline (<65 vs. ≥65 years), smoking status (current, ever, or never) and BMI (cut points at 18.5, 24.9, and 29.9 kg/m<sup>2</sup>). For acute coronary heart disease and stroke and asthma and COPD effect modification for education (low, medium, or high) was also assessed. Effect modification was evaluated by introducing interaction terms into the model and tested by the Wald test.
7. We compared effect estimates obtained with the Europeanwide LUR model and the published ESCAPE local models. ESCAPE models were study-area specific but based on fewer sites (20–40 for PM<sub>2.5</sub> and 40–80 for NO<sub>2</sub>)
8. In the analysis of natural-cause, cardiovascular, cerebrovascular, ischemic heart disease, diabetes, and cardiometabolic mortality and of the incidence of coronary and cerebrovascular events, additional adjustment for road-traffic noise was conducted. For six cohorts traffic noise at the residential address was available. For CEANS, DCH, and DNC, the Nordic Prediction method was used. For HNR and KORA, the German VDUS/RLS-90 method was used. More detail on noise exposure assessment has been provided previously (Stafoggia et al. 2014). Reporting noise effect estimates is beyond the scope of ELAPSE and the topic of other projects.

We evaluated the impact of measurement error on the magnitude of the effect estimates by a comparison of effect estimates derived from predictions originating from the five-fold cross-validation of exposure estimation. All models were validated using hold-out validation using

random subsets (20%) of the monitoring datasets, stratified by site type (background and traffic) and region of Europe. Hence, in the exposure model validation approach we developed five exposure models based on stratified random samples of 80% of the monitoring sites and compared the robustness of exposure model predictions and health effect estimates between these five models. This provided information about uncertainty with respect to model identification, which has been identified as an important source of uncertainty in recent LUR validation work (Basagaña et al. 2012; Wang et al. 2012). We further applied regression calibration (Carroll et al. 1995; Keogh and White 2013). We estimated the attenuation factor from a simple linear regression of the measured concentrations at the ESCAPE monitoring sites versus the estimated concentrations from the ELAPSE model for PM<sub>2.5</sub> and NO<sub>2</sub>. ESCAPE monitoring data were not used in model development for PM<sub>2.5</sub> and NO<sub>2</sub>. For BC and O<sub>3</sub> we needed another procedure as BC models were developed based upon ESCAPE data and O<sub>3</sub> was not measured in ESCAPE. Therefore, for BC and O<sub>3</sub> we regressed ELAPSE model predictions from the hold-out validation procedure at ESCAPE and Air-Base sites, respectively, on measured concentrations. Appendix 1, section M9, provides further details.

#### **MODEL SPECIFICATION FOR ANALYSIS OF ADMINISTRATIVE COHORTS**

All analyses were performed per individual cohort because privacy regulations prevented data transfer to a central database. We selected all available individual-level covariates and all area-level variables available at the neighborhood and regional scale in our main model. As all administrative cohorts were analyzed individually, we preferred maximal adjustment per cohort rather than a common model with few covariates. We specified three confounder models as for the pooled cohort.

Model 1 included only age (time axis), sex (strata), and calendar time (year[s] of enrollment). Model 2 included all individual-level variables available within each cohort. Availability of these covariates differs by administrative cohort. Briefly, the English administrative cohort has information on individual lifestyle covariates from primary-care records (e.g., smoking and BMI), but no information on individual SES. The other administrative cohorts have individual data on demographic variables — such as education (Swiss, Roman, and Norwegian), household income (Dutch, Danish, and Norwegian) or employment status (Roman, Danish, and Norwegian) — that characterize SES. The main model was a model with the maximum number of variables available in the cohort. Model 3 added area-level SES variables at the regional and neighborhood

spatial scale. More area-level variables were included in model 3 than in the pooled cohort, as cohorts were analyzed individually. In the main model, we added area-level variables as covariates in the standard Cox model, with no random area effect. In all models we did adjust the standard errors of the effect estimates for clustering inherent in the data due to the subjects' residence in the same neighborhood. The variance correction applies a robust or sandwich-type variance estimator to account for the clustering of subjects (Austin 2017; Therneau and Grambsch 2000).

Variables differed by cohort due to availability of data and privacy regulations (Appendix 1, Table M21). The English cohort had more detailed individual information, but only assigned a deprivation score at the neighborhood-level scale. For bureaucratic reasons the assignment of more detailed SES indicators such as income, education, and unemployment separately and the accounting for regional-scale SES — as requested in the manual and followed by the other administrative cohorts — were not possible. To account for regional differences, we therefore included regional area indicators in the main model for the English cohort, whereas it was decided as sensitivity analysis in the other cohorts (which did have multiple neighborhood- and regional-scale SES variables in their models). Multiple years were available for most cohorts. We selected the survey closest to the baseline for the indirect adjustment.

To address the lack of information of lifestyle factors, we explored two approaches: (1) using information on area-level (neighborhood and region) morbidity or mortality from lung cancer, COPD, diabetes, and (2) using indirect adjustment using available survey data. We used age-standardized area-level rates of diseases related to smoking and obesity (lung cancer, COPD, and diabetes). The areas considered were neighborhood and a larger scale, depending on data availability.

We further applied the indirect adjustment method proposed by Shin and colleagues (2014). The method involves the use of ancillary information from surveys that are highly representative of the subjects in the cohort to adjust the estimated HRs for missing confounders, such as smoking and BMI. The method uses the relationship between air pollution exposure and lifestyle in an external (survey) population and applies that relationship to the cohort that is missing lifestyle information. The Shin method further needs a risk function for the missing confounder variable. We applied the Shin method for natural-cause mortality and the key potential confounders, smoking status and BMI. We focused on smoking and BMI because these were considered critical confounders in the WHO systematic review of outdoor air pollution (Chen

and Hoek 2020). Risk estimates for these confounders were obtained from the analysis of the pooled cohort. Table M22 in Appendix 1 provides more detail on the surveys used for the indirect adjustment approach. In the data from large population surveys from the Netherlands, Switzerland, Norway, Belgium, and Denmark, variables such as smoking, alcohol use, exercise, and BMI were available (Appendix 1, Table M22). We applied the method in four steps. First, we assessed the comparability of the survey and cohort. Second, we assessed the quantitative relationship between air pollution and smoking status and BMI. Third, we obtained the HR of smoking and BMI by analyzing the large pooled cohort. Fourth, we used the above information in the adjustment formulas proposed by Shin and colleagues.

As Shin and colleagues (2014) emphasized the need for “assessing the adequacy of the ancillary health studies in representing the cohort study,” in the first step we compared the distributions of air pollution, age, sex, and other variables available in the administrative cohorts between the administrative cohorts and the available surveys. Second, the correlations amongst the variables available in the administrative cohort (i.e., education and income) were examined and compared with the correlations amongst the same variables in the health survey. Concerns should be raised about using the ancillary data if these distributions or correlations are not similar.

The second step was to assess whether BMI and smoking were associated with the air pollution estimates using a multiple linear regression model controlling for the other covariates in model 3. This step examined whether these variables could act as potential confounders in the final epidemiological models. The quantitative results were applied in the adjustment procedure.

Subsequently we adjusted the estimates from the main Cox model by using the estimates of the association between the missing risk factors (BMI and smoking) and the health outcome. Shin and colleagues (2014) proposed that these estimates be obtained from the literature. In ELAPSE we estimated the association between BMI and total mortality from the pooled cohorts' data. The indirect adjustment method incorporates in the formula the association between observed covariates (all variables in the model including air pollution) in the administrative cohort with the missing covariates (BMI and smoking) that are acquired from the ancillary study data. Specifically, the indirectly adjusted parameter is given by  $\hat{\beta} = \hat{\gamma} - \bar{\Delta}\lambda$ , where  $\hat{\gamma}$  the vector of the unadjusted estimates for the air pollution effect,  $\bar{\Delta}$  is a matrix of estimates for the associations between the air pollution exposure and lifestyle from the ancillary dataset and the observed

variables,  $\tilde{\lambda}$  is a vector of the regression parameter estimates of the BMI and smoking risk factors on the health outcome obtained from the pooled cohorts analysis. The variance of the adjusted estimates also incorporates the variance of the estimates of the observed variables, the variance-covariance matrix of the estimates of the missing factors obtained from the pooled cohorts' analysis, and the variance-covariance matrix of the estimates for the associations between observed variables and missing factors based on ancillary survey data.

We performed the following additional analyses to assess the robustness of the findings in our main model:

1. We additionally adjusted for indicator variables for region of the country to allow for variation in health outcomes not accounted for by the confounder variables in the models. We did not include area indicators in our main models as we preferred to include area-level SES variables in the models that have a more logical link with health than simple indicator variables. In Appendix 1, Table M23 lists the area definitions in our cohorts.
2. We used time-varying exposure analyses and the concentration back-extrapolated to the year of recruitment as alternative exposure variables. Residential history was incorporated in these analyses, such that both changes in air pollution spatial patterns and moving residential address were accounted for. A more detailed description of these analyses is provided in the pooled cohort analysis section.
3. For all endpoints we evaluated effect modification by age at baseline (<65 vs. ≥65 years), sex, education (low, medium, or high), smoking status (current, ever, or never), and BMI (cut-points at 18.5, 24.9, and 29.9 kg/m<sup>2</sup>), when this information was available. Effect modification was evaluated by introducing interaction terms into the model and tested by the Wald test.
4. In the analysis of cardiovascular, cerebrovascular, and ischemic heart disease mortality, additional adjustment for road-traffic noise was conducted. For three cohorts, traffic noise at the residential address was available (Dutch, Roman, and Swiss cohorts). National prediction methods were used. Reporting noise effect estimates is beyond the scope of ELAPSE and the topic of other projects.
5. We compared effect estimates obtained with the Europe-wide LUR model and local models.

#### **META-ANALYSIS OF EFFECT ESTIMATES FROM ADMINISTRATIVE COHORTS**

Effect estimates of the seven cohorts from linear (single- and two-pollutant) and subset linear models were combined using meta-analysis techniques. Specifically, we

applied generic variance meta-analytic methods to provide random effects pooled estimates. Random effects meta-analysis was used as we expected heterogeneity between the estimates attributable to a variety of sources, such as country-specific characteristics or different confounding control among different cohorts. Heterogeneity was assessed using the Cochran Chi statistic and the I<sup>2</sup> statistic. We addressed the amount of heterogeneity by estimating a range of estimates for the between-cohorts variance to test the robustness of the findings. Specifically, we applied the Der Simonian and Laird, empirical Bayes, and restricted maximum likelihood (REML) methods with and without the Knapp and Hartung correction (Knapp and Hartung 2003). REML was our default model. We used the *'metagen'* function in R.

#### **COMBINATION OF POOLED COHORT AND ADMINISTRATIVE COHORTS' EFFECT ESTIMATES**

We evaluated the agreement between the estimates from the pooled cohort (with detailed individual confounder control) and those of the administrative cohorts (with larger numbers of observations but less individual-level confounder control). We did not perform a meta-analysis of the (linear) effect estimates from the pooled cohort and the administrative cohorts, but rather evaluated consistency across the two sets of cohorts.

---

## **RESULTS**

---

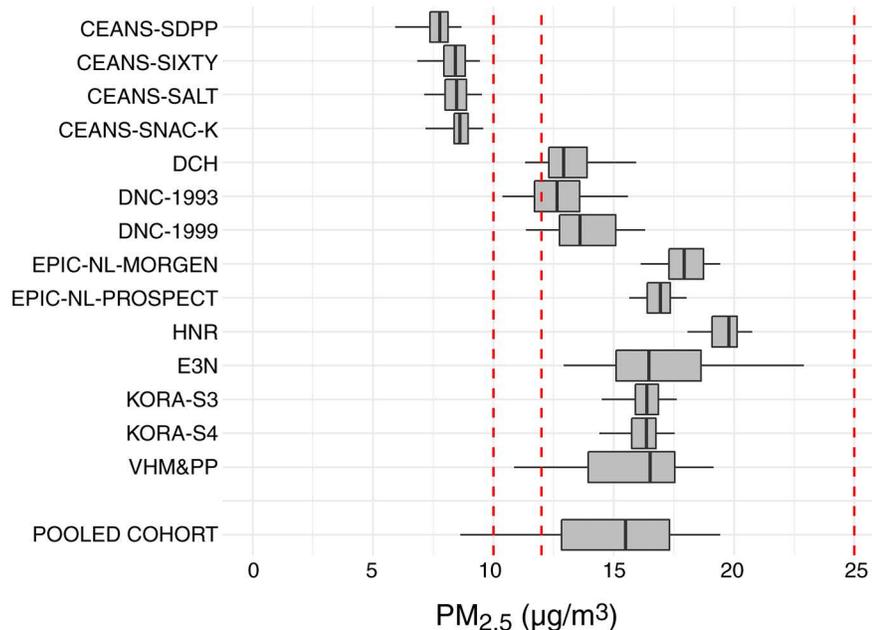
We first present associations in the pooled cohort and then in the administrative cohorts. In the pooled cohort we subsequently discuss the mortality findings and then the findings for incidence of cardiovascular disease, lung cancer, asthma, and COPD. In all sections, we distinguish descriptive analyses of exposure and population, main analyses, and additional analyses of the association between air pollution and health outcomes. Each section (e.g., pooled cohort mortality) ends with a summary of the main findings. At the end of the Results chapter, two overview tables summarize the main findings for the pooled cohort and the administrative cohorts for mortality and morbidity.

### **POOLED COHORT**

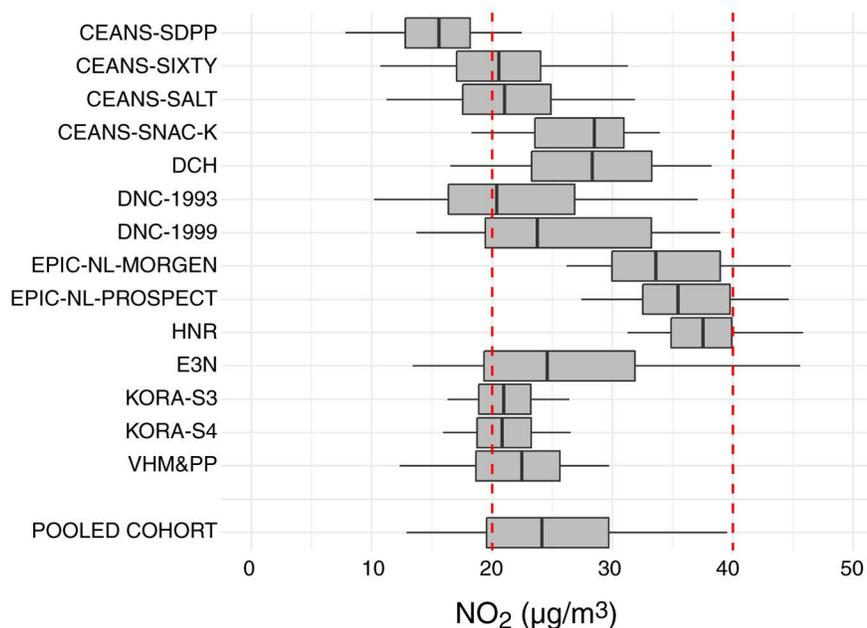
We first show exposure distributions by cohort and covariate distributions by cohort, and then the key epidemiological findings.

#### **Exposure**

Figures 2 and 3 and Appendix 2, (available on the HEI website) Table P1 show the exposure distributions for



**Figure 2. Description of PM<sub>2.5</sub> exposure at participant addresses in the pooled cohort.** The boundary of the box closest to zero indicates P25, the boundary furthest from zero indicates P75, the bold line in the middle of the box indicates P50, and the whiskers indicate P5 and P95. Red dotted lines indicate the WHO air quality guideline (10 µg/m<sup>3</sup>), the U.S. EPA NAAQS (12 µg/m<sup>3</sup>), and the EU Limit Value (25 µg/m<sup>3</sup>) — all annual averages. (Adapted from Hvitfeldt et al 2021; Creative Commons license CC BY-NC-ND 4.0)



**Figure 3. Description of NO<sub>2</sub> exposure at participant addresses in the pooled cohort.** The boundary of the box closest to zero indicates P25, the boundary furthest from zero indicates P75, the bold line in the middle of the box indicates P50; and the whiskers indicate P5 and P95. Red dotted lines indicate the WHO air quality guideline (40 µg/m<sup>3</sup>) and the WHO Health Risks of Air Pollution in Europe health impact quantification threshold (20 µg/m<sup>3</sup>) — all annual averages. P = percentile. (Adapted from Hvitfeldt et al 2021; Creative Commons license CC BY-NC-ND 4.0)

PM<sub>2.5</sub> and NO<sub>2</sub>, respectively, based on the main ELAPSE models. Exposure distributions for BC and O<sub>3</sub> are shown in Appendix 2, Figures P1 and P2. Large north-to-south gradients in air pollution between cohorts were observed. Especially for BC and NO<sub>2</sub>, contrasts within cohorts were substantial. Contrast of O<sub>3</sub> exposure within cohorts was very small, for example, in the large VHM&PP cohort, the 5th and 95th percentiles of warm season O<sub>3</sub> exposure were 87 and 98 µg/m<sup>3</sup>, respectively. For 98% of the pooled cohort, O<sub>3</sub> exposures were between 60 and 100 µg/m<sup>3</sup>, and for 68% the exposures were between 80 and 100 µg/m<sup>3</sup>. For PM<sub>2.5</sub>, exposure contrast (on a relative scale) within cohorts was larger than for O<sub>3</sub>, but smaller than for BC and NO<sub>2</sub>. In the large VHM&PP cohort, the 5th and 95th percentiles of PM<sub>2.5</sub> exposure were 11 and 19 µg/m<sup>3</sup>, respectively. Virtually all subjects had 2010 PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below the EU limit values. More than 50,000 and 25,000 subjects had a residential PM<sub>2.5</sub> exposure below the U.S. EPA NAAQS (12 µg/m<sup>3</sup>) and the WHO guideline (10 µg/m<sup>3</sup>), respectively.

Correlations between the pollutants are shown per cohort in Appendix 2, Table P2. Especially BC and NO<sub>2</sub> were highly correlated in all cohorts. PM<sub>2.5</sub> was moderately to highly correlated with BC and NO<sub>2</sub>. O<sub>3</sub> was negatively correlated with PM<sub>2.5</sub> and with especially NO<sub>2</sub> and BC. The negative correlations for O<sub>3</sub> were especially high for annual and winter season averages. In the remainder of this report, we therefore focus on warm season O<sub>3</sub>.

Figure P3 in Appendix 2 illustrates the time trends in air pollution concentrations in four regions in Europe based upon the DEHM model (used for back-extrapolation). The model estimates background concentrations at a 26 × 26 km spatial scale. We present modeled concentrations because of better completeness compared with monitoring data. In Appendix 1, section M4, we document the comparability of modeled and monitored trends in areas and periods with available monitoring data. Figure P3 in Appendix 2 shows that concentrations of PM<sub>2.5</sub>, BC, and NO<sub>2</sub> have decreased substantially over time whereas O<sub>3</sub> concentrations have been more stable.

### Population Characteristics

Table 4 shows the main descriptive data of the individual cohorts contributing to the pooled cohort analyses. Further descriptive details of the individual cohorts are in Appendix 2, section P1. The VHM&PP cohort comprises nearly 50% of the pooled cohort. Relatively few subjects were excluded from the main analysis because of missing covariate, exposure, or outcome data (85% of the enrolled population). Cohorts differ in all characteristics, supporting the decision to perform analyses stratified by

subcohort. Four sub-cohorts exclusively included women (DNC-1993, DNC-1999, EPIC-NL Prospect, and E3N). The mean age varied substantially across cohorts. The percentage of subjects who were overweight ranged from 21% in the French female E3N cohort to 74% in the German HNR cohort. The percentage of current smokers ranged from 13% in the French female E3N cohort to 36% in the Danish DCH cohort. The relatively low percentage of included subjects in the KORA cohorts was due to lack of availability of data on area-level SES in part of the study area. Information on recruitment and follow-up is not included consistently in the design papers quoted in Appendix 2, section P1, and nonresponse research has not been performed by all cohorts. We expect that recruitment into the cohorts was not a random process, for example, subjects with more education may have participated more in the cohorts.

The population included in our main analytical model differed from the subjects lost due to missing covariates (Table P3). Subjects with missing covariates were younger and experienced slightly higher air pollution exposure. In the cohorts with the lowest exposure levels, the percentages of subjects with missing data on covariates were small. Of the 106,728 subjects in the Swedish and Danish cohorts, 99,520 subjects were included in the main analysis (93%) (Table 4). In Appendix 2, “Missing values” subsection of section P3, we documented that these losses did not result in selection bias.

### Natural- and Cause-Specific Mortality Analyses

**Single-Pollutant Analyses** Table 5 shows the HRs for associations between air pollution and natural-cause and cause-specific mortality. A large number of events occurred during the follow-up period of approximately 20 years on average. HRs were positive and statistically significant for almost all outcomes and PM<sub>2.5</sub>, NO<sub>2</sub>, and BC. Only the association between PM<sub>2.5</sub> and respiratory disease was not statistically significant. Effect estimates were similar for most causes of death. The largest HRs were found for diabetes mortality, but with wider CIs because of the smaller number of deaths compared with other outcomes. HRs are significantly negative for all outcomes and O<sub>3</sub> in the warm season.

Table 6 shows the HRs in three a priori defined confounder models. Overall, moderate decreases in HRs were found from the least-adjusted model 1 to the most-adjusted model 3. Inclusion of individual-level confounders tended to decrease the HRs (model 2 versus 1), whereas adjustment for area-level confounders increased HRs modestly (Model 3 versus model 2).

**Table 4.** Description of the Study Populations

Cohort	N <sup>a</sup>	N in Main Model (%) <sup>b</sup>	Person-Year Follow-up	Mean Age	Female (%)	BMI $\geq 25$ kg/m <sup>2</sup> n, (%)	Current Smoker (%)
CEANS-SDPP	7,835	7,727 (98.6)	122,606	47.1	4,727 (61)	3,982 (52)	2,038 (26)
CEANS-SIXTY	4,180	3,969 (95)	61,328	60.0	2,068 (52)	2,545 (65)	839 (21)
CEANS-SALT	6,724	6,176 (91.9)	64,140	57.8	3,417 (55)	2,458 (40)	1,311 (21)
CEANS-SNACK	3,248	2,830 (87.1)	20,970	72.9	1,767 (62)	1,501 (53)	404 (14)
DCH	56,308	53,647 (95.3)	974,868	56.7	28,134 (52)	30,071 (56)	19,459 (36)
DNC-1993	19,664	17,043 (86.7)	318,351	56.2	17,043 (100)	4,783 (28)	6,383 (37)
DNC-1999	8,769	8,128 (92.7)	117,390	47.9	8,128 (100)	2,439 (30)	2,325 (29)
E3N	53,521	39,006 (72.9)	650,664	53.0	39,006 (100)	8,058 (21)	5,060 (13)
EPIC-NL-MORGEN	20,711	18,302 (88.4)	308,257	42.9	10,060 (55)	8,986 (50)	6,359 (35)
EPIC-NL-PROSPECT	16,194	14,570 (90)	239,372	57.7	14,570 (100)	7,978 (55)	3,335 (23)
HNR	4,809	4,733 (98.4)	56,857	59.7	2,382 (50)	3,480 (74)	1,113 (24)
KORA-S3	4,566	2,572 (56.3)	40,080	49.4	1,308 (51)	1,722 (67)	519 (20)
KORA-S4	4,257	2,281 (53.6)	29,528	49.3	1,173 (51)	1,563 (69)	523 (23)
VHM&PP	170,250	144,383 (84.8)	3,335,141	42.1	81,117 (56)	61,249 (43)	28,916 (20)
Total	381,036	325,367 (85.4)	6,339,553	48.7	214,900 (66)	140,815 (43)	78,584 (24)

<sup>a</sup> Population size is the number of subjects for which information was transferred to Utrecht University for construction of the pooled cohort. These numbers may differ slightly from the original cohort populations shown in Appendix 2, section P1.

<sup>b</sup> Numbers refer to the mortality analyses; small differences with incidence analyses due to different confounders and exclusion of prevalent cases.

**Table 5.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause and Cause-Specific Mortality<sup>a</sup>

	Natural Cause	Cardio-vascular Disease	Ischemic Heart Disease	Cerebro-vascular Disease	Respiratory	COPD	Diabetes	Cardio-metabolic
Deaths (N)	47,131	15,542	7,265	3,740	2,865	1,711	1,034	16,576
PM <sub>2.5</sub>	1.130 (1.106, 1.155)	1.135 (1.095, 1.176)	1.111 (1.056, 1.169)	1.128 (1.048, 1.214)	1.054 (0.961, 1.156)	1.131 (1.002, 1.278)	1.316 (1.144, 1.514)	1.146 (1.107, 1.186)
NO <sub>2</sub>	1.086 (1.070, 1.102)	1.089 (1.060, 1.120)	1.098 (1.053, 1.145)	1.068 (1.011, 1.129)	1.101 (1.038, 1.168)	1.141 (1.056, 1.233)	1.238 (1.112, 1.378)	1.098 (1.069, 1.128)
BC	1.081 (1.065, 1.098)	1.085 (1.055, 1.116)	1.078 (1.033, 1.125)	1.075 (1.016, 1.138)	1.084 (1.020, 1.151)	1.119 (1.034, 1.211)	1.240 (1.112, 1.382)	1.095 (1.065, 1.125)
O <sub>3</sub>	0.896 (0.878, 0.914)	0.887 (0.854, 0.922)	0.870 (0.821, 0.921)	0.882 (0.817, 0.953)	0.890 (0.821, 0.966)	0.861 (0.774, 0.957)	0.744 (0.645, 0.859)	0.877 (0.845, 0.910)

<sup>a</sup> N = 325,367. HR (95% CI) (95% confidence interval) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC, 0.5 × 10<sup>-5</sup>/m; and O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for cohort ID, age, sex, year of baseline visit, smoking (status, duration, intensity, and intensity squared), BMI, marital status, employment status, and 2001 neighborhood-level mean income.

**Table 6.** Hazard Ratios in Models with Increasing Level of Adjustment for Confounders for Natural-Cause, Cardiovascular, and Respiratory Mortality

Pollutant	Model 1 <sup>a</sup> HR (95% CI)	Model 2 <sup>b</sup> HR (95% CI)	Model 3 <sup>c</sup> HR (95% CI)
<b>Natural-Cause Mortality</b>			
PM <sub>2.5</sub>	1.152 (1.127, 1.177)	1.121 (1.097, 1.146)	1.130 (1.106, 1.155)
NO <sub>2</sub>	1.113 (1.097, 1.129)	1.070 (1.055, 1.086)	1.086 (1.070, 1.102)
BC	1.113 (1.096, 1.129)	1.069 (1.053, 1.085)	1.081 (1.065, 1.098)
O <sub>3</sub>	0.844 (0.828, 0.862)	0.899 (0.881, 0.918)	0.896 (0.878, 0.914)
<b>Cardiovascular Mortality</b>			
PM <sub>2.5</sub>	1.136 (1.096, 1.177)	1.121 (1.082, 1.162)	1.135 (1.095, 1.176)
NO <sub>2</sub>	1.094 (1.065, 1.123)	1.064 (1.036, 1.093)	1.089 (1.060, 1.120)
BC	1.094 (1.064, 1.125)	1.065 (1.035, 1.095)	1.085 (1.055, 1.116)
O <sub>3</sub>	0.858 (0.826, 0.891)	0.899 (0.866, 0.934)	0.887 (0.854, 0.922)
<b>Respiratory Mortality</b>			
PM <sub>2.5</sub>	1.070 (0.974, 1.175)	1.025 (0.934, 1.125)	1.054 (0.961, 1.156)
NO <sub>2</sub>	1.134 (1.071, 1.201)	1.052 (0.992, 1.115)	1.101 (1.038, 1.168)
BC	1.128 (1.063, 1.197)	1.046 (0.985, 1.111)	1.084 (1.020, 1.151)
O <sub>3</sub>	0.788 (0.727, 0.855)	0.890 (0.820, 0.966)	0.890 (0.821, 0.966)

<sup>a</sup> Model 1 is adjusted for age (time axis), sex (strata), and calendar year of recruitment.

<sup>b</sup> Model 2 adds individual-level confounders to model 1 (Table 5).

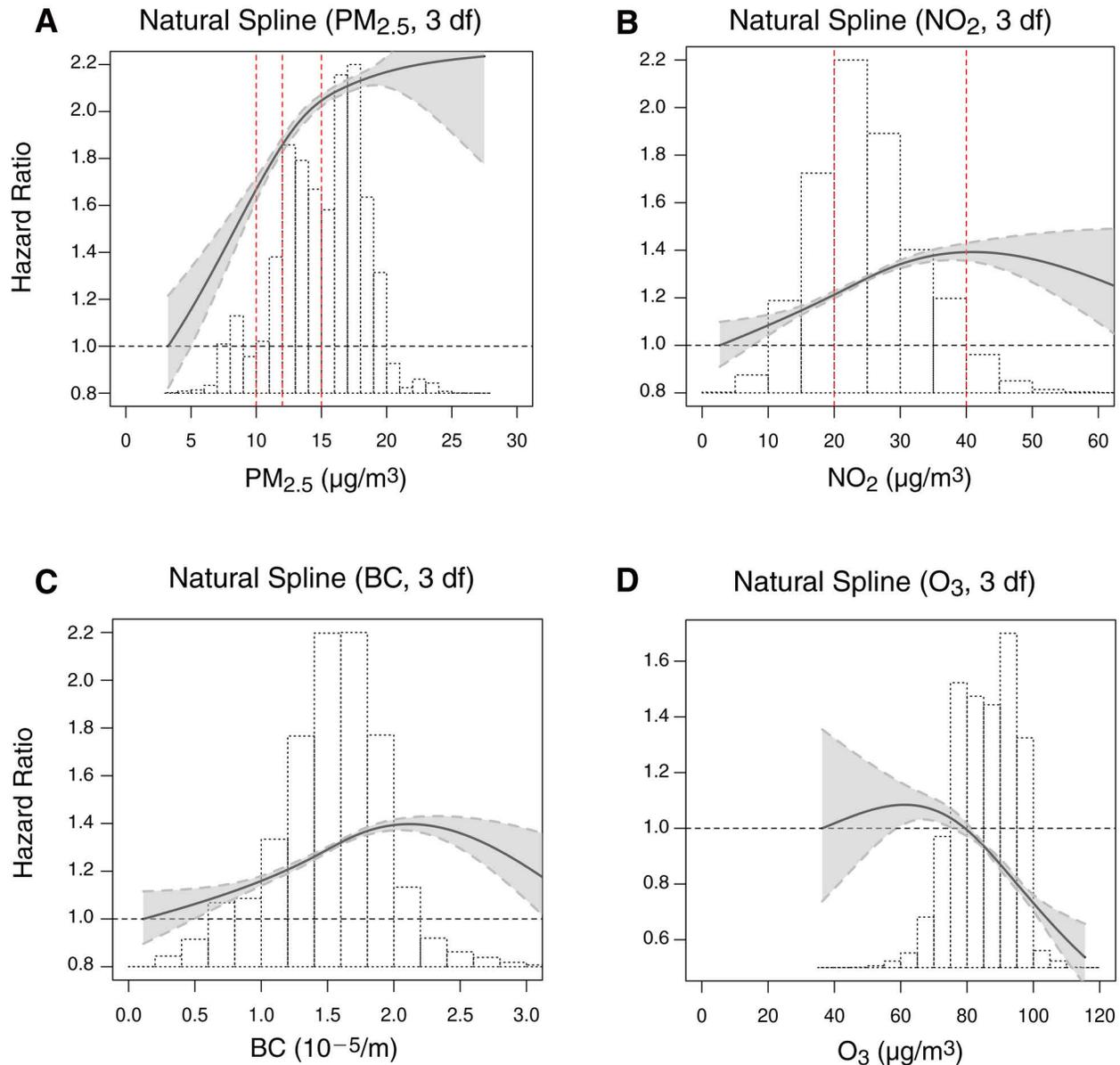
<sup>c</sup> Model 3 adds neighborhood income-level to model 2.

**Concentration–Response Functions** Figure 4 shows shapes of concentration–response functions for natural-cause mortality in the pooled cohort based upon natural splines with three degrees of freedom. Spline functions tend to be steeper at low concentrations for PM<sub>2.5</sub>, BC, and NO<sub>2</sub>, consistent with a supralinear function. At the low and high end of the distribution, confidence limits are wide because of scarce data, as illustrated by the histogram. We prefer to show the full curves but note that, for example, for NO<sub>2</sub> exposures higher than 40 µg/m<sup>3</sup> we cannot distinguish between a flat or a decreasing curve. For O<sub>3</sub> the negative association is dominated by the trend in a narrow range of exposures between 70 and 100 µg/m<sup>3</sup>. Concentration–response functions for natural-cause mortality with natural splines of two and four degrees of freedom were similar to those with three degrees of freedom (Appendix 2, Figures P4 and P5) and penalized splines (Appendix 2, Figure P6). The penalized spline functions were substantially less smooth than the natural splines related to the automatic selection of degrees of

freedom. We consider these patterns more difficult to interpret. The interpretation of the shape of concentration–response plots is complicated because cohorts from different countries with different exposure levels were included. This issue applied especially for PM<sub>2.5</sub>, where the gradient of exposure across Europe was large and the low end of the exposure distribution is primarily in the north of Europe. For NO<sub>2</sub> and BC this was much less of an issue as subjects residing in major cities and near major roads had higher exposure in all countries.

Concentration–response functions for cause-specific mortality were, in general, similar to the natural-cause mortality indicating mostly supralinear curves (Appendix 2, Figures P7 to P13).

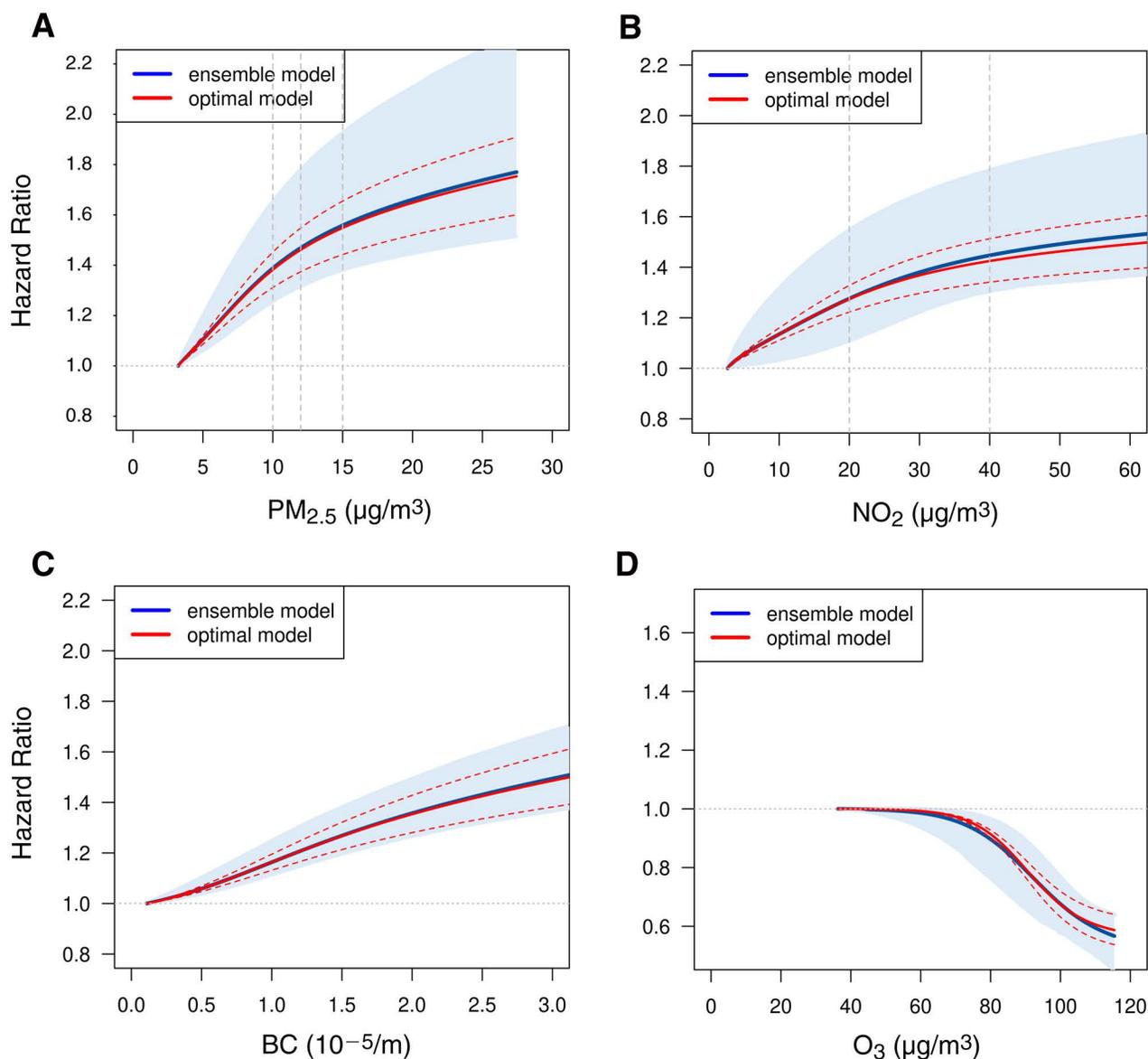
The shape-constrained health impact functions (SCHIF) in Figure 5 generally confirm the results of the natural splines shown in Figure 4 with steeper positive associations between pollution and natural-cause mortality at the low end of the distribution. The optimal and ensemble



**Figure 4. Natural cubic splines (three degrees of freedom) for associations between air pollution exposure and natural-cause mortality, adjusted model 3.** Red dotted lines are air quality limit and guideline values. X-axis truncated at 30, 60, and 120  $\mu\text{g}/\text{m}^3$  for  $\text{PM}_{2.5}$ ,  $\text{NO}_2$ , and  $\text{O}_3$ , respectively, and at  $3 \times 10^{-5}/\text{m}$  for BC. Hazard ratios are expressed relative to minimum exposure. Histograms are of exposure distributions. (From Strak et al. [In press])

model functions show very similar shapes. The uncertainty around the ensemble model is larger (blue shaded area) than around the optimal model (red dotted line). The optimal is a single model, the ensemble model is a weighted average from multiple models. The SCHIF functions deviate from the natural spline functions at the high

end of the distributions, for example, for  $\text{PM}_{2.5}$ ,  $\text{NO}_2$ , and BC. Figure 5 suggests an increase in HR whereas the splines suggest a flattening or decrease with wide CIs. The difference is likely related to the parametric nature of SCHIF versus the more empirical nature of the spline functions. SCHIF functions for cause-specific mortality



**Figure 5. SCHIF for associations between air pollution exposure and natural-cause mortality.** Vertical dotted lines are air quality limit and guideline values. Hazard ratios are expressed relative to minimum exposure. (From Strak et al. [In press])

are generally similar to those for natural-cause mortality (Appendix 2, Figures P14 to P20). Most but not all show a supralinear curve for PM<sub>2.5</sub>, NO<sub>2</sub>, and BC. The SCHIFs for respiratory and COPD mortality and NO<sub>2</sub> and for diabetes mortality and PM<sub>2.5</sub>, NO<sub>2</sub>, and BC suggest a lower slope at low versus high concentrations.

**Subset and Threshold Analysis** Table 7 shows the HRs observed for successively lower concentrations of pollutants

for natural-cause mortality. HRs remain elevated and significant for PM<sub>2.5</sub> even when all observations higher than 12 µg/m<sup>3</sup> are removed from the analysis. The HR below 10 µg/m<sup>3</sup> is similar to the HR with all observations, but with wider CIs. Effect estimates tend to be higher over the low concentration ranges. HRs remain elevated and significant for NO<sub>2</sub> even when all observations higher than 20 µg/m<sup>3</sup> are removed from the analysis. HRs remain elevated and significant for BC even when all observations

**Table 7.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause Mortality in Subset Analysis<sup>a</sup>

Pollutant	Subset	N	HR (95% CI)
<b>PM<sub>2.5</sub></b>			
	Full dataset	325,367	1.130 (1.106, 1.155)
	<25 µg/m <sup>3</sup>	325,339	1.131 (1.107, 1.156)
	< 20 µg/m <sup>3</sup>	316,540	1.138 (1.113, 1.164)
	<15 µg/m <sup>3</sup>	151,250	1.257 (1.193, 1.324)
	<12 µg/m <sup>3</sup>	52,528	1.296 (1.140, 1.474)
	<10 µg/m <sup>3</sup>	25,422	1.146 (0.931, 1.410)
<b>NO<sub>2</sub></b>			
	Full dataset	325,367	1.086 (1.070, 1.102)
	<40 µg/m <sup>3</sup>	310,643	1.101 (1.083, 1.119)
	<30 µg/m <sup>3</sup>	247,039	1.114 (1.088, 1.140)
	<20 µg/m <sup>3</sup>	88,510	1.099 (1.033, 1.170)
<b>BC</b>			
	Full dataset	325,367	1.081 (1.065, 1.098)
	<3 × 10 <sup>-5</sup> /m	324,757	1.084 (1.067, 1.101)
	<2.5 × 10 <sup>-5</sup> /m	320,632	1.091 (1.074, 1.109)
	<2 × 10 <sup>-5</sup> /m	296,371	1.107 (1.087, 1.128)
	<1.5 × 10 <sup>-5</sup> /m	142,032	1.125 (1.086, 1.165)
	<1 × 10 <sup>-5</sup> /m	35,406	1.041 (0.942, 1.150)
	<0.5 × 10 <sup>-5</sup> /m	5,013	1.173 (0.568, 2.422)
<b>O<sub>3</sub></b>			
	Full dataset	325,367	0.896 (0.878, 0.914)
	<120 µg/m <sup>3</sup>	325,367	0.896 (0.878, 0.914)
	<100 µg/m <sup>3</sup>	320,522	0.891 (0.872, 0.910)
	<80 µg/m <sup>3</sup>	98,840	0.979 (0.940, 1.018)
	<60 µg/m <sup>3</sup>	1,724	0.618 (0.337, 1.135)

<sup>a</sup>HR (95% CI) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC, 0.5 × 10<sup>-5</sup>/m; O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for cohort ID, age, sex, year of baseline visit, smoking (status, duration, intensity, intensity squared), BMI, marital status, employment status, and 2001 neighborhood-level mean income.

higher than 1.5 × 10<sup>-5</sup>/m are removed from the analysis. Effect estimates are higher over the low concentration ranges. HRs for O<sub>3</sub> attenuate toward unity at concentrations below 80 µg/m<sup>3</sup>. We note that the subsets contain different cohorts especially for PM<sub>2.5</sub> where there are large differences between cohorts (Figures 2 and 3). All cohorts contribute to the analysis of HRs below 15 µg/m<sup>3</sup> PM<sub>2.5</sub>,

though the Dutch and German cohort contributed little (Appendix 2, Table P1). The Stockholm, Danish, French, and Austrian cohorts contribute to the analysis below 10 and 12 µg/m<sup>3</sup> PM<sub>2.5</sub>, but the analysis below 10 µg/m<sup>3</sup> PM<sub>2.5</sub> is dominated by the Stockholm cohort. All cohorts except HNR and the Dutch cohorts contributed to the analysis of NO<sub>2</sub> concentrations below 20 µg/m<sup>3</sup>. The populations included in the different subsets may additionally differ in important aspects, for example, there are more rural populations in the lower exposure subsets.

Similar patterns were found for cause-specific mortality (Appendix 2, Tables P4 to P6), though CIs were wider related to the smaller number of deaths. For PM<sub>2.5</sub>, HRs remain elevated for subjects with concentrations below 12 but not 10 µg/m<sup>3</sup> for cardiovascular mortality (Table P4). For NO<sub>2</sub>, HRs were elevated and higher for subjects with concentrations below 20 µg/m<sup>3</sup> compared to the full population. For BC, associations remain elevated for subjects with concentrations all below 1.5 × 10<sup>-5</sup>/m. For the respiratory mortality outcomes (Table P5), HRs for PM<sub>2.5</sub> and BC remain elevated for subjects with the lowest concentrations though with wide CIs. For NO<sub>2</sub>, HRs were smaller for subjects with concentrations below 30 and 20 µg/m<sup>3</sup>. The subset analyses for cardiometabolic mortality (Table P6) show the same results as for cardiovascular disease mortality, as the two categories largely overlap. For PM<sub>2.5</sub>, HRs were elevated for diabetes mortality for subjects with concentrations all below 15 but not 12 µg/m<sup>3</sup>. For NO<sub>2</sub>, associations remained elevated for subjects with concentrations all below 20 µg/m<sup>3</sup>. For BC, associations remain elevated in all subsets.

In Appendix 2, Tables P7 to P10 show the results of the threshold analyses for natural-cause and cause-specific mortality. For PM<sub>2.5</sub>, NO<sub>2</sub>, and BC the no-threshold model had the lowest AIC (indicating better performance) for natural-cause mortality (Appendix 2, Table P7). For PM<sub>2.5</sub>, the AIC is clearly lower for the linear model compared with the model with a threshold at 10 µg/m<sup>3</sup>, but similar to models with a threshold at 5 or 7.5 µg/m<sup>3</sup> (Table P7). We somewhat arbitrarily interpret differences of AIC of at least one unit as indicating a meaningful difference. The rejection of a threshold at 10 µg/m<sup>3</sup> is consistent with the graphical and subset analyses. The scarcity of PM<sub>2.5</sub> exposures well below 10 µg/m<sup>3</sup> limits the power to detect thresholds at 5 or 7.5 µg/m<sup>3</sup>. For NO<sub>2</sub>, models with thresholds of 20 and 15 µg/m<sup>3</sup> have clearly poorer performance than the linear model.

For the three cardiovascular mortality endpoints, there was no difference in AIC between the four models for PM<sub>2.5</sub>, indicating that neither a threshold at 10 µg/m<sup>3</sup>, nor

a linear shape can be excluded (Table P8). For NO<sub>2</sub>, the pattern for cardiovascular mortality was similar to the pattern for natural mortality (no evidence of a threshold at 20 or 15 µg/m<sup>3</sup>). For BC, models with a threshold at 1 × 10<sup>-5</sup>/m performed best for cardiovascular disease and ischemic heart disease but not cerebrovascular disease mortality. For O<sub>3</sub>, models with a threshold at 80 µg/m<sup>3</sup> performed best. For respiratory and COPD mortality, there was no difference in AIC between the four models for PM<sub>2.5</sub> and NO<sub>2</sub>, indicating thresholds at, respectively, 10 or 20 µg/m<sup>3</sup> cannot be excluded (Table P9). For diabetes mortality and PM<sub>2.5</sub>, there was no difference in AIC between the four models (Table P10). For BC and NO<sub>2</sub>, the models with the highest threshold performed less than the other models.

**Two-Pollutant Analyses** Table 8 shows results of the two-pollutant analyses for natural-cause mortality. We present all two-pollutant models, also when the correlation between pollutants was high. We have flagged in the tables and text where we judge that limited interpretation of a two-pollutant model is possible. Effect estimates for PM<sub>2.5</sub> attenuated modestly from 1.13 to 1.08–1.09 after adjustment for NO<sub>2</sub>, BC, and O<sub>3</sub>. All adjusted PM<sub>2.5</sub> effect estimates remained statistically significant. CIs increased only mildly in two-pollutant models, consistent with the moderate correlations of other pollutants with PM<sub>2.5</sub>, suggesting that these models lead to interpretable results. Effect estimates for NO<sub>2</sub> attenuated from 1.09 to 1.05 after adjustment for PM<sub>2.5</sub> or O<sub>3</sub> and remained significant, with only mild increases in the CIs. Effect estimates for BC were attenuated from 1.08 to 1.04 after adjustment for PM<sub>2.5</sub> or

O<sub>3</sub> and remained significant. Two-pollutant models of BC and NO<sub>2</sub> are difficult to interpret because of the very high correlation ( $R > 0.8$  in most cohorts) between BC and NO<sub>2</sub> (Appendix 2, Table P2). Effect estimates for O<sub>3</sub> remained statistically significantly negative in two-pollutant models, with only a small shift toward unity.

For cardiovascular mortality, the same patterns as for natural-cause mortality were found (Appendix Table P11). Effect estimates for PM<sub>2.5</sub> were attenuated modestly in two-pollutant models, whereas the effect estimates for NO<sub>2</sub> and especially BC were reduced substantially.

For respiratory mortality, the nonsignificant single-pollutant HR of PM<sub>2.5</sub> was reduced to below unity in all two-pollutant models (Appendix 2, Table P12). Effect estimates for NO<sub>2</sub> remained most stable in two-pollutant models. Effect estimates for BC were robust against adjustment for PM<sub>2.5</sub> but not NO<sub>2</sub> and O<sub>3</sub>.

**Additional Analyses**

In this section, we present a large number of sensitivity analyses with respect to exposure specification and confounder control. In general, associations found in these sensitivity analyses were similar to the associations we found in the main analyses.

**Time-Varying and Baseline Exposure Analyses** Tables 9 and 10 show the HRs from applying exposure at baseline and the time-varying exposure analysis instead of the 2010 exposure in the main model. HRs and CIs were generally smaller than in the main analysis, especially for PM<sub>2.5</sub> when using the back-extrapolated exposure at baseline

**Table 8.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause Mortality in Two-Pollutant Models<sup>a,b</sup>

Pollutant	Single Pollutant HR (95% CI)	HR (95% CI) Adjusted for PM <sub>2.5</sub>	HR (95% CI) Adjusted for NO <sub>2</sub>	HR (95% CI) Adjusted for BC	HR (95% CI) Adjusted for O <sub>3</sub>
PM <sub>2.5</sub>	1.130 (1.106, 1.155)	NA	1.083 (1.054, 1.113)	1.092 (1.062, 1.123)	1.089 (1.061, 1.117)
NO <sub>2</sub>	1.086 (1.070, 1.102)	1.050 (1.031, 1.070)	NA	1.074 (1.038, 1.112)	1.053 (1.032, 1.074)
BC	1.081 (1.065, 1.098)	1.039 (1.019, 1.060)	1.012 (0.977, 1.048)	NA	1.044 (1.024, 1.065)
O <sub>3</sub>	0.896 (0.878, 0.914)	0.935 (0.913, 0.957)	0.940 (0.914, 0.966)	0.930 (0.906, 0.955)	NA

<sup>a</sup> NA = Not applicable;  $N = 325,367$ . HR (95% CI) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC, 0.5 × 10<sup>-5</sup>/m; O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for cohort ID, age, sex, year of baseline visit, smoking (status, duration, intensity, intensity squared), BMI, marital status, employment status, and 2001 neighborhood-level mean income.

<sup>b</sup> Two-pollutant models of BC and NO<sub>2</sub> are difficult to interpret because of high correlation between BC and NO<sub>2</sub>.

**Table 9.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause Mortality in Time-Varying Analyses: Baseline Exposure

Pollutant	Main Model HR (95% CI) <sup>a</sup>	Back-Extrapolated Baseline Exposure (ratio method) HR (95% CI)	Back-Extrapolated Baseline Exposure (difference method) HR (95% CI)
PM <sub>2.5</sub>	1.130 (1.106, 1.155)	1.048 (1.038, 1.059)	1.058 (1.042, 1.075)
NO <sub>2</sub>	1.086 (1.070, 1.102)	1.056 (1.045, 1.067)	1.082 (1.067, 1.098)
BC	1.081 (1.065, 1.098)	1.046 (1.034, 1.058)	1.072 (1.057, 1.088)
O <sub>3</sub>	0.896 (0.878, 0.914)	0.918 (0.902, 0.935)	0.916 (0.899, 0.933)

<sup>a</sup> Main model 3 restricted to subjects available in the back-extrapolated baseline exposure analysis ( $N = 325,342$ ).

**Table 10.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause Mortality in Time-Varying Analyses: Time-Varying Exposure Including Residential Mobility to Adjust for Time Trends in Mortality and Pollution

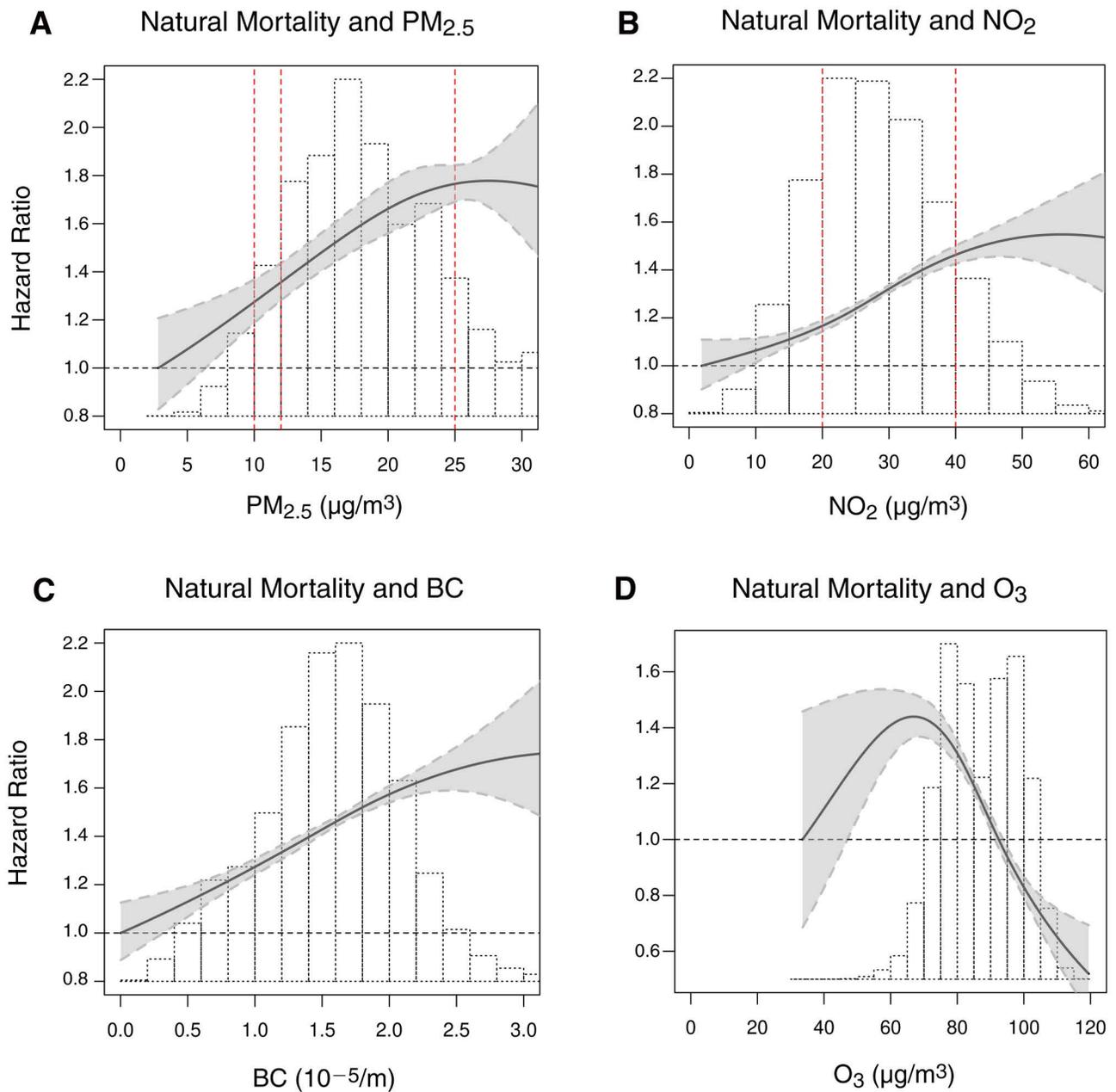
Pollutant	Main Model HR, Reduced Set of Cohorts HR (95% CI) <sup>a</sup>	Strata for 1 Year <sup>b</sup> (ratio method) HR (95% CI)	Strata for 1 Year (difference method) HR (95% CI)	Strata for 5 Year (ratio method) HR (95% CI)	Strata for 5 Year (difference method) HR (95% CI)
PM <sub>2.5</sub>	1.123 (1.090, 1.157)	1.123 (1.093, 1.153)	1.136 (1.103, 1.170)	1.085 (1.061, 1.109)	1.081 (1.057, 1.107)
NO <sub>2</sub>	1.104 (1.083, 1.126)	1.109 (1.089, 1.130)	1.112 (1.091, 1.133)	1.112 (1.093, 1.133)	1.116 (1.095, 1.138)
BC	1.098 (1.076, 1.120)	1.108 (1.086, 1.131)	1.103 (1.081, 1.124)	1.085 (1.064, 1.107)	1.102 (1.081, 1.124)
O <sub>3</sub>	0.887 (0.864, 0.910)	0.932 (0.920, 0.945)	0.932 (0.920, 0.944)	0.897 (0.887, 0.907)	0.895 (0.885, 0.906)

<sup>a</sup> Main model 3 restricted to subjects available in the time-varying analysis ( $N = 185,585$ ).

<sup>b</sup> The 1-year and 5-year strata refer to calendar time periods we used to adjust for time trends in mortality.

(Table 9). HRs were similar to the HR in the main analysis in the time-varying analyses (Table 10). Figures P3 show the temporal trends in PM<sub>2.5</sub>, NO<sub>2</sub>, BC, and O<sub>3</sub> concentrations based upon the Danish DEHM model, which was used for back-extrapolation. In Appendix 2, Figure P21 shows the distribution of back-extrapolated exposures in comparison with the 2010 exposures. Especially PM<sub>2.5</sub> exposures were substantially higher at baseline than in 2010. NO<sub>2</sub> and BC exposures were higher at baseline as well, but the difference was smaller than for PM<sub>2.5</sub>. O<sub>3</sub> exposures were only modestly higher at baseline. The smaller HRs for the back-extrapolated baseline exposure analysis, are probably due to relating the same health outcome to a larger absolute difference in exposure related to applying spatially varying ratios higher than unity to the 2010 spatial surface.

When we applied exposure surfaces from models based upon AirBase monitoring data of 2000 and 2005 to estimate exposure instead of the back-extrapolated exposure based upon the DEHM model, HRs were mostly consistent with the 2010 model. HRs were 1.056 (CI = 1.042, 1.069) for NO<sub>2</sub> in 2000, 1.102 (CI = 1.079, 1.126) for NO<sub>2</sub> in 2005, 1.156 (CI = 1.125, 1.188) for PM<sub>2.5</sub> in 2013, 0.925 (CI = 0.909, 0.941) for warm-season O<sub>3</sub> in 2000, and 0.908 (CI = 0.892, 0.924) for warm-season O<sub>3</sub> in 2005. For PM<sub>2.5</sub> there were insufficient monitoring sites to develop models for the years 2000 and 2005. Figure 6 shows the results of the natural spline models in time-varying exposure analyses. This documents that the observation of associations at low pollution levels was also found in the time-varying exposure analysis and thus not fully explained by the use of the relatively recent — and therefore relatively low — 2010 exposure. The modest



**Figure 6. Natural cubic splines (three degrees of freedom) for associations between time-varying exposure and natural-cause mortality.** Red dotted lines are air quality limit and guideline values. Hazard ratios are expressed relative to minimum exposure. Histograms show exposure distribution. (From Strak et al. [In press])

shift of the curve despite the downward trends in exposure is related to the large number of events that occurred in the later years of follow-up. In 2010–2017, 33% of all deaths occurred; in 2005–2017, 63% of all deaths occurred, and in 2000–2017 84% of all deaths occurred.

When the start of follow-up was defined as the year 2000, 2005, and 2008 such that the exposure more closely represented exposure during most of the follow-up time, we found HRs were smaller than in the main analysis though still statistically significant (Appendix 2, Table P13).

#### **Alternative Adjustment for Differences Between Cohorts**

In Appendix 2, Figures P22 to P25 and Tables P14 and P15 show the results of the sensitivity analyses for natural-cause mortality with respect to alternative adjustment for differences between cohorts (Figure P22), assessing impact of individual cohorts on the HR (Figures P23 to P25), adjustment for additional confounders (Table P14), and application of multiple imputations to impute missing values of confounders (Table P15). In general HRs were robust to these alternative specifications both for natural-cause and cause-specific mortality.

The methods to adjust for differences between subcohorts resulted in very similar effect estimates and CIs (Figure P22). Not adjusting for cohort at all resulted in smaller though still significant effect estimates for  $PM_{2.5}$ , BC, and  $NO_2$ .

HRs were robust when individual cohorts were excluded from the analysis, especially for  $PM_{2.5}$  (Figure P23). HRs were modestly reduced for  $NO_2$  and BC when the large VHM&PP and DCH cohorts were excluded, but still highly significant. An individual cohort analysis further showed that in most cohorts a positive association with  $PM_{2.5}$ ,  $NO_2$ , and BC was found (Figure P24). HRs were similar to the full population HRs for the subsets of cohorts used in the three incidence analyses (Figure P25).

**Adjustment for Additional Confounders** Further adjustment for education, smoking amount, and duration in past smokers, as well as dietary factors in a pooled cohort including cohorts that had information on these factors, did not affect effect estimates (Appendix 2, Table P14). Further adjustment for occupational status (blue vs. white collar job) and additional area-level SES variables also did not affect HRs. Effect estimates were not ( $PM_{2.5}$ ) or mildly ( $NO_2$ , BC, and  $O_3$ ) affected by exclusion of specific cohorts, such as the large VHM&PP cohort having no data on education, smoking, and dietary factors.

#### **Evaluation of Loss of Information Due to Missing Covariate Data**

HRs were almost identical to the main model HRs when we used multiple imputation to estimate missing covariate data and when we used the full study population (Table P15). We furthermore found no difference in HRs from the minimally adjusted model 1 in the full study population and in the complete case study population, suggesting that selection bias due to excluding subjects with missing covariates was not important (Table P16).

#### **Comparison with Local ESCAPE and Dispersion Models**

In Appendix 2, Table P17 provides the correlations between the study-area specific ESCAPE models and the ELAPSE exposure model per cohort. In this analysis, the DNC, most of the French E3N, and the Aarhus part of DCH are not included because of missing ESCAPE exposure estimates. In the four Stockholm cohorts a zero correlation between the ESCAPE and ELAPSE models for all three pollutants was found. In the other cohorts, the correlation varied substantially from low to high. The mean within-cohort correlation for  $PM_{2.5}$ ,  $NO_2$ , and BC was 0.45, 0.66, and 0.57, respectively. For the Stockholm, Danish, and HNR cohorts, estimates from local dispersion models were available (Table P17). The correlations for the Stockholm cohort of the ELAPSE exposure estimates were much higher for BC (0.51 to 0.83 in the four subcohorts) than for the ESCAPE models. For  $PM_{2.5}$ , the correlations for the Stockholm cohorts were low but not zero (0.08 to 0.51).

In Appendix 2, Table P18 compares HRs estimated with the study-area specific ESCAPE models and the ELAPSE exposure model for those cohorts that had effect estimates available for both models ( $N = 209,885$ ). HRs for natural-cause mortality were significant for both models and somewhat larger for ESCAPE for  $PM_{2.5}$  and smaller for  $NO_2$  and BC compared with ELAPSE. HRs were smaller compared with the full ELAPSE population (Table 5). For cardiovascular mortality, HRs for the ESCAPE model were smaller and statistically nonsignificant. For respiratory mortality, HRs for both models were statistically nonsignificant with no clear pattern of effect size between the models.

**Measurement Error Analyses** In Appendix 2, Tables P19 and P20 show the results of the measurement error analyses. HRs were robust to using the five hold-out validation models instead of the main model based upon the full dataset (Table P19). Application of regression calibration resulted in very small changes in the HR and CI (Table P20). The correction was applied in model 1, but the correction likely results in similar changes in model 3, as the HRs from models 1 and 3 were very similar (Table 6). The measurement error correction only accounted for errors in

our model predicting ambient concentrations. We did not have validation data from personal exposure monitoring to use as a gold standard in our regression calibration application. Furthermore, it should be noted that our epidemiological models assessed the association with outdoor air pollution, so in principle our gold standard should be personal exposure from outdoor sources. As a consequence, the model uses the slope from the measurement error model and mostly accounts for the classical error incorporated in the complex mix of classical and Berkson error in air pollution exposure estimates as the variability in individual exposure is not captured by the model.

**Adjustment for Traffic Noise** In Appendix 2, Table P21 shows the results of additional adjustment for traffic noise for natural-cause mortality, the cardiovascular mortality endpoints, and diabetes mortality. Traffic noise was available for 109,021 subjects, 30% of the full population. Traffic noise was available for Stockholm, two Danish, and two German cohorts. Further adjustment for traffic noise resulted in small decreases in effect estimates for natural-cause, cardiovascular, and cardiometabolic mortality. For cerebrovascular mortality, further adjustment for traffic noise resulted in small decreases in effect estimates for PM<sub>2.5</sub> and somewhat larger decreases for NO<sub>2</sub> and BC. For ischemic heart disease mortality, effect estimates did not change, but in this restricted pooled cohort, generally null associations were found. For diabetes, HRs were reduced most for PM<sub>2.5</sub>.

**Effect Modification** Figure P26 in Appendix 2 shows the results of effect modification analyses conducted for PM<sub>2.5</sub> and NO<sub>2</sub>. We found no consistent differences in HRs between age categories, smoking status, and BMI — the covariates available in all cohorts. HRs were smaller in the elderly for PM<sub>2.5</sub>, but larger for NO<sub>2</sub>. HRs were larger in current smokers for PM<sub>2.5</sub>, but there was no difference at all for NO<sub>2</sub>. No differences in HRs were found between normal weight, overweight, and obese subjects. The small underweight group tended to have lower HRs, but CIs were wide for this subgroup. Because we used sex as a stratum in the analysis, we did not perform an interaction analysis for sex.

**Associations with PM<sub>2.5</sub> Exposure from MAPLE** PM<sub>2.5</sub> from the MAPLE method was slightly lower than PM<sub>2.5</sub> from our main ELAPSE model (Appendix 2, Table P22 and Figures P27–P30). The mean PM<sub>2.5</sub> concentration was 13.6 µg/m<sup>3</sup> (standard deviation = 3.9 µg/m<sup>3</sup>) for MAPLE versus 15.0 µg/m<sup>3</sup> (standard deviation = 3.2 µg/m<sup>3</sup>) for ELAPSE. Maps of ELAPSE and MAPLE PM<sub>2.5</sub> showed very similar patterns across Europe (Appendix 2, Figure P27). In individual cohorts, differences between ELAPSE and

MAPLE were large however, such as the Danish and German KORA cohort, in which the MAPLE estimates were substantially lower. The difference between the Stockholm and Danish cohorts was very small for the MAPLE method. ESCAPE measurements in 2010 in Copenhagen and Augsburg showed an average of 11 and 14 µg/m<sup>3</sup>, respectively (Eeftens et al. 2012b), suggesting that especially the KORA MAPLE estimates may be too low. Routine monitoring data from Denmark were higher than both the ELAPSE and MAPLE modeled concentrations. The overall correlation between the two model estimations was high (Appendix 2, Figure P28). The median correlation of MAPLE exposure and ELAPSE PM<sub>2.5</sub> per cohort was moderate but also highly variable: 0.37 (CI = -0.14, 0.72) (Table P23).

Larger differences were found for PM<sub>2.5</sub> exposure from earlier periods. Figure P30 shows that the PM<sub>2.5</sub> exposure in 2001 (the earliest available year from MAPLE) was slightly lower in most cohorts compared with 2010. In some cohorts, such as the Dutch and German HNR cohort, the 2001 exposure were substantially lower than in 2010. This trend differs substantially from the trend applied in ELAPSE based upon the DEHM model (Figure P21).

Table 11 compares linear associations between natural-cause and cause-specific mortality with PM<sub>2.5</sub> from the main ELAPSE model and the MAPLE method. HRs were very similar between the two exposure models for natural-cause mortality and most cause-specific outcomes, with the exception of COPD mortality. The HR for PM<sub>2.5</sub> using the MAPLE 2001 estimate (instead of the 2010 estimates in Table 11) was 1.118 (CI = 1.093, 1.144), which was much larger than the estimate for the ELAPSE back-extrapolated exposure: 1.048 (CI = 1.038, 1.059) in Table 9. The HRs were also comparable between ELAPSE and MAPLE exposure when we restricted the data to the cohorts available for the incidence analyses (Figure P31).

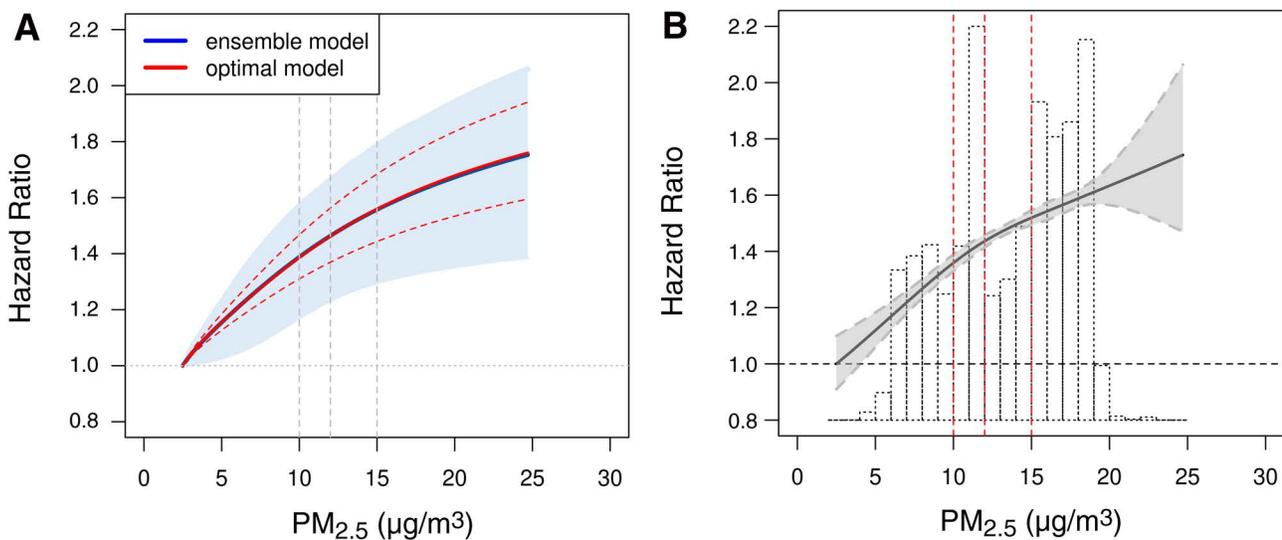
The shape of the concentration–response function using natural splines and SCHIF is shown in Figure 7 for natural-cause mortality and in Figure P32 for cardiovascular and respiratory mortality. For natural-cause and cause-specific mortality, both splines and SCHIF suggest a supralinear association, consistent with the main ELAPSE model (Figures 4 and 5).

**Associations with Large Spatial Scale Ozone** O<sub>3</sub> exposure from our main model at a 100 × 100 m spatial scale was moderately correlated with the two larger spatial scale estimates we evaluated as an additional analysis. The median (min–max) of cohort-specific correlations with the DEHM warm-season estimate at a 50 × 50 km scale was 0.39 (0.02–0.82). The median (min–max) of cohort-specific correlations with the MACC annual average (warm season

**Table 11.** Hazard Ratios for Associations Between PM<sub>2.5</sub> from ELAPSE and MAPLE Exposure Model and Natural-Cause and Cause-Specific Mortality<sup>a</sup>

	Natural Cause HR (95% CI)	Cardio-vascular Disease HR (95% CI)	Ischemic Heart Disease HR (95% CI)	Cerebro-vascular Disease HR (95% CI)	Respiratory HR (95% CI)	COPD HR (95% CI)	Diabetes HR (95% CI)	Cardio-metabolic HR (95% CI)
PM <sub>2.5</sub> ELAPSE	1.134 (1.109, 1.159)	1.137 (1.097, 1.178)	1.113 (1.057, 1.171)	1.130 (1.050, 1.217)	1.050 (0.957, 1.152)	1.134 (1.004, 1.281)	1.312 (1.140, 1.509)	1.147 (1.108, 1.188)
PM <sub>2.5</sub> MAPLE	1.117 (1.094, 1.139)	1.105 (1.068, 1.143)	1.132 (1.077, 1.189)	1.080 (1.008, 1.158)	1.088 (0.998, 1.185)	1.062 (0.949, 1.189)	1.244 (1.092, 1.417)	1.112 (1.076, 1.150)

<sup>a</sup> N = 323,782; HR (95% CI) presented for 5 µg/m<sup>3</sup> increment in PM<sub>2.5</sub>. Main model adjusted for cohort ID, age, sex, year of baseline visit, smoking (status, duration, intensity, intensity squared), BMI, marital status, employment status, and 2001 neighborhood-level mean income.



**Figure 7.** SCHIF (A) and natural spline function (B, 3df) for natural-cause mortality and MAPLE PM<sub>2.5</sub>. Vertical dotted lines are air quality limit and guideline values. Hazard ratios are expressed relative to minimum exposure. Histogram shows exposure distribution.

not available) estimate at a 10 × 10 km scale was 0.37 (−0.35 −0.70) (Appendix 2, Table P23).

For natural-cause mortality and cardiovascular mortality, associations with larger spatial scale O<sub>3</sub> exposure estimates were negative and similar in magnitude compared with our main fine spatial scale exposure (Table P24). For respiratory mortality, we found a nonsignificant positive effect estimate for DEHM at 50 × 50 km.

### Summary of the Mortality Findings in the Pooled Cohort

Virtually all subjects had 2010 PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below the EU limit values. More than 50,000 and 25,000 subjects had a residential PM<sub>2.5</sub> exposure below the U.S. EPA NAAQS (12 µg/m<sup>3</sup>) and the WHO air quality guideline (10 µg/m<sup>3</sup>), respectively. We found significant positive associations between PM<sub>2.5</sub>, NO<sub>2</sub>, and BC and natural-cause, respiratory, cardiovascular, and

diabetes mortality. In our main model, the HRs (95% CI) were 1.13 (CI = 1.11, 1.16) for an increase of 5 µg/m<sup>3</sup> PM<sub>2.5</sub>, 1.09 (CI = 1.07, 1.10) for an increase of 10 µg/m<sup>3</sup> NO<sub>2</sub>, and 1.08 (CI = 1.06, 1.10) for an increase of 0.5 × 10<sup>-5</sup>/m BC for natural-cause mortality. Associations with O<sub>3</sub> were negative, both in the fine spatial scale main ELAPSE model and in the large spatial scale exposure models. For PM<sub>2.5</sub>, NO<sub>2</sub>, and BC, we generally observed a supralinear association with steeper slopes at low exposures and no evidence of a concentration below which no association was found. Subset analyses further confirmed that these associations remained at low levels: below 10 µg/m<sup>3</sup> for PM<sub>2.5</sub> and below 20 µg/m<sup>3</sup> for NO<sub>2</sub>. HRs were similar to the full cohort HRs for subjects with exposures below the EU limit values for PM<sub>2.5</sub> and NO<sub>2</sub>, the U.S. NAAQS for PM<sub>2.5</sub>, and the WHO Air Quality Guidelines for PM<sub>2.5</sub> and NO<sub>2</sub>. The mortality associations were robust to alternative specifications of exposure, including different time-varying exposure models and PM<sub>2.5</sub> from the MAPLE study. Time-varying exposure natural spline analyses confirmed associations at low pollution levels. HRs in two-pollutant models were attenuated but remained elevated and statistically significant for PM<sub>2.5</sub> and NO<sub>2</sub>. Associations with O<sub>3</sub> attenuated but remained negative in two-pollutant models with NO<sub>2</sub>, BC, and PM<sub>2.5</sub>.

#### **Incidence of Acute Coronary Heart Disease, Stroke, Asthma, COPD, and Lung Cancer**

**Descriptive Data** Incidence analyses have been conducted for a more limited number of cohorts than for the mortality analyses. Table 12 presents the included cohorts per endpoint, total number of participants, number of events and person years, and key population characteristics for the pooled population. Distributions per cohort have been shown in Table 4 and in detail in Appendix 2, section P1. The numbers per cohort differ slightly from Table 4 because of different confounders and exclusion of prevalent cases of cancer, coronary, or stroke events.

#### **Main Analysis of Association with Incidence of Disease**

**Single-Pollutant Models** Table 13 shows the HRs for linear associations between pollutants and incidence of acute coronary heart disease, stroke, lung cancer, asthma, and COPD in single-pollutant models. Positive associations were found for PM<sub>2.5</sub>, NO<sub>2</sub>, and BC with all outcomes. Associations with O<sub>3</sub> were negative, some of them significant. For acute coronary heart disease, associations were stronger and statistically significant for NO<sub>2</sub> compared with PM<sub>2.5</sub>. PM<sub>2.5</sub> was more consistently associated with lung cancer than NO<sub>2</sub> and BC. Stroke, asthma, and COPD were associated with PM<sub>2.5</sub>, NO<sub>2</sub>, and BC. Effect estimates were high for asthma and COPD incidence.

**Concentration-Response Functions** Natural spline functions showed a sublinear association between NO<sub>2</sub> and acute coronary heart disease with little indication of an association below about 20 µg/m<sup>3</sup> (Figure 8). For PM<sub>2.5</sub>, the slope was steeper at low concentrations. For stroke incidence, supralinear functions were found for PM<sub>2.5</sub>, BC, and NO<sub>2</sub> (Figure 9): the positive slopes were steeper at low concentrations, with slopes leveling off above 15 µg/m<sup>3</sup> for PM<sub>2.5</sub> and 40 µg/m<sup>3</sup> for NO<sub>2</sub>. Figure 10 shows for lung cancer incidence a positive slope for PM<sub>2.5</sub> at low concentrations, leveling off at about 18 µg/m<sup>3</sup>. SCHIF for associations between PM<sub>2.5</sub> and NO<sub>2</sub>, and acute coronary heart disease, stroke, and lung cancer incidence, generally showed similar patterns (Figure 11). Collectively, these analyses suggested that the association with these three endpoints extends to the lowest levels of PM<sub>2.5</sub> and NO<sub>2</sub>, with the exception of acute coronary heart disease and NO<sub>2</sub>. In Appendix 2, Figures P33 to P38 document that spline functions with two and four degrees of freedom show similar patterns. Penalized splines generally showed similar patterns as well (Appendix 2, Figures P39–P41).

Figures P42 to P45 show the spline (three degrees of freedom) and SCHIF functions for asthma and COPD incidence. Spline functions suggest supralinear associations, with steeper slopes at lower concentrations, leveling off at higher concentration levels. SCHIF functions were smoother and suggested linear to supralinear functions.

**Table 12.** Description of the Pooled Study Populations per Endpoint

Variable / Cohorts	Acute Coronary Heart Disease (all except VHM & PP and E3N)	Stroke (all except VHM & PP and E3N)	Lung Cancer (all, except KORA)	Asthma (CEANS, DCH, DNC)	COPD (CEANS, DCH, DNC)
Total participants	157,265	157,265	368,697	105,906	106,094
Main model participants	137,148	137,148	307,550	98,326	98,508
Person-years at risk	2,168,806	2,168,806	5,561,379	1,634,458	1,637,916
Number of events	10,071	6,950	3,956	1,965	4,928
Age, years (mean $\pm$ SD)	54 $\pm$ 9	54 $\pm$ 9	48.3 $\pm$ 13.4	55.8 $\pm$ 7.5	55.8 $\pm$ 7.5
Women (%)	67	67	66	66	66
BMI ( $\geq$ 25 kg/m <sup>2</sup> ) (%)	50	50	43	48	48
Current smokers (%)	31	31	24	33	33
Former smokers (%)	30	30	18	30	30
Cigarettes/day (mean $\pm$ SD) <sup>a</sup>	9 $\pm$ 11	9 $\pm$ 11	15.2 $\pm$ 8.9	9.2 $\pm$ 10.4	9.2 $\pm$ 10.4
Years of smoking (mean $\pm$ SD) <sup>a</sup>	16 $\pm$ 16	16 $\pm$ 16	25.3 $\pm$ 13.1	17.1 $\pm$ 16.5	17.0 $\pm$ 16.5
Not employed (%) <sup>b</sup>	28	28	29	24	24
Married/living with partner (%)	72	72	72	71	71
Primary school or less (%)	14	14	NA	14	14

<sup>a</sup> For current and former smokers; for lung cancer for current smokers.

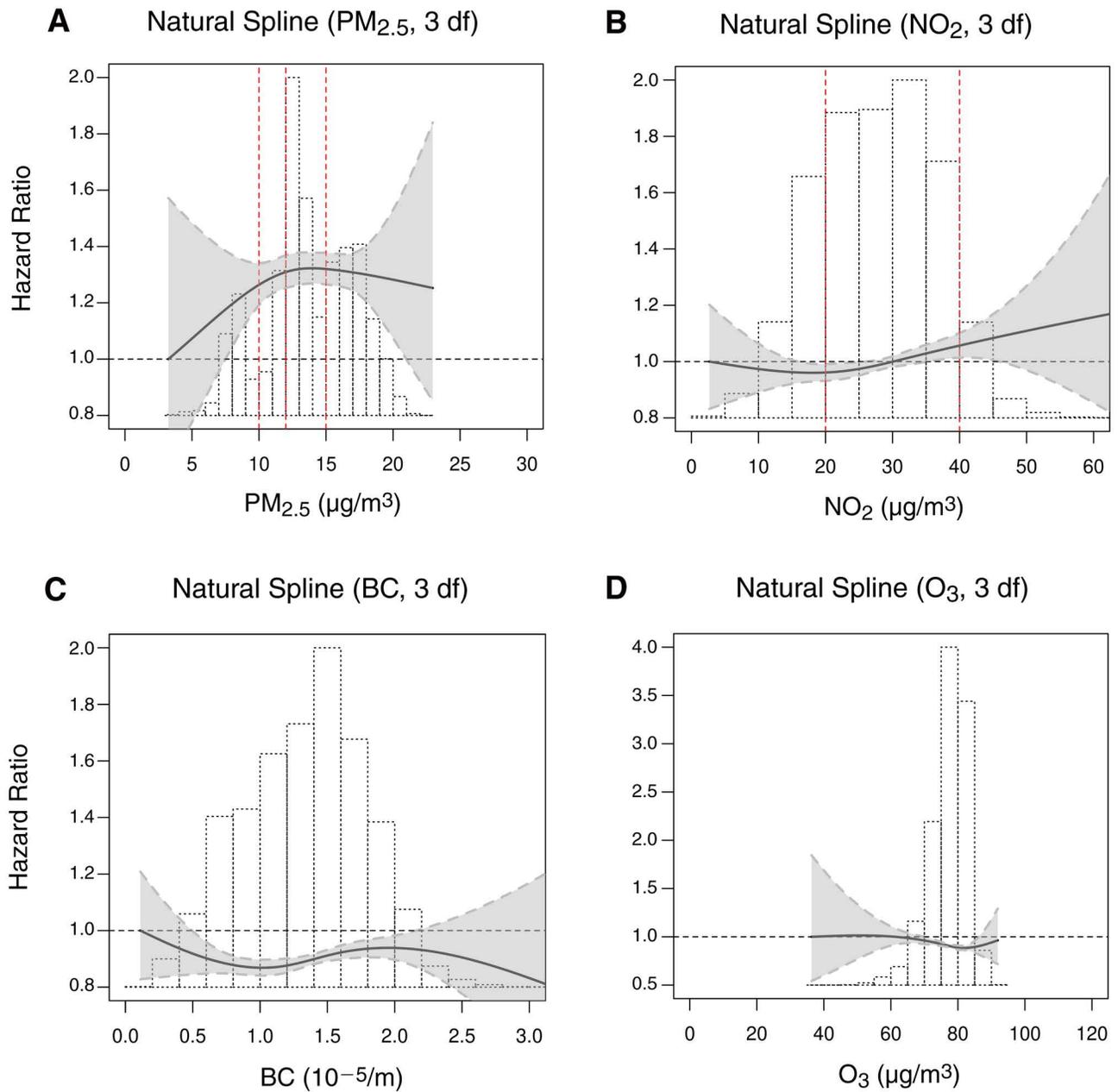
<sup>b</sup> Includes unemployed, home workers, and retired subjects.

**Table 13.** Hazard Ratios for Associations Between Air Pollution and Incidence of Acute Coronary Heart Disease, Stroke, Lung Cancer, Asthma, and COPD<sup>a, b</sup>

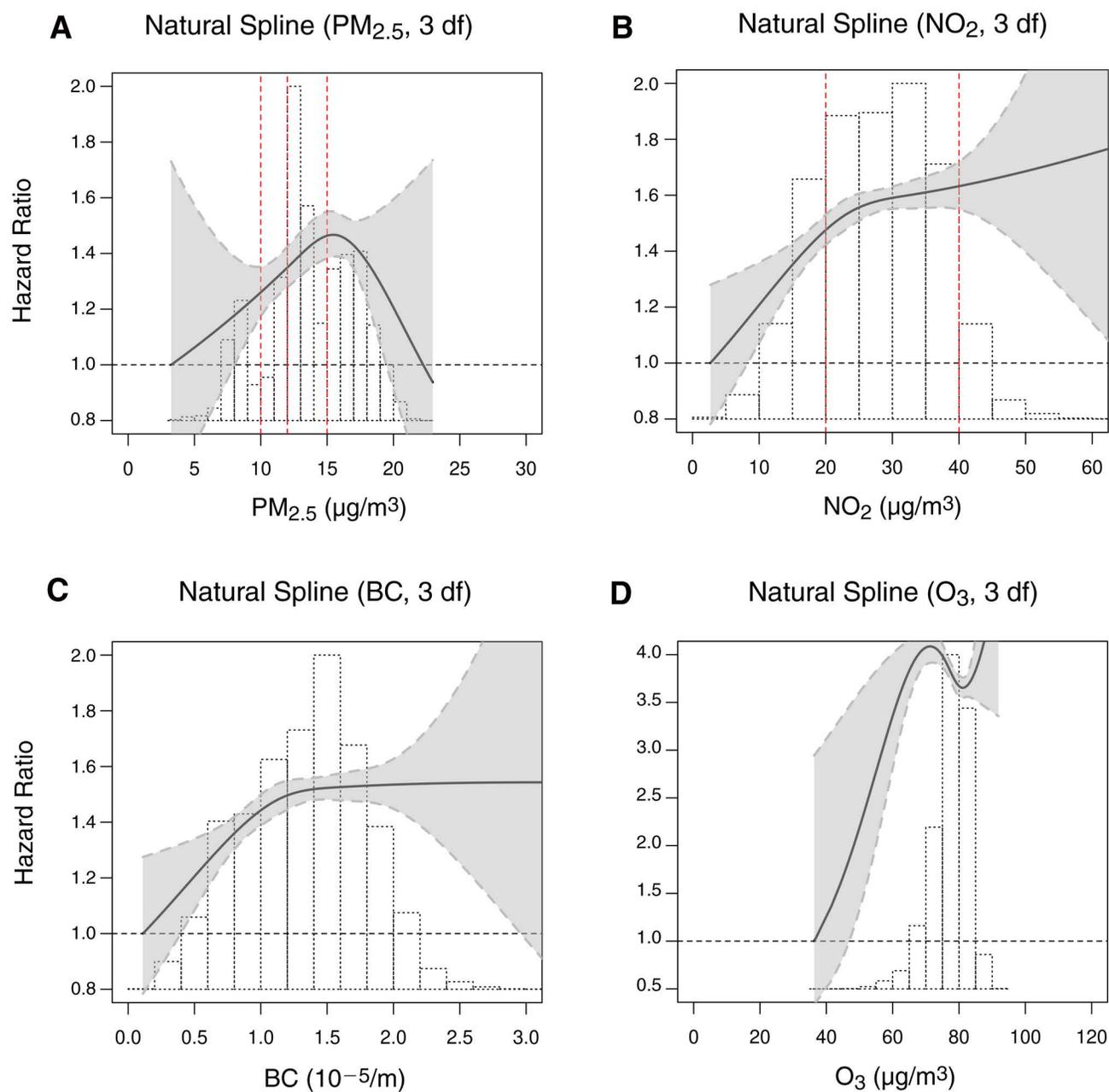
	Acute Coronary Heart Disease HR (95% CI)	Stroke HR (95% CI)	Lung Cancer HR (95% CI)	Asthma HR (95% CI)	COPD HR (95% CI)
PM <sub>2.5</sub>	1.02 (0.95, 1.10)	1.10 (1.01, 1.21)	1.13 (1.05, 1.23)	1.22 (1.04, 1.43)	1.17 (1.06, 1.29)
NO <sub>2</sub>	1.04 (1.01, 1.07)	1.08 (1.04, 1.12)	1.02 (0.97, 1.07)	1.17 (1.10, 1.25)	1.11 (1.06, 1.16)
BC	1.02 (0.99, 1.06)	1.06 (1.02, 1.10)	1.02 (0.97, 1.07)	1.15 (1.08, 1.23)	1.11 (1.06, 1.15)
O <sub>3</sub>	0.94 (0.90, 0.98)	0.96 (0.91, 1.01)	0.95 (0.89, 1.02)	0.90 (0.81, 0.99)	0.99 (0.93, 1.05)

<sup>a</sup> Acute coronary heart disease and stroke,  $N = 137,148$ ; lung cancer,  $N = 307,550$ ; asthma,  $N = 98,326$ ; COPD,  $N = 98,508$ . PM<sub>2.5</sub> per 5  $\mu\text{g}/\text{m}^3$ , NO<sub>2</sub> per 10  $\mu\text{g}/\text{m}^3$ , BC per  $0.5 \times 10^{-5}/\text{m}$ , O<sub>3</sub> per 10  $\mu\text{g}/\text{m}^3$ .

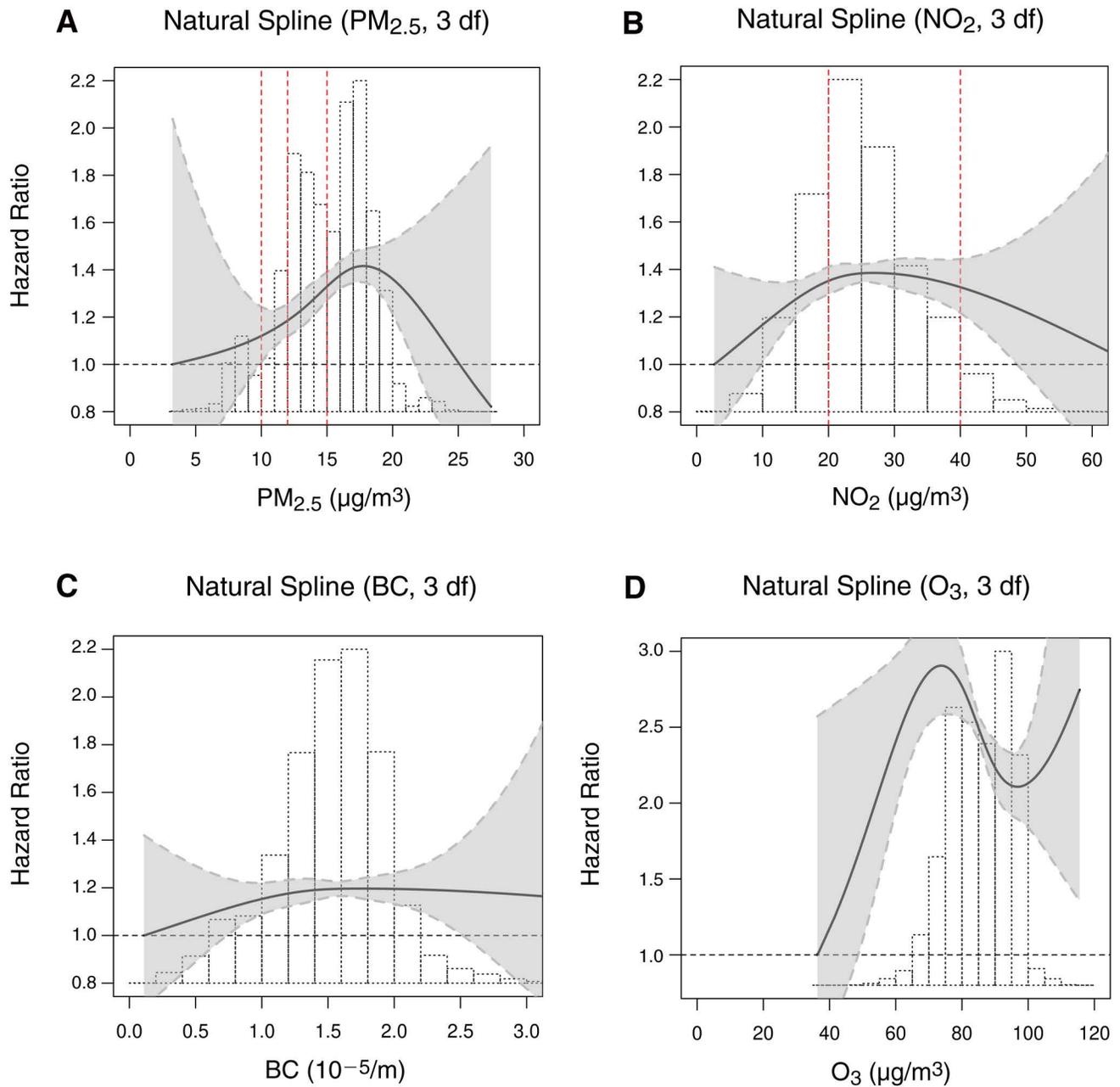
<sup>b</sup> Hazard ratio adjusted for cohort ID, age, sex, year of baseline visit, smoking (status, duration, intensity, intensity squared), BMI category, marital status, employment status, and 2001 neighborhood-level mean income (acute coronary heart disease, stroke, asthma, and COPD incidence additionally adjusted for educational status).



**Figure 8.** Natural cubic splines (three degrees of freedom) for associations between air pollution exposure and acute coronary heart disease incidence. Red dotted lines are air quality limit and guideline values. Hazard ratios are expressed relative to minimum exposure. Histograms reflect exposure distributions. (Adapted from Wolf et al. [In press])

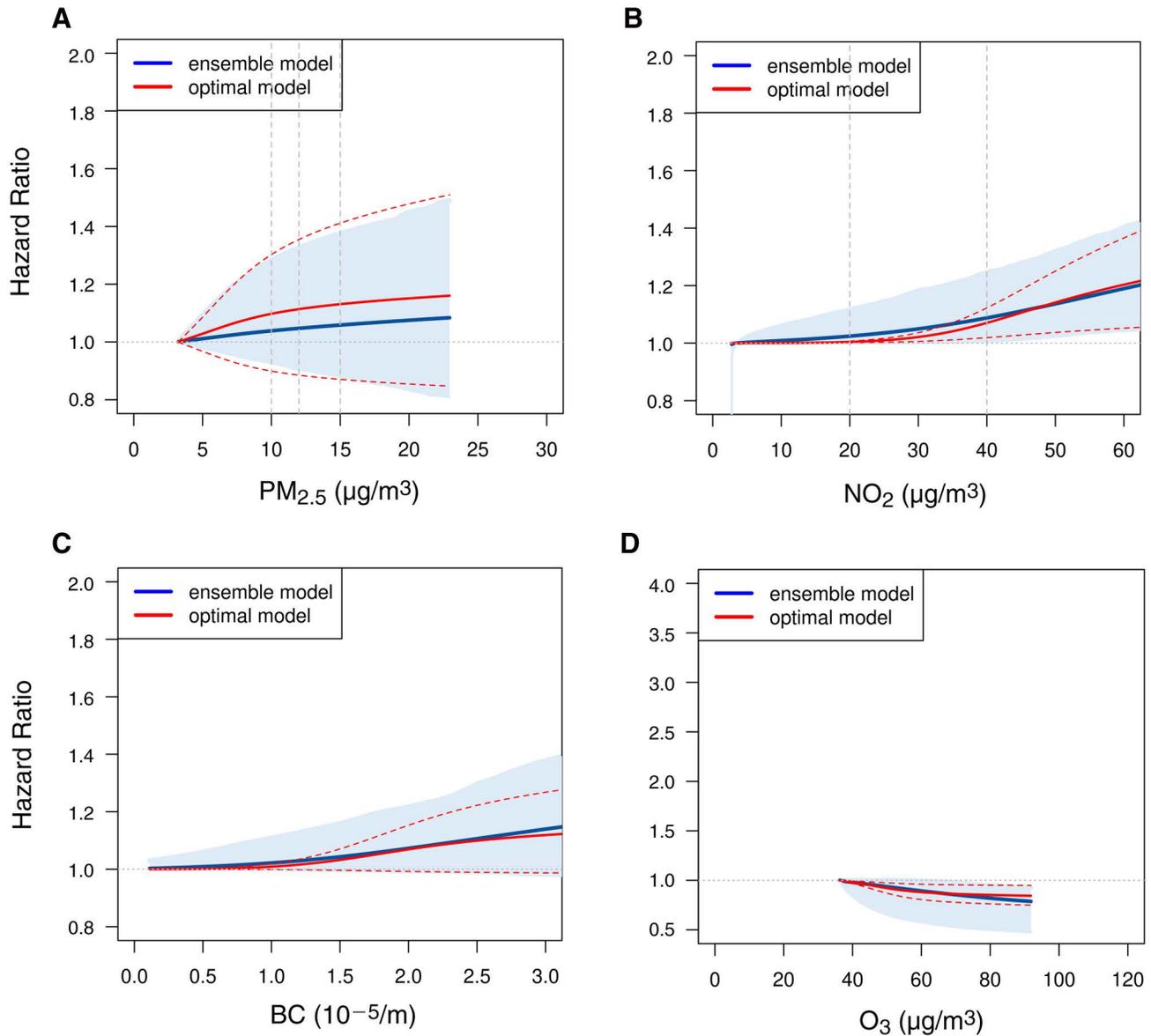


**Figure 9.** Natural cubic splines (three degrees of freedom) for associations between air pollution exposure and stroke incidence. Red dotted lines are air quality limit and guideline values. Hazard ratios expressed relative to minimum exposure. Histograms reflect exposure distributions. (Adapted from Wolf et al. [In press])



**Figure 10. Natural cubic splines (three degrees of freedom) for associations between air pollution exposure and lung cancer incidence.** Red dotted lines are air quality limit and guideline values. Hazard ratios expressed relative to minimum exposure. Histograms reflect exposure distributions. (Adapted from Hvitfeldt et al 2021; Creative Commons license CC BY-NC-ND 4.0)

## Acute Coronary Heart Disease Incidence



**Figure 11. SCHIF for associations between air pollution exposure and acute coronary heart disease, stroke, and lung cancer incidence.** Vertical dotted lines are air quality limit and guideline values. Hazard ratios expressed relative to minimum exposure. (Adapted from Hvitfeldt et al 2021 [Creative Commons license CC BY-NC-ND 4.0] and Wolf et al. [In press])

(Figure 11 continues next page.)

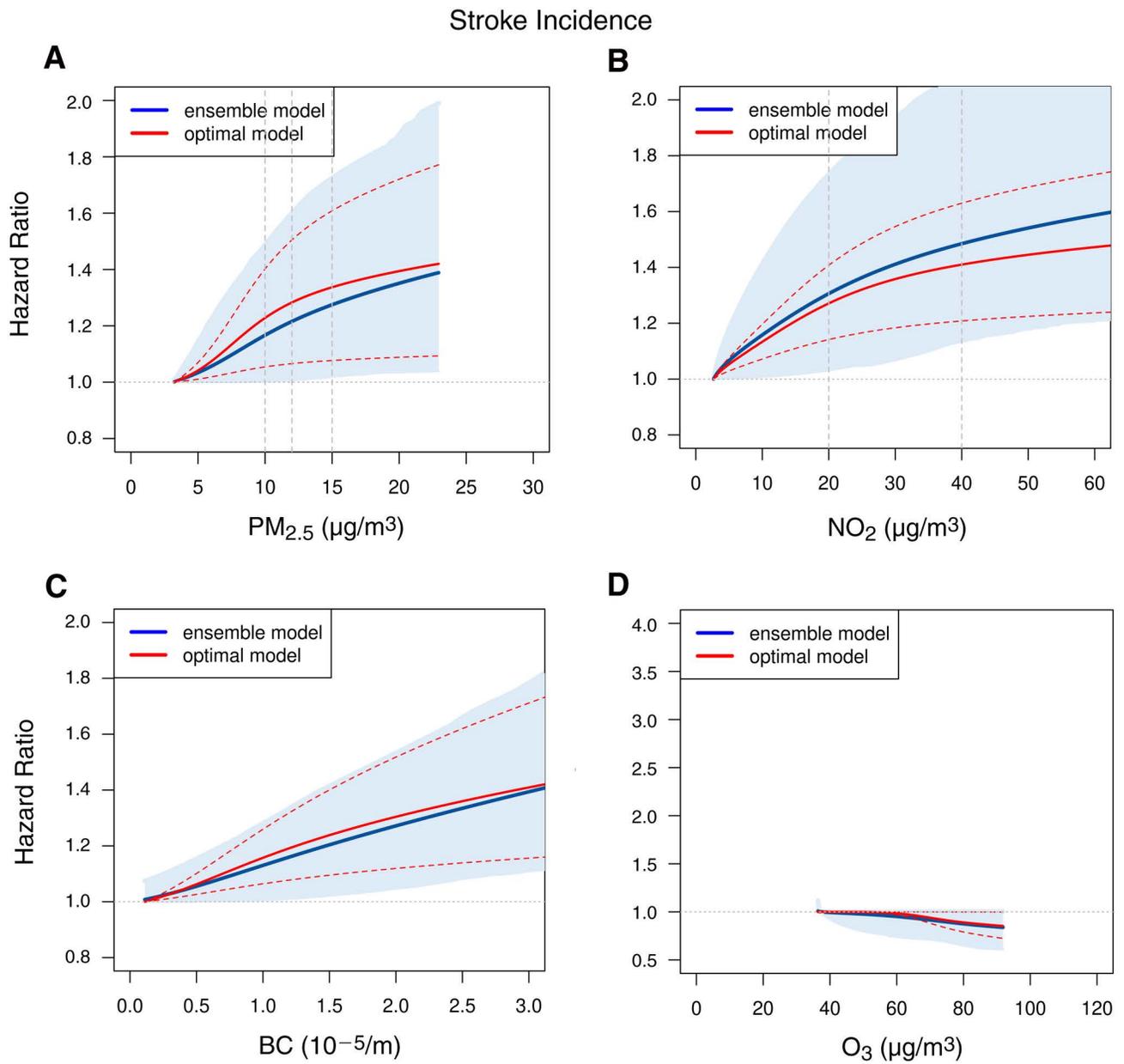


Figure 11 (Continued).

Lung Cancer Incidence

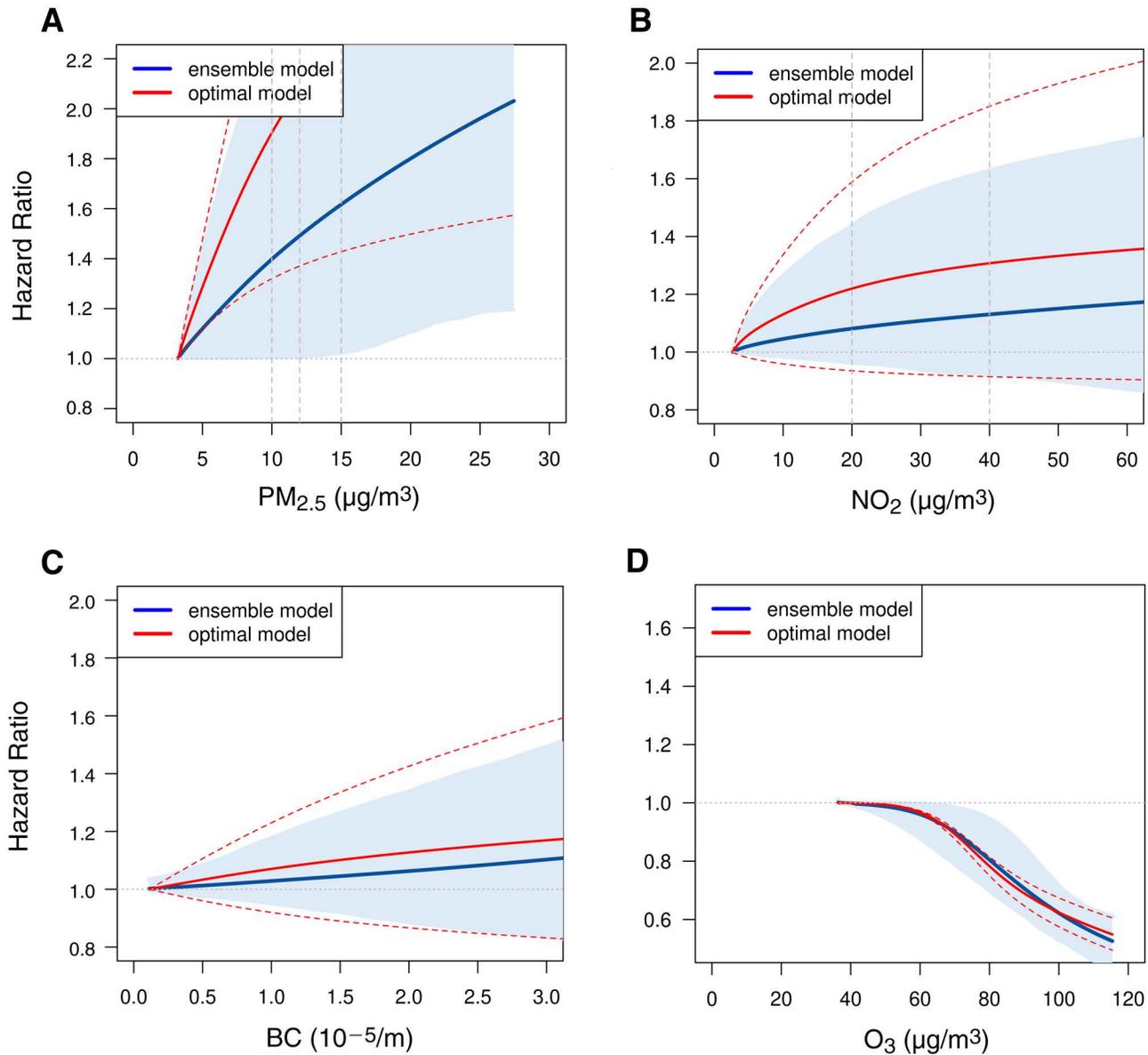


Figure 11 (Continued).

### Subset Analysis

Table 14 shows the results of the subset analyses for acute coronary heart disease and stroke incidence. For PM<sub>2.5</sub>, effect estimates for incidence of acute coronary heart disease were larger at low concentrations, though nonsignificant. For NO<sub>2</sub> effect estimates were inconsistent below about 30 µg/m<sup>3</sup>. For stroke incidence, PM<sub>2.5</sub> effect estimates were larger for subsets below 15 and 12 µg/m<sup>3</sup> compared with the full range HR, but not for the below 10 µg/m<sup>3</sup> subset. For NO<sub>2</sub> and BC, effect estimates slightly increased in the subsets with lowest concentrations. All cohorts contributed to the analysis of HRs below 15 µg/m<sup>3</sup> PM<sub>2.5</sub>, though the Dutch and German cohort contributed little (Appendix 2, Table P1). The Stockholm and Danish cohorts contributed to the analysis below 10 and 12 µg/m<sup>3</sup> PM<sub>2.5</sub>, but the analysis below 10 µg/m<sup>3</sup> PM<sub>2.5</sub> is dominated by the Stockholm cohort. All cohorts except HNR and the Dutch cohorts contributed to the analysis of NO<sub>2</sub> concentrations below 20 µg/m<sup>3</sup>.

Table 15 shows the results of the subset analyses for lung cancer incidence. For PM<sub>2.5</sub>, effect estimates were similar in subsets with wide CIs below 12 µg/m<sup>3</sup>. For NO<sub>2</sub>, associations became somewhat larger at low concentrations, with wider CIs. We note that the subsets contained different cohorts especially for PM<sub>2.5</sub> where there were large differences between cohorts (Figures 1 and 2). All cohorts contributed to the analysis of HRs below 15 µg/m<sup>3</sup> PM<sub>2.5</sub>, though the Dutch and German cohort contributed little (Table P1). The Stockholm, Danish, French, and Austrian cohorts contributed to the analysis below 10 and 12 µg/m<sup>3</sup> PM<sub>2.5</sub>, but the analysis below 10 µg/m<sup>3</sup> PM<sub>2.5</sub> was dominated by the Stockholm cohort. All cohorts except HNR and the Dutch cohorts contributed to the analysis of NO<sub>2</sub> concentrations below 20 µg/m<sup>3</sup>.

Appendix Table P25 shows results of subset analyses for asthma and COPD incidence. For PM<sub>2.5</sub>, effect estimates were larger at low concentrations, but with wide CI below 12 µg/m<sup>3</sup>. For NO<sub>2</sub> effect estimates remained consistent below 30 µg/m<sup>3</sup> and possibly 20 µg/m<sup>3</sup> (asthma). All cohorts contributed to the analysis below 10 and 12 µg/m<sup>3</sup> PM<sub>2.5</sub>, but the analysis below 10 µg/m<sup>3</sup> PM<sub>2.5</sub> was dominated by the Stockholm cohort. All cohorts contributed to the analysis of NO<sub>2</sub> concentrations below 20 µg/m<sup>3</sup>.

In Appendix 2, Table P26 shows results of threshold analyses for all incidence outcomes. For acute coronary heart disease and stroke, model performance was almost identical for the linear model and the three threshold models for PM<sub>2.5</sub>. For NO<sub>2</sub>, model performance was almost identical for the linear model and the first two threshold models but slightly worse for acute coronary heart disease than for a threshold at 20 µg/m<sup>3</sup>. For lung cancer

incidence, model performance was almost identical for the linear model and the two first threshold models for PM<sub>2.5</sub>, but slightly better than for a model with a threshold of 10 µg/m<sup>3</sup>. For NO<sub>2</sub>, model performance was almost identical for the linear model and all threshold models. For asthma and COPD admissions, model performance was almost identical for the linear model and the three threshold models for PM<sub>2.5</sub>. For NO<sub>2</sub>, model performance was almost identical for the linear model and the first two threshold models but better (lower AIC) than for a threshold at 20 µg/m<sup>3</sup>.

**Two-Pollutant Models** In Appendix 2, Tables P27–P31 show the results of two-pollutant analyses. For acute coronary heart disease, HRs for NO<sub>2</sub> remained significantly positive after adjustment for PM<sub>2.5</sub>, but HRs were attenuated after adjustment for O<sub>3</sub> (Table P27). The negative effect estimate for O<sub>3</sub> was attenuated and became insignificant after adjustment for NO<sub>2</sub>. For stroke, the HR for PM<sub>2.5</sub> was attenuated toward essentially unity after adjustment for NO<sub>2</sub> and BC (Table P28). HRs for NO<sub>2</sub> and BC remained stable after adjustment for PM<sub>2.5</sub> and O<sub>3</sub>. Two-pollutant models of NO<sub>2</sub> and BC are difficult to interpret because of the high mutual correlation (the width of the CIs is almost doubled). For lung cancer incidence, HRs for PM<sub>2.5</sub> were not affected by adjustment for NO<sub>2</sub>, BC, and O<sub>3</sub>. For the other three pollutants, HRs were essentially unity in two-pollutant models (Table P29). For asthma and COPD hospital admissions, the HR for PM<sub>2.5</sub> was attenuated toward essentially unity after adjustment for NO<sub>2</sub> and BC (Tables P30 and 31). HRs for NO<sub>2</sub> and BC remained stable after adjustment for PM<sub>2.5</sub> and O<sub>3</sub>. Significant positive associations between O<sub>3</sub> and COPD hospital admission were found after adjustment for NO<sub>2</sub> and BC.

### Additional Analyses

**Time-Varying and Baseline Exposure Models** Tables P32 and P33 show the HRs from models with exposure at baseline and the time-varying analysis instead of the 2010 exposure in the main model.

For acute coronary heart disease, HRs for exposure at baseline were similar to the 2010 exposure. For the other outcomes, HRs were substantially smaller for PM<sub>2.5</sub> exposure and slightly smaller for NO<sub>2</sub> and BC at baseline, though still statistically significant for stroke and lung cancer. CIs were smaller for the baseline exposure especially with back-extrapolation with the ratio method. This can be explained by the higher exposure contrast in past exposures (Appendix 2, Figure P21). When the same health response is related to a larger contrast (on the absolute scale), a lower slope is observed. For acute coronary

**Table 14.** Hazard Ratios for Associations Between Air Pollution and Acute Coronary Heart Disease and Stroke Incidence in Subset Analysis<sup>a</sup>

Pollutant	Subset	N	Acute Coronary Heart Disease HR (95% CI)	Stroke HR (95% CI)
<b>PM<sub>2.5</sub></b>				
	Full dataset	137,148	1.02 (0.95, 1.10)	1.10 (1.01, 1.21)
	<25 µg/m <sup>3</sup>	137,148	1.02 (0.95, 1.10)	1.10 (1.01, 1.21)
	<20 µg/m <sup>3</sup>	135,564	1.02 (0.94, 1.10)	1.10 (1.01, 1.21)
	<15 µg/m <sup>3</sup>	85,525	1.06 (0.94, 1.20)	1.20 (1.04, 1.38)
	<12 µg/m <sup>3</sup>	34,875	1.15 (0.88, 1.50)	1.25 (0.90, 1.72)
	<10 µg/m <sup>3</sup>	20,271	1.15 (0.80, 1.65)	0.91 (0.60, 1.38)
<b>NO<sub>2</sub></b>				
	Full dataset	137,148	1.04 (1.01, 1.07)	1.08 (1.04, 1.12)
	<40 µg/m <sup>3</sup>	127,332	1.04 (1.00, 1.07)	1.08 (1.04, 1.13)
	<30 µg/m <sup>3</sup>	79,177	1.01 (0.95, 1.07)	1.17 (1.09, 1.25)
	<20 µg/m <sup>3</sup>	29,468	1.05 (0.92, 1.21)	1.16 (0.97, 1.40)
<b>BC</b>				
	Full dataset	137,148	1.02 (0.99, 1.06)	1.06 (1.02, 1.10)
	<3 × 10 <sup>-5</sup> /m	137,141	1.02 (0.99, 1.06)	1.06 (1.02, 1.10)
	<2.5 × 10 <sup>-5</sup> /m	136,723	1.02 (0.99, 1.05)	1.06 (1.02, 1.10)
	<2 × 10 <sup>-5</sup> /m	128,852	1.02 (0.99, 1.06)	1.06 (1.02, 1.11)
	<1.5 × 10 <sup>-5</sup> /m	84,629	1.02 (0.97, 1.08)	1.11 (1.04, 1.19)
	<1 × 10 <sup>-5</sup> /m	34,172	0.88 (0.76, 1.01)	1.17 (0.98, 1.40)
	<0.5 × 10 <sup>-5</sup> /m	4,852	0.76 (0.32, 1.79)	1.73 (0.48, 6.23)
<b>O<sub>3</sub></b>				
	Full dataset	137,148	0.94 (0.90, 0.98)	0.96 (0.91, 1.01)
	<120 µg/m <sup>3</sup>	137,148	0.94 (0.90, 0.98)	0.96 (0.91, 1.01)
	<100 µg/m <sup>3</sup>	137,148	0.94 (0.90, 0.98)	0.96 (0.91, 1.01)
	<80 µg/m <sup>3</sup>	89,272	0.93 (0.88, 0.98)	0.98 (0.91, 1.05)
	<60 µg/m <sup>3</sup>	1,601	1.64 (0.81, 3.31)	0.29 (0.09, 0.96) <sup>b</sup>

<sup>a</sup> HR (95% CI) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC, 0.5×10<sup>-5</sup>/m; O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for cohort ID, age, sex, year of baseline visit, smoking (status, duration, intensity, intensity squared), BMI, marital status, employment status, educational status, and 2001 neighborhood-level mean income.

<sup>b</sup> Unstable estimate.

**Table 15.** Hazard Ratios for Associations Between Air Pollution and Lung Cancer Incidence in Subset Analysis<sup>a</sup>

Pollutant	Subset	N	HR (95% CI)
<b>PM<sub>2.5</sub></b>			
	Full dataset	307,550	1.13 (1.05, 1.23)
	<25 µg/m <sup>3</sup>	307,524	1.13 (1.05, 1.23)
	<20 µg/m <sup>3</sup>	299,514	1.15 (1.06, 1.25)
	<15 µg/m <sup>3</sup>	145,078	1.09 (0.90, 1.30)
	<12 µg/m <sup>3</sup>	49,834	1.21 (0.75, 1.96)
	<10 µg/m <sup>3</sup>	23,554	2.01 (0.80, 5.04)
<b>NO<sub>2</sub></b>			
	Full dataset	307,550	1.02 (0.97, 1.07)
	<40 µg/m <sup>3</sup>	233,303	1.04 (0.98, 1.10)
	<30 µg/m <sup>3</sup>	247,039	1.05 (0.96, 1.13)
	<20 µg/m <sup>3</sup>	83,229	1.06 (0.84, 1.33)
<b>BC</b>			
	Full dataset	307,550	1.02 (0.97, 1.07)
	<3 × 10 <sup>-5</sup> /m	306,984	1.02 (0.97, 1.07)
	<2.5 × 10 <sup>-5</sup> /m	303,132	1.02 (0.97, 1.08)
	<2 × 10 <sup>-5</sup> /m	280,462	1.03 (0.97, 1.09)
	<1.5 × 10 <sup>-5</sup> /m	134,046	1.03 (0.92, 1.15)
	<1 × 10 <sup>-5</sup> /m	33,435	1.32 (0.93, 1.86)
	<0.5 × 10 <sup>-5</sup> /m	4,730	4.02 (0.12, 129.09) <sup>b</sup>
<b>O<sub>3</sub></b>			
	Full dataset	307,550	0.95 (0.89, 1.02)
	<120 µg/m <sup>3</sup>	307,550	0.95 (0.89, 1.02)
	<100 µg/m <sup>3</sup>	302,968	0.95 (0.89, 1.02)
	<80 µg/m <sup>3</sup>	93,882	1.08 (0.96, 1.21)
	<60 µg/m <sup>3</sup>	1,664	0.38 (0.06, 2.34)

<sup>a</sup> HR (95% CI) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC, 0.5×10<sup>-5</sup>/m; O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for cohort ID, age, sex, year of baseline visit, smoking (status, duration, intensity, intensity squared), BMI, marital status, employment status, and 2001 neighborhood-level mean income.

<sup>b</sup> Unstable estimate.

heart disease, HRs were similar for NO<sub>2</sub> in the time-varying analyses and inconsistent dependent on the adjustment for time trends with 1- or 5-year strata for PM<sub>2.5</sub> and BC. For stroke, HRs were larger for the time-varying analyses. For lung cancer, HRs for PM<sub>2.5</sub> were similar for the time-varying analyses with 1-year strata and much smaller with 5-year strata. We note that the time-varying analyses for lung cancer are less interpretable given the latency period for lung cancer. For asthma and COPD, HRs from the time-varying analyses were either smaller or larger (both significant), depending on

the use of 1- or 5-year strata. Overall, the baseline and time-varying exposure analyses support the pattern of associations with the main model, but the magnitude of the association may differ substantially depending on the specification of the exposure model.

The time-varying exposure natural splines further documented that associations were found at low pollution levels (Appendix 2, Figures P46 to P50). The time-varying exposure analysis is not based upon the 2010 exposure metric, which represents the relatively low pollution value at the end of follow-up of most cohorts.

### **Alternative Adjustment for Differences Between Sub-Cohorts**

Tables P34–36 and Figures P51 to P55 in Appendix 2 show the results of the sensitivity analyses for acute coronary heart disease, stroke, and lung cancer incidence with respect to alternative adjustment for differences between cohorts (Figures P51–53), adjustment for additional confounders (Tables P34–P36), and the impact of dropping individual cohort (Figures P54 and P55). Associations were robust to alternative adjustment for differences between cohorts and adjustment for additional confounders. Similar to the mortality analyses, the alternative methods to adjust for differences between cohorts had little impact on effect estimates and CIs. Not adjusting for cohort at all did affect the HR, though in different directions depending on pollutant and outcome (Figures P51–P53). In the confounder sensitivity analysis, we did observe changes in HR when specific cohorts were left out, especially in the lung cancer incidence analyses, where dropping the large VHM&PP cohort reduced the association with  $PM_{2.5}$  to null with our main confounder model. An analysis with dropping cohorts one by one, showed that HRs decreased when dropping the VHM&PP cohort and increased when dropping the E3N cohort (Figure P54). An analysis per individual cohort, showed that in all cohorts except E3N a higher risk for lung cancer in association with  $PM_{2.5}$  was found (Figure P55). For stroke and acute coronary heart disease, fewer confounder sensitivity analyses were conducted than in the lung cancer and mortality analyses, because the main model already included education and smoking intensity for ever-smokers (due to the VHM&PP cohort not having data on these outcomes).

**Comparison with Local Exposure Models** Table P37 in Appendix 2 compares HRs estimated with the study-area specific ESCAPE models and the ELAPSE exposure model. For acute coronary heart disease, associations for the ESCAPE models were slightly weaker for  $NO_2$  and BC and stronger for  $PM_{2.5}$ . For stroke, associations for the ESCAPE models were weaker for  $NO_2$  and BC and stronger for  $PM_{2.5}$ . The HR for  $PM_{2.5}$  ELAPSE model was lower than in the full population for stroke. For lung cancer, association with the ESCAPE models were slightly stronger. For asthma and COPD, a comparison is not informative, as one of the three cohorts included did not have ESCAPE exposures at all (DNC) and for another cohort (DCH) only the Copenhagen part of the cohort had ESCAPE exposures available.

**Adjustment for Traffic Noise** Table P38 shows the results of additional adjustment for traffic noise for acute coronary heart disease and stroke. Effect estimates did not change materially after adjustment, consistent with the

moderate correlation between air pollution and road-traffic noise in the cohorts.

**Effect Modification** In Appendix 2, Figures P56 to 60 show the results of effect modification for the five incidence outcomes. For acute coronary heart disease, HRs were consistently higher for the elderly compared with the nonelderly for both  $NO_2$  and  $PM_{2.5}$  (Figure P56). There was also a tendency toward higher effect estimates in subjects with low education. Effect estimates were similar across smoking status and BMI categories. For stroke, HRs were higher in the elderly for  $PM_{2.5}$ , but not  $NO_2$  (Figure P57). HRs did not differ consistently for the other variables. For lung cancer, there were no important differences in HRs for any of the variables. HRs for never smokers were identical to HRs in current smokers (Figure P58). For asthma and COPD there was a tendency toward higher effect estimates in subjects with low education. For COPD, HRs were higher in the nonelderly and in current smokers, a pattern not observed for asthma.

### **Associations with $PM_{2.5}$ Exposure from MAPLE**

Descriptive statistics are presented in Appendix 2, Table P22. Table 16 compares linear associations between the ELAPSE and MAPLE  $PM_{2.5}$  main exposure model. For acute coronary heart disease and stroke, associations with  $PM_{2.5}$  from MAPLE were substantially stronger, while for lung cancer the associations with  $PM_{2.5}$  from MAPLE were much weaker. For asthma and COPD admissions, effect estimates were similar for the two models. We additionally compared the baseline exposure for ELAPSE with the MAPLE 2001 exposure, the earliest year of available MAPLE  $PM_{2.5}$  data. For all endpoints, HRs were larger for MAPLE  $PM_{2.5}$  in 2001 and the HRs for MAPLE 2001 were larger than for MAPLE 2010.

In Appendix 2, Figure P61 shows the SCHIF and spline functions for  $PM_{2.5}$  MAPLE. The functions mostly suggested steeper slopes at low levels for stroke and lung cancer consistent with the corresponding  $PM_{2.5}$  ELAPSE models. For the other outcomes, MAPLE functions were more consistent with lower slopes in the low end of the distribution, different from ELAPSE  $PM_{2.5}$ .

**Associations with Large Spatial Scale Ozone** Table P39 shows associations with  $O_3$  assessed with the larger spatial scale DEHM and MACC models. Associations for the large-scale DEHM model were more positive, especially for stroke, asthma, and COPD. For lung cancer, associations were more negative with the larger scale model. For the annual average MACC model, associations were even more negative than with the ELAPSE model, except for lung cancer.

**Table 16.** Hazard Ratios for Associations Between PM<sub>2.5</sub> from ELAPSE and MAPLE Exposure Model and Incidence<sup>a</sup>

	Acute Coronary Heart Disease HR (95% CI)	Stroke HR (95% CI)	Lung Cancer HR (95% CI)	Asthma HR (95% CI)	COPD HR (95% CI)
PM <sub>2.5</sub> ELAPSE	1.019 (0.943, 1.102)	1.105 (1.015, 1.203)	1.136 (1.049, 1.231)	1.273 (1.078, 1.504)	1.186 (1.071, 1.312)
PM <sub>2.5</sub> MAPLE	1.080 (1.014, 1.150)	1.171 (1.086, 1.263)	1.029 (0.956, 1.107)	1.313 (1.145, 1.505)	1.223 (1.123, 1.333)
PM <sub>2.5</sub> ELAPSE baseline, ratio	1.00 (0.96, 1.04)	1.05 (1.00, 1.09)	1.058 (1.020, 1.097)	1.037 (0.959, 1.121)	1.029 (0.981, 1.081)
PM <sub>2.5</sub> MAPLE 2001	1.146 (1.035, 1.269)	1.310 (1.158, 1.482)	1.065 (0.977, 1.162)	1.541 (1.235, 1.922)	1.402 (1.215, 1.618)

<sup>a</sup> N = 136,245 for acute coronary heart disease and stroke; N = 306,104 for lung cancer; N = 97,413 for asthma; N = 97,593 for COPD. HR (95% CI) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>. Main model adjusted for cohort ID, age, sex, year of baseline visit, smoking (status, duration, intensity, intensity squared), BMI, marital status, employment status, and 2001 neighborhood-level mean income. Acute coronary heart disease, stroke, asthma, and COPD incidence were additionally adjusted for educational status.

### Summary of the Associations with Incidence of Disease in the Pooled Cohort

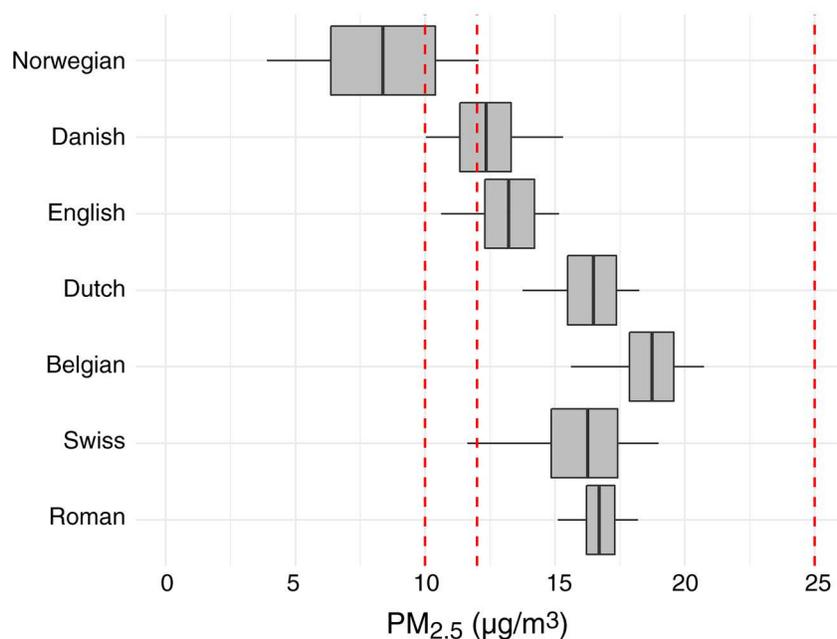
We found significant positive associations between PM<sub>2.5</sub>, NO<sub>2</sub>, and BC and incidence of stroke, asthma, and COPD hospital admissions. NO<sub>2</sub> was furthermore significantly related to acute coronary heart disease and PM<sub>2.5</sub> with lung cancer incidence. We generally observed linear to supralinear associations with no evidence of a threshold, with the exception of the association between NO<sub>2</sub> and acute coronary heart disease, which was sublinear. Subset analyses documented that PM<sub>2.5</sub> associations remained below 20 µg/m<sup>3</sup> and possibly 12 µg/m<sup>3</sup>. NO<sub>2</sub> associations remained below 30 µg/m<sup>3</sup> and in some cases 20 µg/m<sup>3</sup>. In two-pollutant models, NO<sub>2</sub> was most consistently associated with acute coronary heart disease, stroke, asthma, and COPD hospital admissions. PM<sub>2.5</sub> was not associated with these outcomes in two-pollutant models with NO<sub>2</sub>. PM<sub>2.5</sub> was the only pollutant that was associated with lung cancer incidence in two-pollutant models. Associations with O<sub>3</sub> were negative though generally not statistically significant in two-pollutant models.

### ADMINISTRATIVE COHORTS

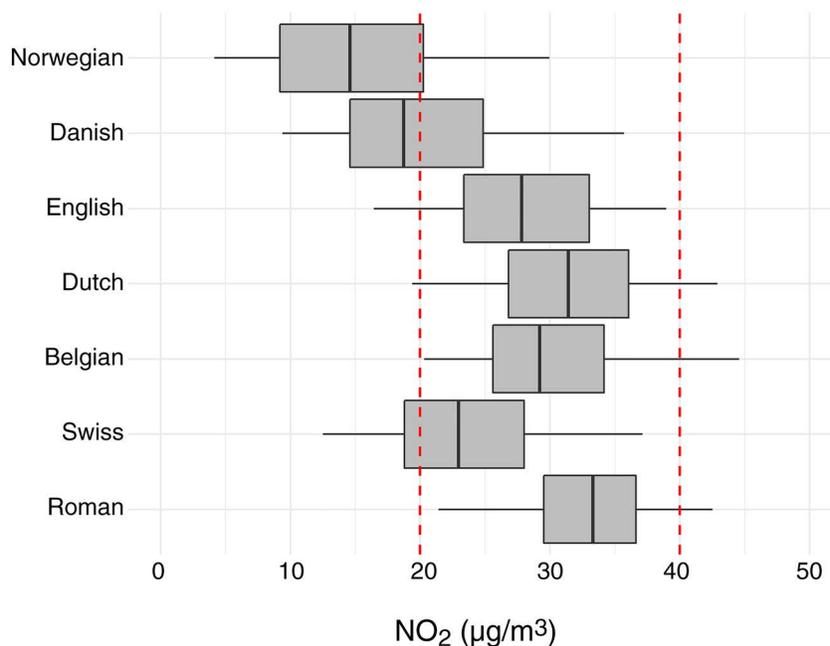
In the administrative cohorts, only associations with natural-cause and cause-specific mortality have been analyzed. After providing descriptive statistics of exposure and population, we present the main analyses and an extensive series of additional analyses related especially to definition of exposure.

#### Exposure

Figures 12 and 13 show the exposure distributions for PM<sub>2.5</sub> and NO<sub>2</sub>, respectively, for each of the seven administrative cohorts, based on the main ELAPSE models. Exposure distributions for BC and O<sub>3</sub> are shown in Figure A1 and Table A1 in Appendix 3 (the administrative cohorts appendix [available on the HEI website]). Large differences between cohorts were observed. Contrasts within cohorts were moderate for PM<sub>2.5</sub> and higher for NO<sub>2</sub>. The Norwegian cohort PM<sub>2.5</sub> exposures were predominantly below 10 µg/m<sup>3</sup>, the current WHO guideline. In the other cohorts, fewer than 5% of the subjects had an exposure below 10 µg/m<sup>3</sup>. Virtually all subjects had 2010 PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below the EU limit values. More than 3.9 million and 1.9 million subjects had a residential PM<sub>2.5</sub> exposure below the U.S.



**Figure 12. Description of PM<sub>2.5</sub> exposure at participant addresses in administrative cohorts.** The boundary of the box closest to zero indicates P25, the boundary furthest from zero indicates P75, the bold line in the middle of the box indicates P50, and the whiskers indicate P5 and P95. Dotted lines are WHO air quality guideline (10 µg/m<sup>3</sup>), U.S. EPA standard (12 µg/m<sup>3</sup>), and EU limit value (25 µg/m<sup>3</sup>), respectively. P = percentile. (Adapted from Stafoggia et al. [In review])



**Figure 13. Description of NO<sub>2</sub> exposure at participant addresses in administrative cohorts.** The boundary of the box closest to zero indicates P25, the boundary furthest from zero indicates P75, the bold line in the middle of the box indicates P50, and the whiskers indicate P5 and P95. Red dotted lines indicate the WHO air quality guideline (40 µg/m<sup>3</sup>) and the WHO Health Risks of Air Pollution in Europe health impact quantification threshold (20 µg/m<sup>3</sup>) — all annual averages. P = percentile. (Adapted from Stafoggia et al. [In review])

EPA NAAQS (12 µg/m<sup>3</sup>) and the WHO guideline (10 µg/m<sup>3</sup>), respectively.

Correlations between pollutants by cohort are shown in Appendix 3, Table A2. Correlations between PM<sub>2.5</sub> and NO<sub>2</sub> were generally moderately high to high. Correlations between PM<sub>2.5</sub> and BC were similar to the PM<sub>2.5</sub>–NO<sub>2</sub> correlations. NO<sub>2</sub> and BC were highly correlated for all cohorts ( $R > 0.84$ ). O<sub>3</sub> had moderate to high negative correlations with all three pollutants, especially NO<sub>2</sub> and BC.

### Descriptive Data on Population and Mortality Outcomes

Tables 17 and 18 contain population characteristics and the numbers of natural-cause and cause-specific deaths in the seven administrative cohorts. A very large number of subjects and a large number of deaths have been included in the analysis. With the exception of the English cohort, no individual lifestyle data were available. All cohorts, except the English cohort, did have individual- and area-level information on SES. The administrative cohorts were able to adjust for more area-level variables than in the pooled cohort analyses. All cohorts, except the English and Roman cohort, included both regional- and neighborhood-scale socioeconomic variables. In the Roman cohort, regional-scale variability was not applicable. In the English cohort, complicated privacy regulations prevented linking of data other than the deprivation score. We thus adjusted for an area indicator variable in the main analysis in the English cohort. The variables included in the analysis for the cohorts differed in availability of data. For example, education was available for nearly the entire population in the Norwegian, Swiss, and Roman cohorts, but data on education was very incomplete for the Dutch and Danish population. As cohorts were analyzed individually, we included the richest set of predictors, taking into account the percentage of missing data. Few missing data occurred, with the exception of the Belgian cohort where about 15% of the cohort had missing data.

A very large number of deaths and person-years of follow-up were available for analysis in each cohort. Mortality rates expressed per person-year of follow-up were fairly similar across the cohorts (Table 18). The distribution of causes of deaths differed somewhat between the seven cohorts (Table 18), possibly reflecting differences in coding of cause of death between countries in addition to population differences. Overall cardiovascular disease mortality ranged from 25% (English cohort) to 39% (Roman cohort) of all natural-cause deaths. Respiratory deaths contributed 6%–7% in the Roman and Swiss cohorts and 10%–14% in the other cohorts.

### Main Analyses of Associations with Natural-Cause and Cause-Specific Mortality

**Single-Pollutant Models** Figure 14 shows forest plots of the effect estimates of the linear models for natural-cause mortality and PM<sub>2.5</sub>, NO<sub>2</sub>, BC, and O<sub>3</sub> in the seven administrative cohorts, including the random effects pooled estimate and an indication of heterogeneity. Forest plots for cause-specific mortality are shown in Appendix 3, Figures A2 to A8. Table 19 shows the meta-analytic summary HRs for natural-cause and cause-specific mortality and the four main pollutants of the linear single-pollutant models in the seven cohorts.

Positive associations that were statistically significant with natural-cause mortality were found for PM<sub>2.5</sub>, NO<sub>2</sub>, and BC in meta-analyses of the seven cohorts for these three pollutants (Table 19). HRs were elevated for all cohorts (Figure 14). Large heterogeneity was found of HRs between cohorts, with relatively high PM<sub>2.5</sub> effect estimates for the Danish cohort and low estimates for the Belgian, Dutch, English, and Swiss cohort. Related to the large heterogeneity, the weights of all cohorts to the random effect summary HR were nearly identical, despite fairly large differences in the width of the individual cohorts' intervals. The large heterogeneity of HRs across cohorts also explains the larger width of the summary HR CI compared with some of the individual cohorts estimates. The width of the CI differed more between cohorts for PM<sub>2.5</sub> than for NO<sub>2</sub> and BC, related to more similar exposure contrast for NO<sub>2</sub> and BC compared with PM<sub>2.5</sub>. Differences in precision of effect estimates between cohorts are further affected by large differences in follow-up time and number of events. O<sub>3</sub> was negatively associated with natural-cause mortality and cause-specific mortality in these single-pollutant analyses, a pattern found in most cohorts.

Combined HRs for cardiovascular, respiratory, and lung cancer mortality were elevated and mostly statistically significant (Table 19). HRs for respiratory and lung cancer mortality were higher than for natural-cause mortality. HRs for cardiovascular mortality were slightly lower than for natural-cause mortality. Positive associations with cardiovascular mortality were found in all cohorts, except the Belgian cohort (Appendix 3, Figure A2). The same pattern of HRs across cohorts was found. Associations with ischemic heart disease were stronger than for stroke mortality (Table 19). Respiratory mortality was associated with NO<sub>2</sub> in all cohorts, PM<sub>2.5</sub> in all cohorts except the English and Swiss cohort, and BC in all cohorts except the English cohort. Lung cancer mortality was associated with NO<sub>2</sub> and BC in all cohorts and PM<sub>2.5</sub> in all cohorts except the Belgian and Roman cohorts. HRs were relatively high in all cohorts compared with the natural-cause mortality HRs. A positive association with diabetes mortality was found of

**Table 17.** Study Characteristics of Administrative Cohorts

	Belgian	Danish	Dutch	English	Norwegian	Roman	Swiss
Total participants ( <i>N</i> )	6,491,801	3,323,612	10,532,360	1,491,124	2,516,192	1,263,712	4,293,521
Participants with complete data ( <i>N</i> )	5,474,548	3,083,235	10,465,727	1,368,740	2,309,001	1,263,712	4,188,175
<b>Individual Level Covariates</b>							
Age (mean ± SD)	52.6 (15.2)	51.3 (13.1)	53.6 (15.1)	58.9 (12.8)	53.9 (15.9)	55.1 (15.4)	52.7 (15.2)
Women (%)	50.6	48.3	51.3	52.5	49.1	54.5	52.0
Smoking status (% current smokers)	NA	NA	NA	36.0 <sup>a</sup>	NA	NA	NA
BMI (mean ± SD) <sup>b</sup>	NA	NA	NA	27.5(5.6) <sup>a</sup>	NA	NA	NA
Country origin (% local)	96.9	94.5	82.8 <sup>c</sup>	NA	NA	NA	83.1
Household income	NA	Deciles	Deciles	NA	Quartiles	NA	NA
Marital status single 1	12.3	10.6 <sup>d</sup>	18.9	NA	17.8	15.3	14
Marital status married 2	68.3	69.4	63.1	NA	59.3	66.3	69.3
Marital status divorced 3	9.9	12.5	10.1	NA	12.3	7.0	8.7
Marital status widowed 4	9.5	7.5	8.0	NA	10.6	11.4	8.1
Education high 3	24.3	22 <sup>d</sup>	NA	NA	23	16.2	22.7
Education low 1	23.8	36.8	NA	NA	31.2	24.9	24.5
Education medium 2	51.9	41.1	NA	NA	45.8	58.8 <sup>e</sup>	52.7
Occupational status 1 (employed/self-employed)	53.3	64.7	NA	NA	65.8	45.8 <sup>f</sup>	61.4
Occstatus (unemployed) 2	5.1	2.7	NA	NA	1.2	5.0	2.2
Occstatus (homemaker) 3	8.5	32.6	NA	NA	0	21.0	14.6
Occstatus (retired) 4	33.2	no value	NA	NA	33	23.5	21.8
Mother tongue (German + Rhaeto-Romansch)	NA	NA	NA	NA	NA	NA	65.1
Mother tongue (French)	NA	NA	NA	NA	NA	NA	19.6
Mother tongue (Italian)	NA	NA	NA	NA	NA	NA	7.4
Mother tongue (other)	NA	NA	NA	NA	NA	NA	8
<b>Area-Level Covariates<sup>g</sup></b>							
Neighborhood income	Linear	Linear	Linear	NA	Linear	Deciles	NA
Neighborhood low education	Linear	Linear	NA	NA	Linear	Linear	Linear
Neighborhood high education	NA	NA	NA	NA	NA	Linear	Linear
Neighborhood unemployment rate	Linear	Linear	Linear	NA	Linear	Linear	Linear
Neighborhood non-Western ethnic	Linear	NA	Linear	NA	NA	NA	NA
Neighborhood score (% lowest)	NA	NA	Linear	Deciles	NA	Quintiles	Linear

M = manual; NA = not applicable; NM = nonmanual.

<sup>a</sup> 46.9% never smokers and 17.1% former smokers.

<sup>b</sup> BMI squared also in model.

<sup>c</sup> In seven categories.

<sup>d</sup> Not included in final model because of large fraction of missing data (HRs almost identical with and without marital status and education level adjustment).

<sup>e</sup> In two categories as high school (32.9) and junior high school (25.9).

<sup>f</sup> Employed in four categories: Employed NM High class (13.5%), Employed NM Low class (15.9%), Employed M (9.9%), Employed other (6.5%). Additional category other (4.8%).

<sup>g</sup> Because all cohorts have different representations, we add whether and how a variable was available and adjusted for in the main model 3.

**Table 18.** Frequency of Health Outcomes in Administrative Cohorts (Main Model)

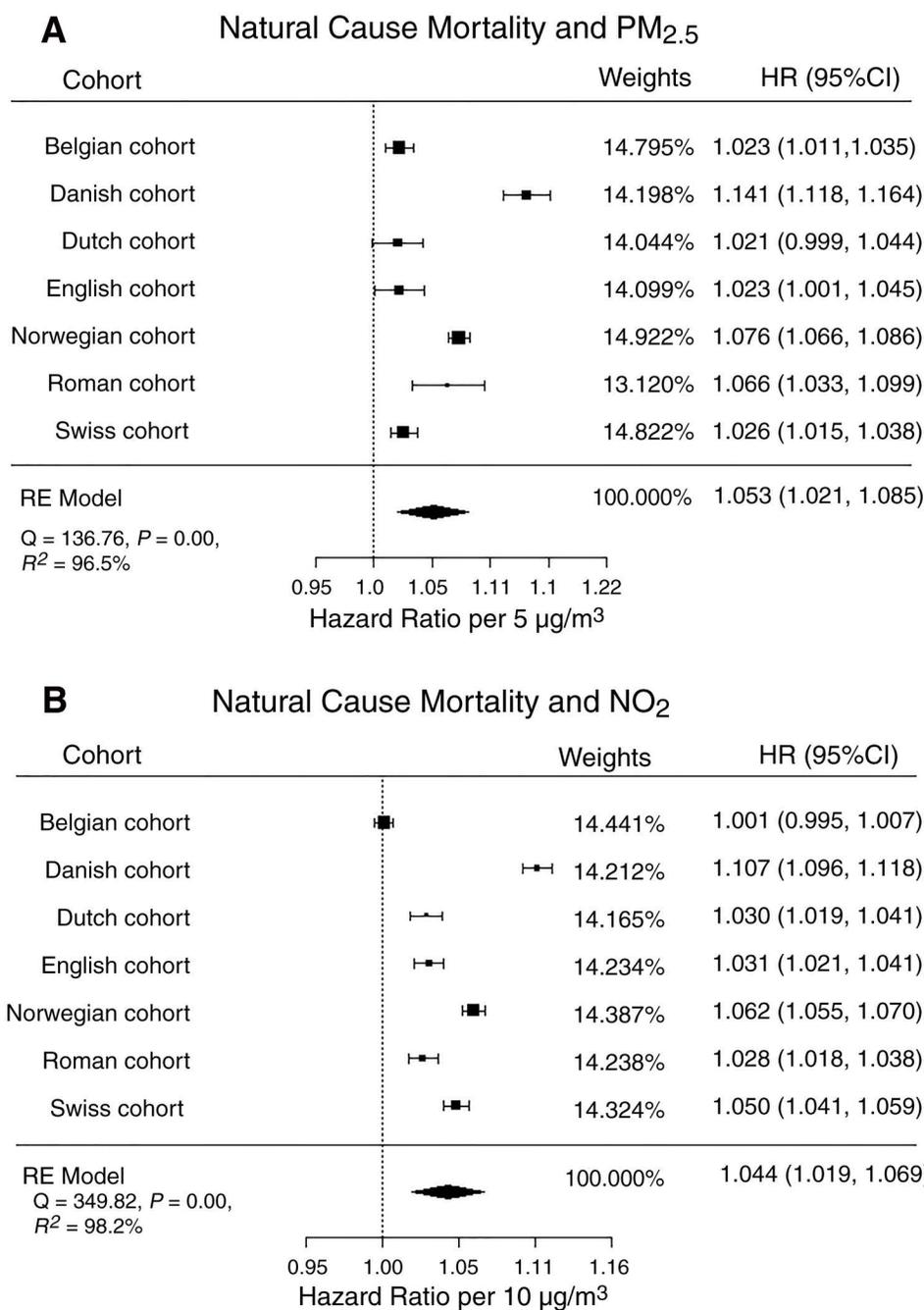
	Belgian	Danish	Dutch	English	Norwegian	Roman	Swiss
Main model (N)	5,474,548	3,083,235	10,465,727	1,368,740	2,309,001	1,263,712	4,188,175
Person-years at risk	54,575,223	42,586,584	50,436,539	9,084,293	32,531,421	15,301,265	53,344,296
Start follow-up	2001	2000	2008	2011	2001	2001	2000
End follow-up	2011	2015	2012	2017	2016	2013	2014
Natural deaths (N)	707,146	714,629	604,309	145,988	524,592	235,543	661,534
Mortality incidence per 1,000 person year	13.0	16.8	12.0	16.1	16.1	15.4	12.4
Deaths <sup>a</sup>							
Cardiovascular disease	234,553 (33)	232,699 (33)	165,000 (27)	36,615 (25)	183,971 (35)	91,278 (39)	241,985 (37)
Stroke	54,910 (8)	63,584 (9)	40,471 (7)	9,668 (7)	46,391 (9)	20,968 (9)	44,045 (7)
Ischemic heart disease	80,585 (11)	90,061 (13)	49,036 (8)	18,485 (13)	76,478 (15)	34,087 (14)	100,519 (15)
Respiratory disease	82,341 (12)	90,158 (13)	63,104 (10)	20,946 (14)	56,857 (11)	14,972 (6)	43,612 (7)
COPD	34,236 (5)	53,136 (7)	29,806 (5)	9,414 (6)	26,046 (5)	9,542 (4)	20,572 (3)
Lung cancer	52,211 (7)	51,881 (7)	49,294 (8)	9,850 (7)	27,525 (5)	19,572 (8)	36,680 (6)
Diabetes	11,610 (2)	18,267 (3)	13,587 (2)	780 (1)	9,510 (2)	9,592 (4)	15,276 (2)
Cardiometabolic	246,163 (34)	250,966 (35)	178,587 (30)	37,395 (26)	193,481 (37)	100,870 (43)	257,261 (39)

<sup>a</sup> Deaths expressed as *n* (%).

**Table 19.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause And Cause-Specific Mortality: Meta-analysis of Seven Administrative Cohorts, Fully Adjusted Model 3

	Natural Cause HR (95% CI)	Cardio-vascular HR (95% CI)	Ischemic Heart Disease HR (95% CI)	Stroke HR (95% CI)	Respiratory HR (95% CI)	COPD HR (95% CI)	Lung Cancer HR (95% CI)	Diabetes HR (95% CI)	Cardio-metabolic HR (95% CI)
PM <sub>2.5</sub>	1.053 (1.021, 1.085)	1.041 (1.010, 1.072)	1.057 (1.007, 1.109)	1.021 (0.980, 1.064)	1.064 (1.013, 1.118)	1.075 (0.995, 1.161)	1.102 (1.036, 1.172)	1.038 (0.974, 1.106)	1.040 (1.008, 1.073)
NO <sub>2</sub>	1.044 (1.019, 1.069)	1.025 (1.006, 1.044)	1.036 (1.007, 1.066)	1.019 (1.003, 1.035)	1.058 (1.024, 1.093)	1.078 (1.034, 1.124)	1.093 (1.053, 1.134)	1.013 (0.958, 1.070)	1.024 (1.004, 1.045)
BC	1.039 (1.018, 1.059)	1.022 (1.004, 1.040)	1.031 (1.007, 1.056)	1.014 (0.997, 1.033)	1.053 (1.021, 1.085)	1.074 (1.028, 1.122)	1.078 (1.038, 1.118)	1.015 (0.969, 1.065)	1.021 (1.003, 1.040)
O <sub>3</sub>	0.953 (0.929, 0.979)	0.976 (0.954, 0.998)	0.971 (0.936, 1.007)	0.986 (0.975, 0.997)	0.948 (0.910, 0.988)	0.935 (0.882, 0.991)	0.924 (0.887, 0.963)	0.984 (0.917, 1.057)	0.976 (0.953, 1.000)

<sup>a</sup> HR (95% CI) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC, 0.5 × 10<sup>-5</sup>/m; O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for age, sex, year of baseline visit, and cohort-specific individual- and area-level SES variables.



**Figure 14. Associations between PM<sub>2.5</sub>, NO<sub>2</sub>, BC, and O<sub>3</sub> and natural-cause mortality.** HRs from meta-analysis of seven administrative cohorts, fully adjusted model 3. (Adapted from Stafoggia et al. [In review])

(Figure 14 continues next page.)

similar magnitude as for cardiovascular mortality, but with wider CI that include unity (Figure A8). HRs for combined cardiometabolic disease mortality were slightly smaller than for cardiovascular mortality.

Tables A3 to A7 in Appendix 3 show HRs of models with increasing adjustment for individual- and area-level

confounders in the individual cohorts for natural-cause and cause-specific mortality. In all cohorts, HRs were sensitive to more complete adjustment for potential confounders. Adjustment for individual- and especially area-level confounders increased the HRs in some cohorts (e.g., Roman, Swiss, and Norwegian) and decreased HRs in

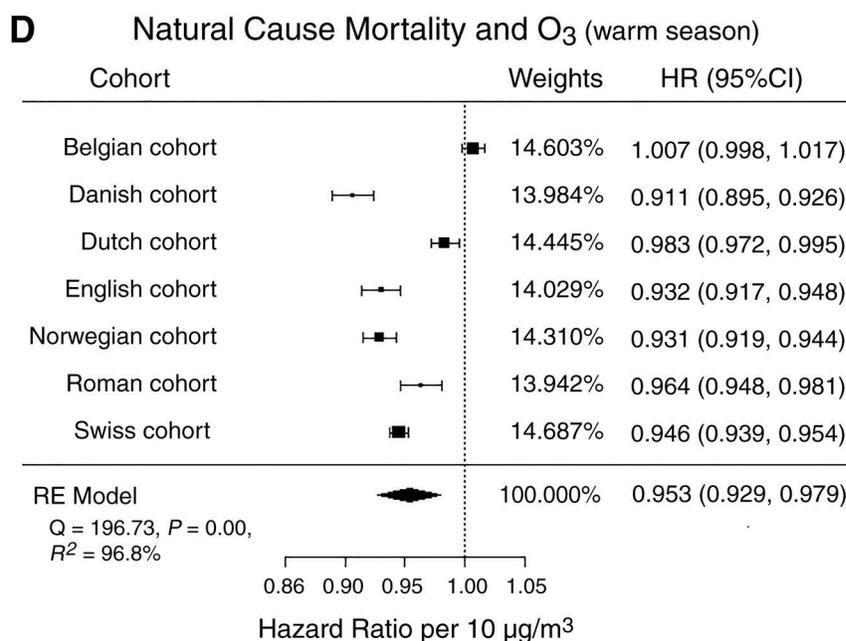
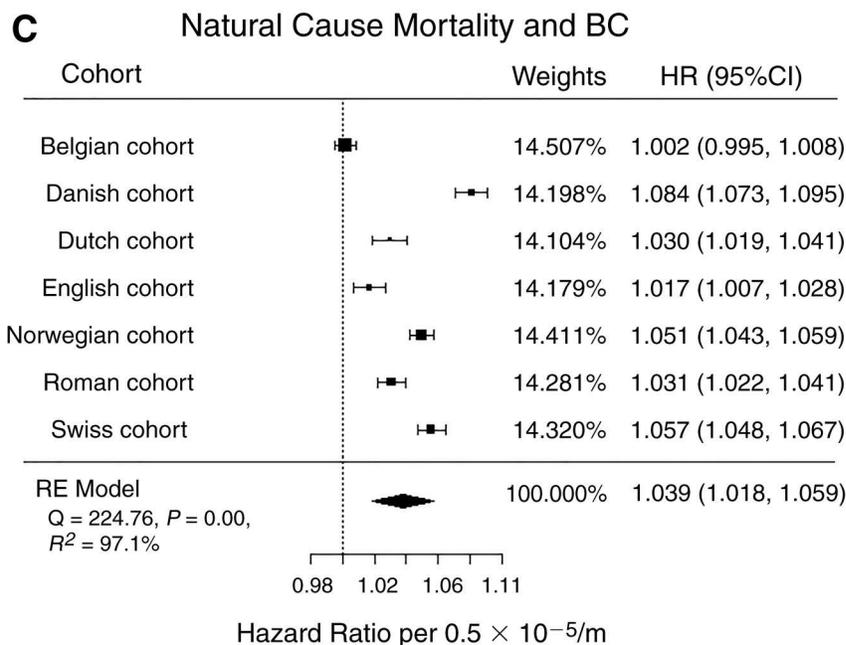
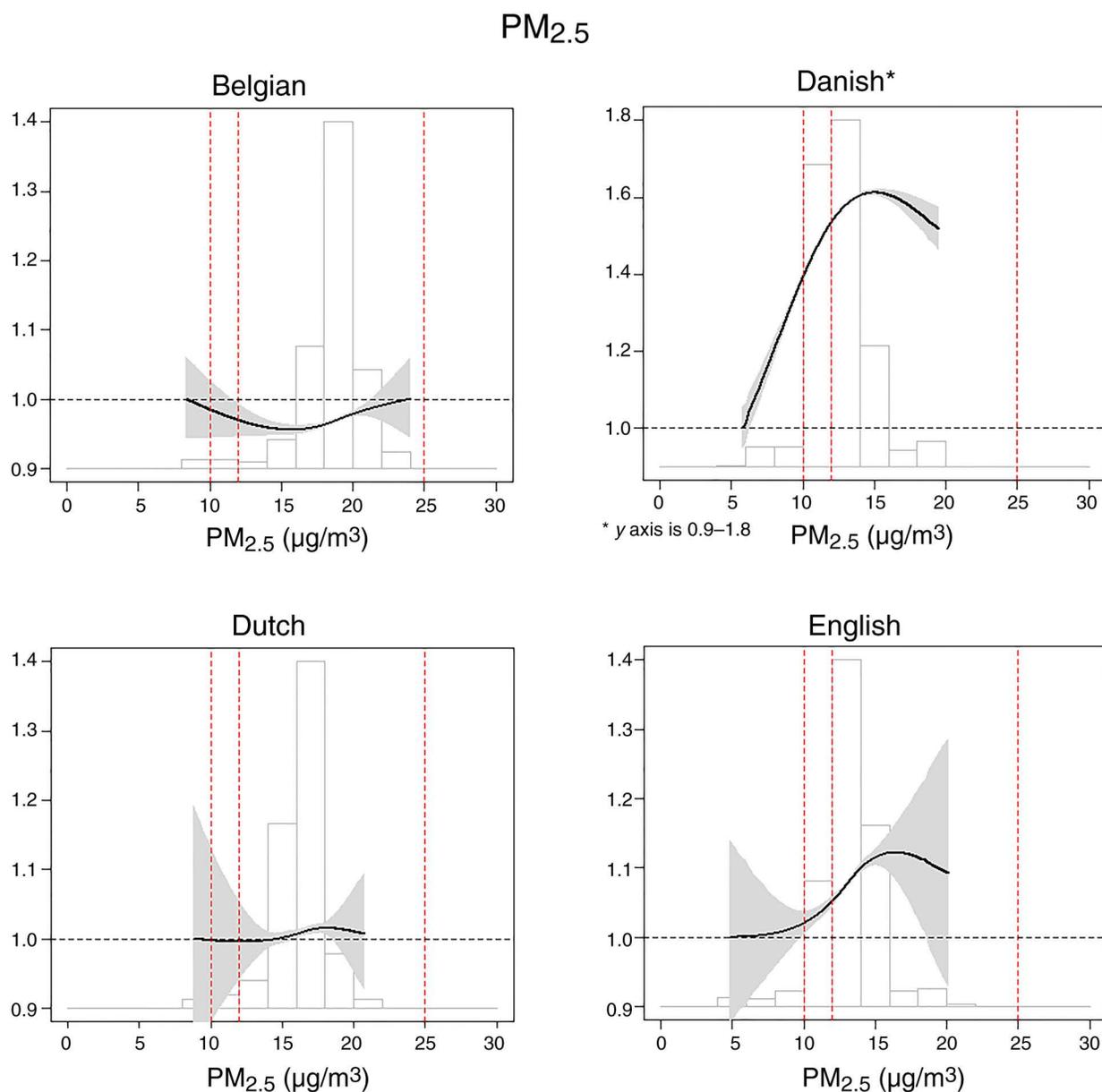


Figure 14 (Continued).

other cohorts (e.g., Dutch and English). Analyses of the minimally adjusted model 1 in the analyzed population with complete information for model 3 and the full study population showed nearly identical HRs, suggesting selection bias was not substantial (Appendix 3, Table A8).

**Concentration–Response Functions** Figures 15 and 16 show the natural spline (three degrees of freedom) and SCHIF functions for natural-cause mortality and PM<sub>2.5</sub> and NO<sub>2</sub> for the seven administrative cohorts. SCHIF and natural spline functions for natural-cause mortality and BC and O<sub>3</sub> are shown in Figures A9 and A10 in Appendix 3.



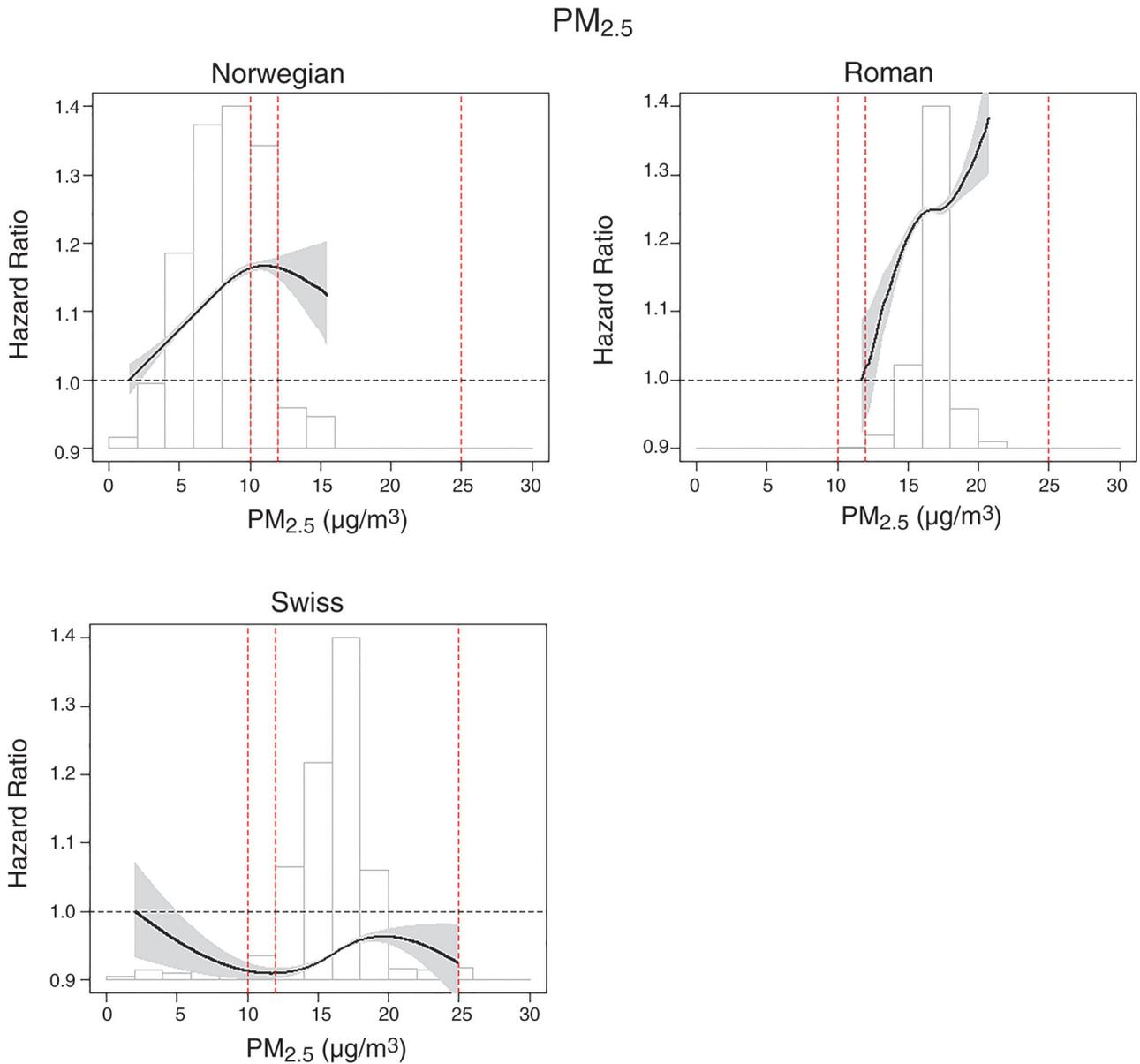
**Figure 15. Natural spline (three degrees of freedom) of the association between natural-cause mortality and PM<sub>2.5</sub> and NO<sub>2</sub> in seven administrative cohorts.** Red dotted lines are air quality limit and guideline values. Hazard ratios expressed relative to minimum exposure. Histograms are of exposure distributions. (Adapted from Klompaker et al 2021 [Creative Commons license CC BY-NC-ND 4.0] and Staffoggia et al. [In review])

(Figure 15 continues next 3 pages.)

Spline functions for cardiovascular, respiratory, and lung cancer mortality are shown in Figures A11 to A13.

The natural spline functions show different shapes in the seven cohorts for PM<sub>2.5</sub>. The cohorts differed substantially in PM<sub>2.5</sub> exposures, so the shapes reflect the concentration–response function in different ranges. In the Danish, English, Norwegian, and Roman cohorts, no evidence of a threshold was found. In the Norwegian and Danish cohorts,

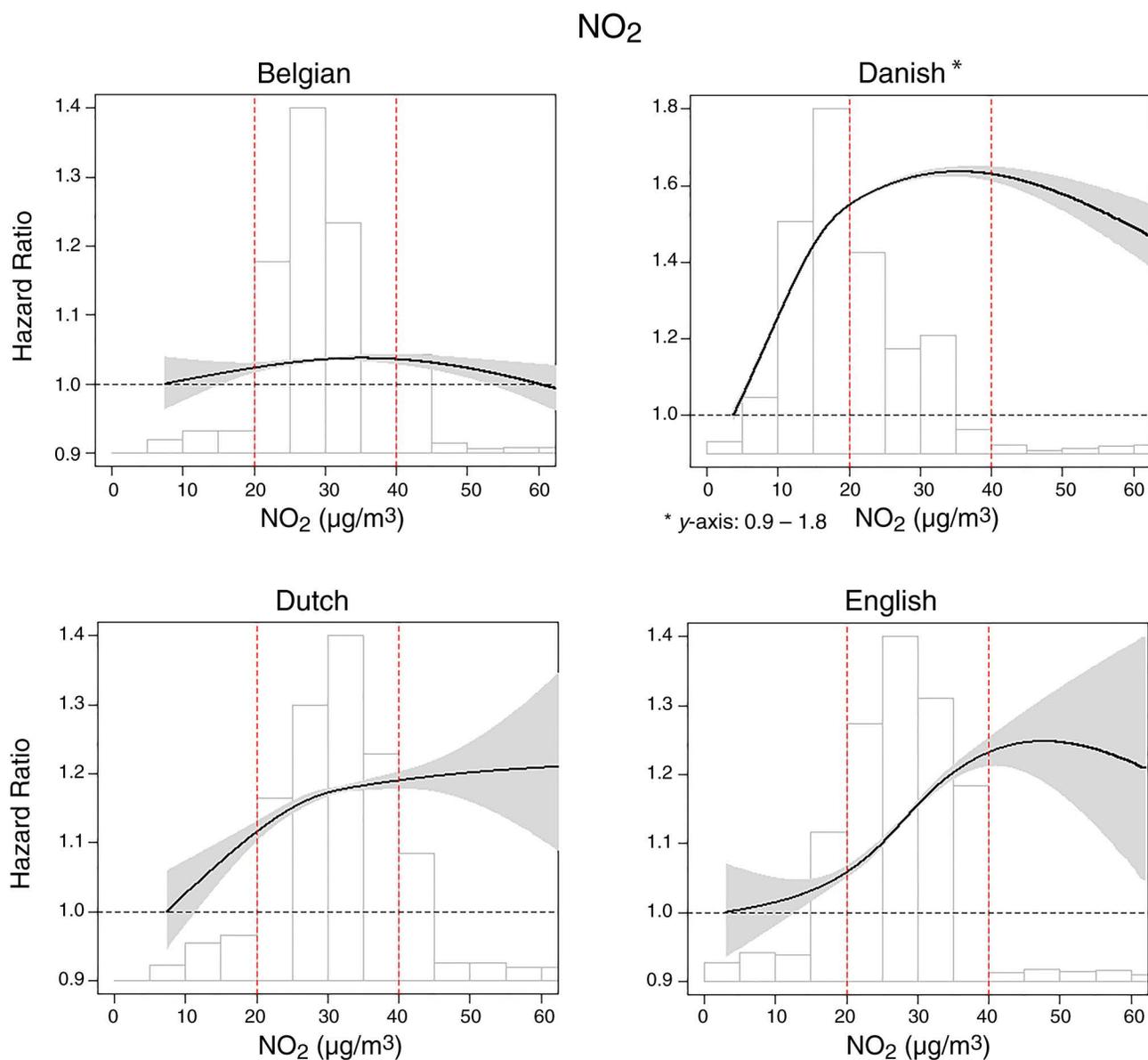
associations between PM<sub>2.5</sub> and mortality remained well below 10 µg/m<sup>3</sup>. In the other cohorts, few subjects had exposures below 10 µg/m<sup>3</sup>. In the Belgian and Swiss cohorts, at the lowest levels little evidence for an association was found. The HRs reflected the mortality risk with the risk at the minimum concentration. In the Swiss cohort, all HRs were below unity, related to a decrease in risk from the minimum until about 10 µg/m<sup>3</sup> (with very sparse data) and then



(Figure 15 Continued.)

an increase of risk in the concentration range with more data. We note that at the low and high end of the distribution, the uncertainty about the shape was large, as indicated by wide CI. The uncertainty was related to scarce data, as illustrated by the histogram. We prefer to show the full curves but note that, for example, for PM<sub>2.5</sub> exposures below about 15 and 9 µg/m<sup>3</sup> in the Belgian and Dutch cohorts, respectively, we cannot distinguish between an increasing, flat, or

decreasing curve. SCHIF functions showed generally similar though somewhat smoother patterns. SCHIF functions also showed no evidence of a threshold for all cohorts except the Swiss and Belgian cohorts (Figure 16). The uncertainty of the function was smaller for HRs close to the minimum value. This is an issue further interpreted in the Discussion section on low-level air pollution.

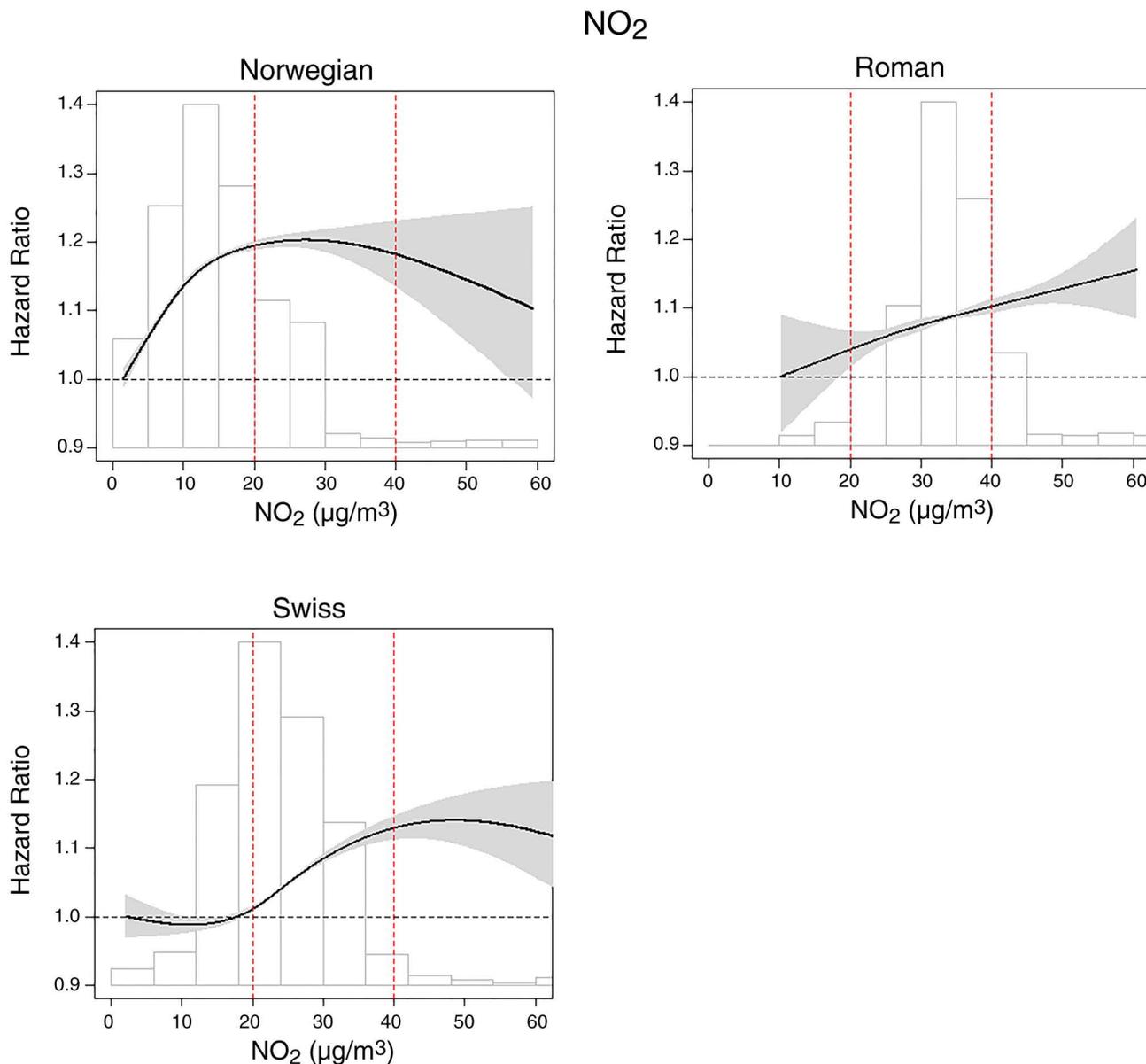


(Figure 15 Continued.)

For NO<sub>2</sub>, in most cohorts a supralinear shape was found in the spline analyses, with steeper slopes at low exposures and a leveling off at higher exposures. Exposure distributions of NO<sub>2</sub> were more similar between cohorts. In the Swiss cohort, little evidence for an association was found for exposures less than about 15 µg/m<sup>3</sup>, followed by a mostly linear increase and levelling off at high exposures. In the SCHIF analyses, patterns were similar: there

was no evidence of a threshold with the exception of the Swiss cohort, where there was no evidence of an association until about 20 µg/m<sup>3</sup>.

The splines for cardiovascular, respiratory, and lung cancer mortality showed more variable patterns of associations compared with natural-cause mortality. In most cohorts, associations remained for the lowest levels. For cardiovascular mortality, most cohorts showed a positive association

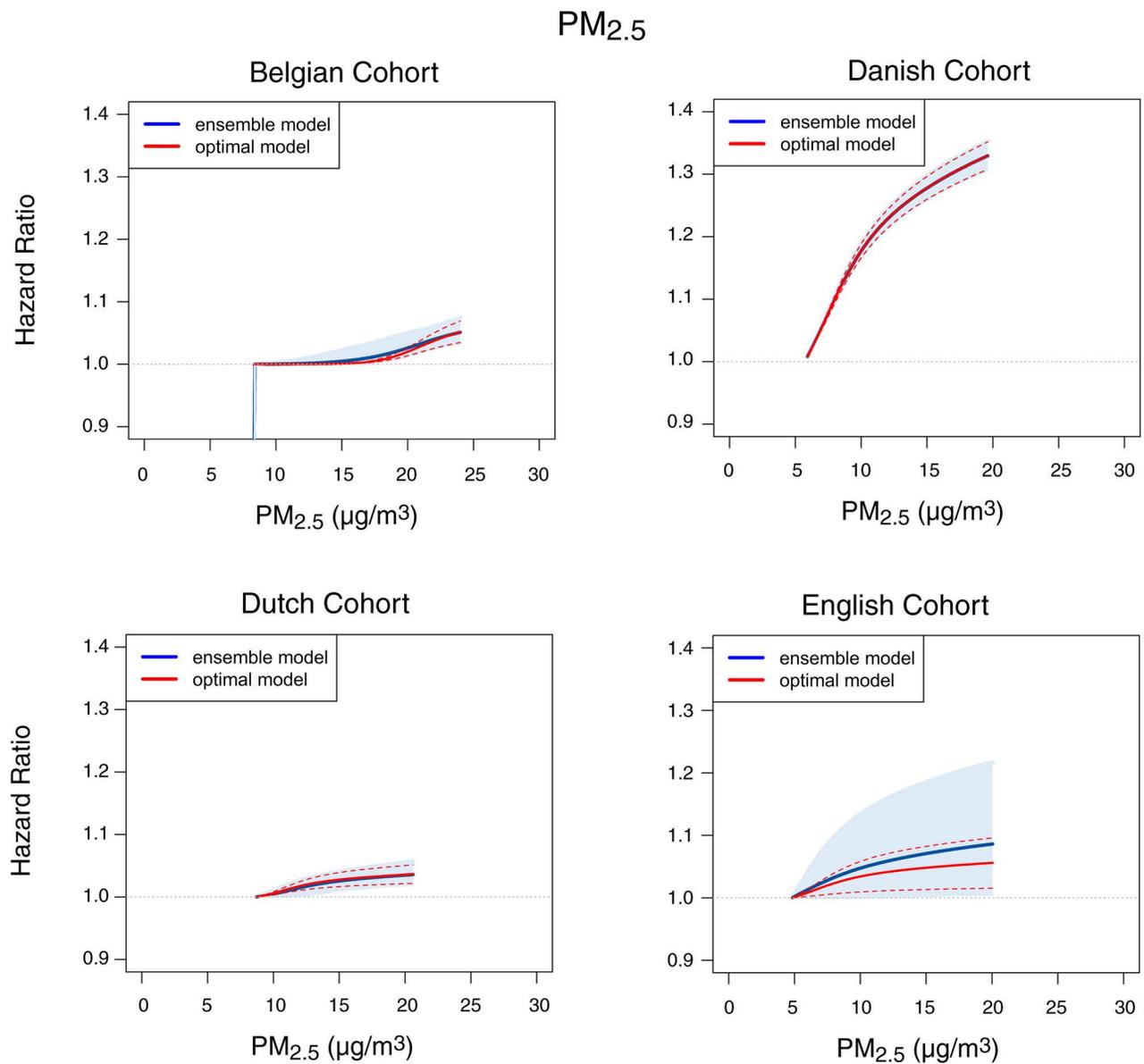


(Figure 15 Continued.)

with PM<sub>2.5</sub> at low levels, with the exception of the English cohort (Appendix 3, Figure A11). In all cohorts there was evidence that NO<sub>2</sub> was associated with cardiovascular mortality at low levels, though the association was weak in the Swiss cohort. For respiratory mortality, most cohorts showed a positive association with PM<sub>2.5</sub> at low levels, with the exception of the Roman and Swiss cohorts (Figure A12). In the Swiss cohort, a negative association was found for the

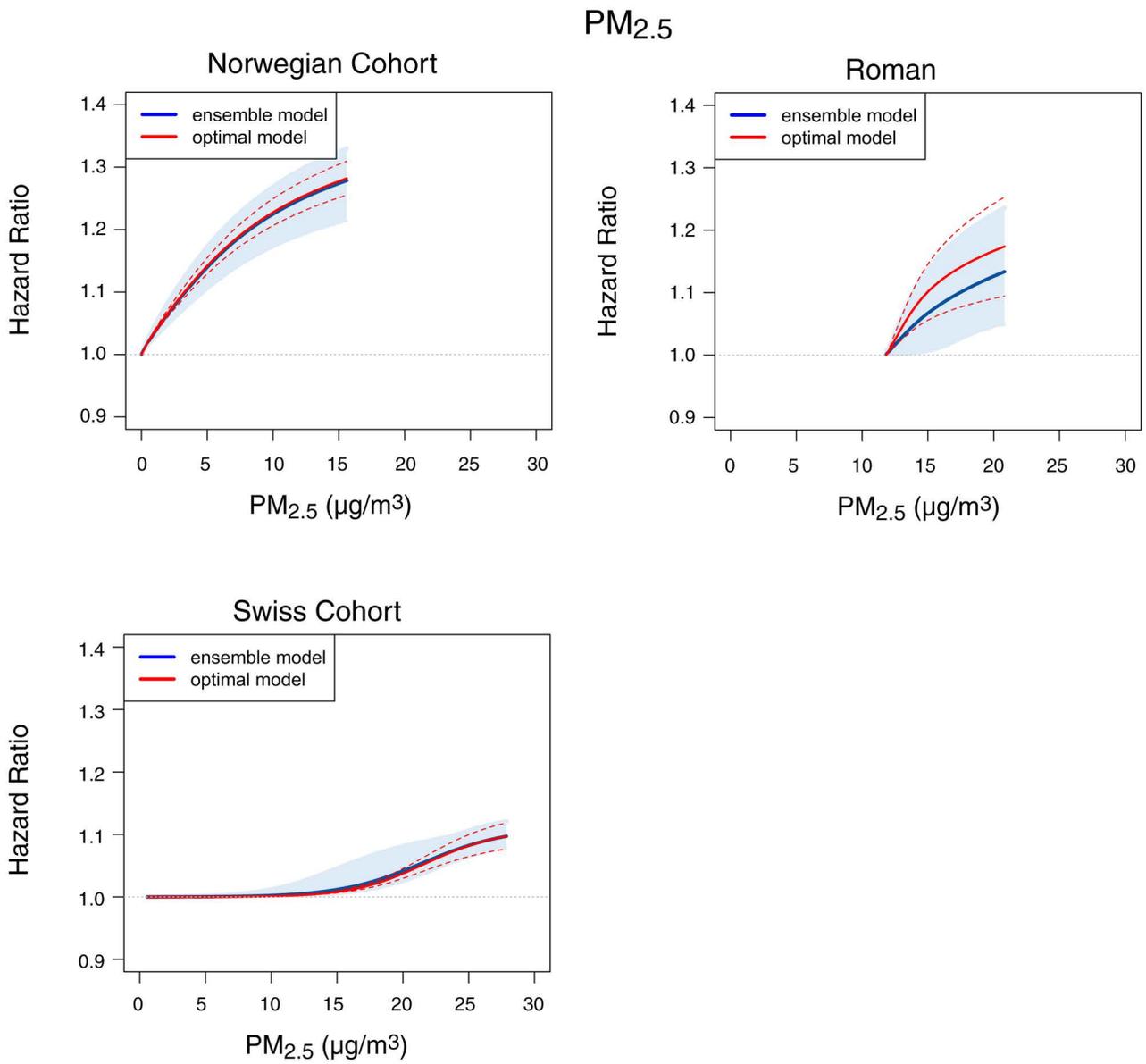
lowest exposures to about 15 µg/m<sup>3</sup>. For lung cancer mortality, no clear concentration–response pattern was found. Associations with PM<sub>2.5</sub> at low levels were not observed for the Belgian, Roman, and Swiss cohorts (Figure A13).

We had no possibility to combine nonparametric functions for the cohorts. The combined analysis of the shape of the concentration–response function was performed using standardized subset analyses.

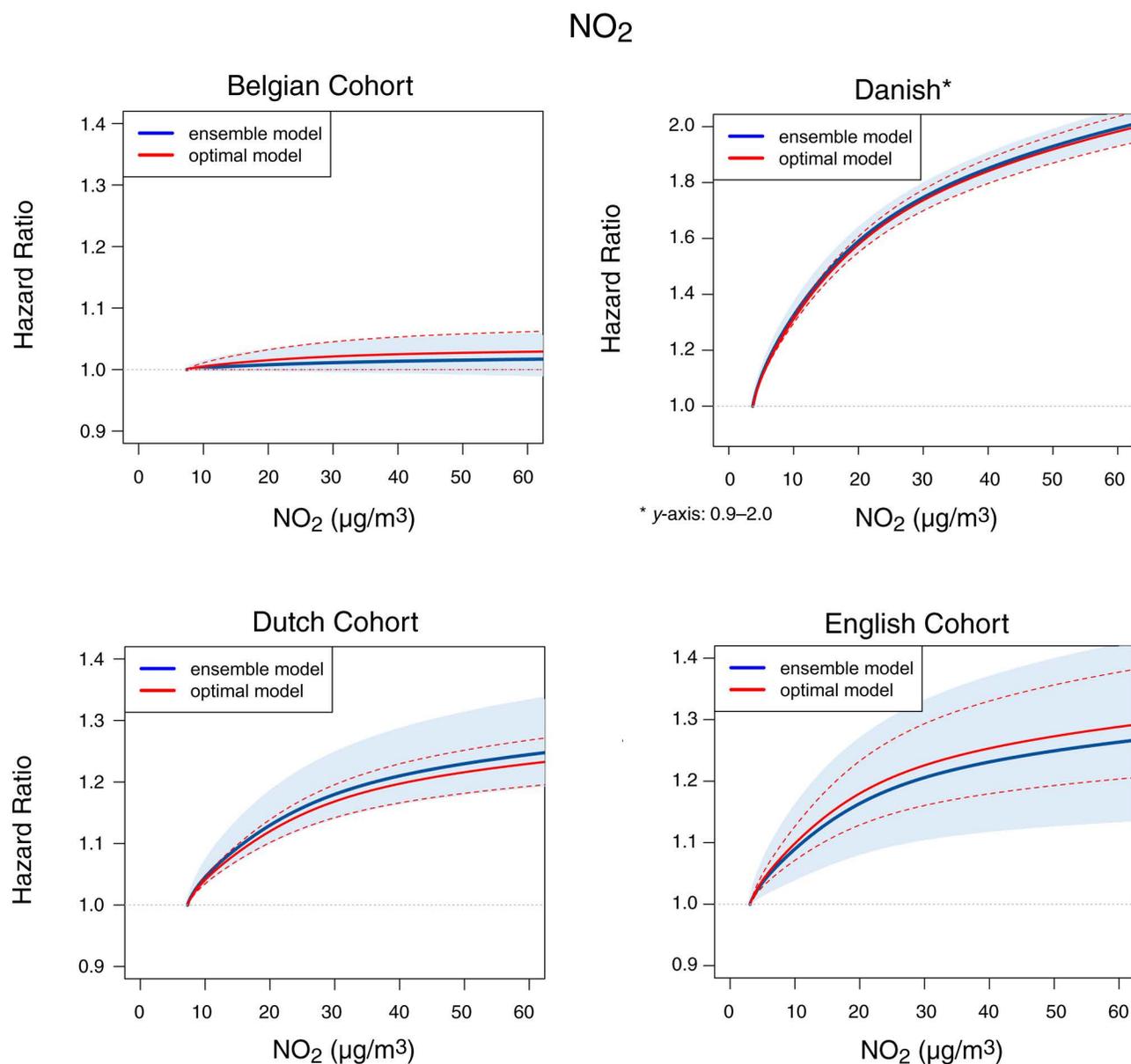


**Figure 16.** SCHIF functions of the associations between natural-cause mortality and PM<sub>2.5</sub> and NO<sub>2</sub> in seven administrative cohorts. Hazard ratios expressed relative to minimum exposure. Sparse data artefacts in some curves (Belgian PM<sub>2.5</sub> and Roman NO<sub>2</sub>) (Adapted from Stafoggia et al. [In review]).

(Figure 16 continues next 3 pages.)



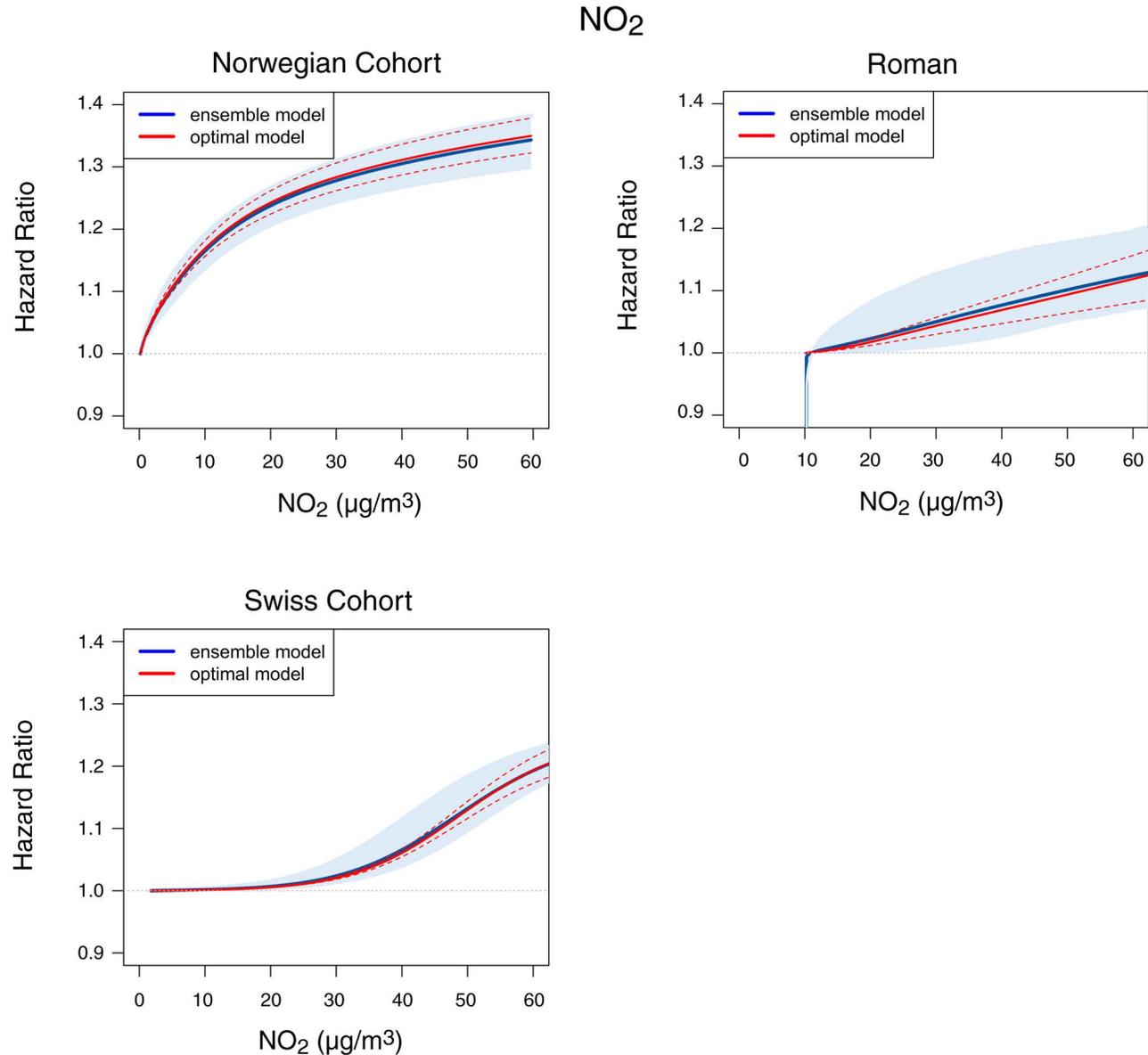
(Figure 16 continued.)



(Figure 16 continued.)

**Subset and Threshold Analysis** Table 20 shows the meta-analytic results of the subset analysis for natural-cause mortality based on the results of the separate subset analyses performed within the seven administrative cohorts. Subset analyses for cardiovascular, respiratory, and lung cancer mortality are shown in Appendix 3, Tables A9 to A11 (meta-analytic estimates). Subset analyses for the separate administrative cohorts are shown in Tables A12 to A16.

The subset analyses supported the observations from the spline and SCHIF analyses of the concentration–response function, in that associations remained stable when limiting the analysis to low pollution levels. For PM<sub>2.5</sub> and NO<sub>2</sub>, HRs were modestly larger when the analysis was restricted to exposures below 10 and 20 µg/m<sup>3</sup>, respectively, compared with HRs for the full range of exposures (Table 20). BC effect estimates for the lowest concentration ranges were modestly larger than for the full range. We



(Figure 16 continued.)

note that by restricting the concentration range, some full cohorts dropped out of the analyses because they had no subjects at low to very low concentrations. All cohorts contributed to the analysis of HRs below 15 µg/m<sup>3</sup> PM<sub>2.5</sub> (Table A11, Figure 12). The Danish, English, Norwegian, and Swiss cohorts contributed to the analysis below 10 and 12 µg/m<sup>3</sup> PM<sub>2.5</sub>. The analysis below 10 µg/m<sup>3</sup> PM<sub>2.5</sub> is dominated by the Norwegian cohort. All cohorts contributed to the analysis of NO<sub>2</sub> concentrations below 20 µg/m<sup>3</sup>.

We excluded effect estimates from very small subsets in a specific cohort with often extremely wide CI from the meta-analysis because these estimates could influence the summary estimates. This is because, when using random effects meta-analysis in the presence of large heterogeneity, each cohort-specific effect estimate receives nearly equal weight. We did not want very small subsets to have that large an influence.

**Table 20.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause Mortality in Subset Analysis: Meta-analysis of Seven Administrative Cohorts

Pollutant	Subset	Cohorts with Data <sup>a</sup> (N)		Cohorts Contributing to Meta-analysis <sup>b</sup> (N)		HR (95% CI)
		Cohorts	Subjects	Cohorts	Subjects	
<b>PM<sub>2.5</sub></b>						
	Full dataset	7	28,146,444	7	28,146,444	1.053 (1.021, 1.085)
	<25 µg/m <sup>3</sup>	7	28,146,444	7	28,146,444	1.053 (1.021, 1.085)
	<20 µg/m <sup>3</sup>	7	27,210,961	7	27,210,961	1.053 (1.022, 1.086)
	<15 µg/m <sup>3</sup>	7	9,703,270	7	9,703,270	1.051 (0.998, 1.108)
	<12 µg/m <sup>3</sup>	7	4,026,755	6	4,026,706	1.095 (1.002, 1.197)
	<10 µg/m <sup>3</sup>	6	1,921,220	4	1,920,292	1.078 (1.046, 1.111)
<b>NO<sub>2</sub></b>						
	Full dataset	7	28,146,444	7	28,146,444	1.044 (1.019, 1.069)
	<40 µg/m <sup>3</sup>	7	26,085,008	7	26,085,008	1.049 (1.024, 1.075)
	<30 µg/m <sup>3</sup>	7	16,791,623	7	16,791,623	1.063 (1.031, 1.097)
	<20 µg/m <sup>3</sup>	7	5,881,351	7	5,881,351	1.061 (0.985, 1.143)
<b>BC</b>						
	Full dataset	7	28,146,444	7	28,146,444	1.039 (1.018, 1.059)
	<3 × 10 <sup>-5</sup> /m	7	28,108,712	7	28,108,712	1.040 (1.019, 1.060)
	<2.5 × 10 <sup>-5</sup> /m	7	27,684,442	7	27,684,442	1.042 (1.021, 1.063)
	<2 × 10 <sup>-5</sup> /m	7	24,278,537	7	24,278,537	1.045 (1.022, 1.067)
	<1.5 × 10 <sup>-5</sup> /m	7	13,185,124	6	13,181,589	1.061 (1.032, 1.092)
	<1 × 10 <sup>-5</sup> /m	4	4,177,269	4	4,160,568	1.081 (0.966, 1.210)
	<0.5 × 10 <sup>-5</sup> /m	2	1,316,429	1	1,315,245	1.072 (1.053, 1.091)
<b>O<sub>3</sub></b>						
	Full dataset	7	28,146,444	7	28,146,444	0.953 (0.929, 0.979)
	<120 µg/m <sup>3</sup>	7	28,153,138	7	28,153,138	0.953 (0.929, 0.979)
	<100 µg/m <sup>3</sup>	7	28,112,455	7	28,112,455	0.952 (0.927, 0.978)
	<80 µg/m <sup>3</sup>	7	17,953,952	6	17,953,371	0.974 (0.946, 1.003)
	<60 µg/m <sup>3</sup>	6	522,723	3	522,359	0.965 (0.906, 1.028)

<sup>a</sup> HR (95% confidence interval) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC, 0.5 × 10<sup>-5</sup>/m; O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for age, sex, year of baseline visit, and cohort-specific individual- and area-level confounders.

<sup>b</sup> We excluded effect estimates from very small subsets in a specific cohort with often extremely wide CI from the meta-analysis.

In Appendix 3, Table A12 documents that also within individual cohorts, effect estimates for PM<sub>2.5</sub> and NO<sub>2</sub> for natural-cause mortality remain elevated when restricted to low concentration subsets. We only interpreted subsets with at least 100,000 subjects. In the Norwegian cohort, the HR for the subjects with a PM<sub>2.5</sub> exposure lower than 10 µg/m<sup>3</sup> was virtually identical to the HR in the full Norwegian population. In the Norwegian, English, and Swiss cohorts, HRs remained elevated when exposures above 12 µg/m<sup>3</sup> were removed. Associations in subsets of NO<sub>2</sub> concentrations below 20 µg/m<sup>3</sup> remained in all cohorts with more than 100,000 subjects.

For cardiovascular and respiratory mortality, HRs remained elevated below the EU limit values for PM<sub>2.5</sub> and NO<sub>2</sub> but became weaker for the lowest levels (Table A13 and A14). For lung cancer mortality, associations with PM<sub>2.5</sub> and BC remained at the lowest levels and were stronger than for the main analysis (Table A15). The association with NO<sub>2</sub> remained elevated below 30 µg/m<sup>3</sup>, but not 20 µg/m<sup>3</sup>. For diabetes mortality, associations in some cohorts remained at low levels but with wide CIs (Table A16).

In Appendix 3, Tables A17 to A21 show the results of threshold analyses for the seven cohorts for natural-cause, cardiovascular, respiratory, lung cancer, and diabetes mortality. For PM<sub>2.5</sub>, the no-threshold model had clearly lower AIC than the different threshold models in the Norwegian cohort (Table A17). In the other cohorts, AIC was nearly identical, reflecting the very low number of observations with exposure below 10 µg/m<sup>3</sup> in these cohorts. For NO<sub>2</sub>, the no-threshold model had clearly lower AIC than the different threshold models in the Norwegian cohort. In the Dutch, English, and Swiss cohorts a no-threshold model had better performance than a model with a threshold at 20 µg/m<sup>3</sup>. In the Belgian and Roman cohorts, differences between all models were small. On the basis of these results, we cannot exclude NO<sub>2</sub> thresholds of 10 or 15 µg/m<sup>3</sup>.

**Two-Pollutant Models** Results of the meta-analysis of two-pollutant models are shown in Table 21 for natural-cause mortality and in Appendix 3, Table A22, for cause-specific mortality. Tables A23 to A27 show results of two-pollutant models for each administrative cohort. We present all two-pollutant models, also when the correlation between pollutants was high. We have noted in the tables and text when we judged that limited interpretation of a two-pollutant model was possible.

After adjustment for BC and especially NO<sub>2</sub>, PM<sub>2.5</sub> HRs for natural-cause mortality were strongly attenuated, whereas the HRs for BC and NO<sub>2</sub> remained stable (Table 21). We note that the width of the CI is only modestly increased

in the PM<sub>2.5</sub> two-pollutant models compared with the single-pollutant models, which is consistent with the moderately high correlation of PM<sub>2.5</sub> with NO<sub>2</sub> and BC. This observation suggests that these models lead to interpretable results. Two-pollutant models of NO<sub>2</sub> and BC were more difficult to interpret because of the high correlation and the consistent substantial increase in width of the CI compared with single-pollutant models. After adjustment for BC and NO<sub>2</sub>, the negative associations for O<sub>3</sub> were increased to essentially unity. The attenuation of the PM<sub>2.5</sub> HR after adjustment for NO<sub>2</sub> was found in most cohorts (Table A23). Only in the Belgian, Danish, and Norwegian cohorts did a weak association of PM<sub>2.5</sub> with natural-cause mortality remain. In the other cohorts, the HR was essentially unity with several point estimates below one.

For cardiovascular, respiratory, and lung cancer mortality, the PM<sub>2.5</sub> association was also attenuated in two-pollutant models, especially so with NO<sub>2</sub> (Table A22). HRs for PM<sub>2.5</sub> remained statistically significant for cardiovascular mortality. For cardiovascular, respiratory, and lung cancer mortality, the PM<sub>2.5</sub> association was attenuated in two-pollutant models with NO<sub>2</sub> in most cohorts (Tables A24 to A26). Associations with NO<sub>2</sub> remained in two-pollutant models, especially for lung cancer mortality. Associations with cardiovascular mortality were weakest. Associations with BC remained elevated after adjustment for PM<sub>2.5</sub> but were strongly reduced after adjustment for NO<sub>2</sub>. The associations with diabetes were inconsistent in two-pollutant models but were already weak in single-pollutant models (Table A27).

#### **Additional Analyses of the Association with Mortality**

We performed extensive analyses of the robustness of the associations observed in the main analysis, including indirect adjustment for missing individual confounders, time-varying exposure analysis, comparison with local exposure models, additional adjustment for area of the country, additional adjustment for traffic noise, associations with PM<sub>2.5</sub> exposure estimates from the MAPLE study, and associations with O<sub>3</sub> at a large spatial scale. In general, the associations observed in our main analyses were robust to these alternative specifications and adjustments, as documented in the next sections of the report.

#### **Indirect Adjustment for Missing Individual Confounders**

Indirect adjustment for missing individual confounders was carried out in all but the English administrative cohorts. The English cohort had individual-level confounder data. In Appendix 1, Table M22 shows the surveys used to assess relationships between air pollution

**Table 21.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause Mortality in Two-Pollutant Analysis: Meta-analysis of Seven Administrative Cohorts<sup>a</sup>

Pollutant	Single-Pollutant HR (95% CI)	HR (95% CI) Adjusted For			
		PM <sub>2.5</sub>	NO <sub>2</sub>	BC	O <sub>3</sub>
Natural-Cause Mortality					
PM <sub>2.5</sub>	1.053 (1.021, 1.085)	NA	1.003 (0.982, 1.025)	1.021 (0.997, 1.046)	1.031 (0.999, 1.064)
NO <sub>2</sub>	1.044 (1.019, 1.069)	1.042 (1.02, 1.065)	NA	1.041 (1.009, 1.073)	1.040 (1.012, 1.069)
BC	1.039 (1.018, 1.059)	1.03 (1.012, 1.049)	1.004 (0.985, 1.022)	NA	1.028 (1.005, 1.051)
O <sub>3</sub>	0.953 (0.929, 0.979)	0.972 (0.947, 0.996)	0.987 (0.961, 1.014)	0.976 (0.948, 1.005)	NA

NA = not applicable.

<sup>a</sup> HR (95% CI) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC, 0.5 × 10<sup>-5</sup>/m; O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for age, sex, year of baseline visit, and cohort-specific individual- and area-level confounders. Note: because of the high correlation between BC and NO<sub>2</sub>, two-pollutant models of BC and NO<sub>2</sub> are difficult to interpret.

exposure and missing lifestyle data. National lifestyle and health surveys were used for all cohorts, except the Roman cohort. In Appendix 3, Table A28 presents a comparison between the distribution of population characteristics and exposure between survey and cohort populations. The survey population generally differed (slightly) from the cohort, suggesting the indirect adjustment approach provided a useful assessment of the potential change in effect estimates related to missing lifestyle factors. The exact quantitative results need to be interpreted with caution, however as the magnitude of the relationships may depend on the population. In the Netherlands, the original survey differed substantially from the general population cohort because in the survey the elderly were overrepresented on purpose. We therefore took a stratified random sample from the full survey population such that key characteristics available in both populations were identical. Sampling was feasible, because the survey population was large (~400,000 adults).

In Appendix 3, Table A29 presents the relationships between smoking status, BMI, and other lifestyle factors with air pollution exposure for six administrative cohorts. In all cohorts, air pollution exposure was associated with lifestyle factors after adjusting for the confounders included in the main confounder model. Associations differed in magnitude and direction per cohort. Current smokers had mildly higher PM<sub>2.5</sub> and NO<sub>2</sub> exposures in

the Norwegian and Swiss surveys. In the Roman population, smoking was associated with lower air pollution exposures. In the Dutch survey, only former smokers had higher NO<sub>2</sub> exposures. Overweight and obese subjects did not differ in air pollution exposure from normal-weight subjects in the Swiss and Dutch cohort. In Norway, overweight and obese subjects had lower air pollution exposures. In Rome, BMI was not available.

To apply indirect adjustment, risk estimates for smoking status and BMI in categories for natural-cause mortality were needed. We focused on smoking and BMI because these lifestyle measures were considered critical confounders in the WHO systematic review of outdoor air pollution (Chen and Hoek 2020). The HRs were estimated from the large pooled cohort, specifying Cox models stratified by sex and (sub)cohort, including smoking, BMI, marital status, and mean income at the neighborhood level during 2001 as covariates. These covariates were typically available in the administrative cohorts. The HRs for natural-cause mortality that we used were 1.16 and 2.13 for former and current smoker (compared with never smokers), respectively, and 1.71, 1.08, and 1.44 for underweight, overweight, and obese subjects (compared to normal weight), respectively.

HRs after indirect adjustment for natural-cause mortality are summarized in Table 22 (for PM<sub>2.5</sub> and NO<sub>2</sub>) and Appendix 3, Table A30 (for BC and O<sub>3</sub>). In the Dutch, Swiss,

and Norwegian cohorts, HRs were attenuated but remained (borderline) significant. In the Roman and Belgian cohorts, HRs increased after indirect adjustment. Table 22 also shows the HRs after adjustment for age-standardized area-level lung cancer, COPD, and diabetes rates as alternative approaches to adjust for missing lifestyle factors. In the Roman cohort, HRs were moderately attenuated. In the Norwegian and Danish cohorts, associations were slightly attenuated. In the Belgian cohort, HRs increased slightly

upon additional adjustment. Overall, associations remained after adjustment for missing lifestyle factors.

**Time-Varying Exposure Analysis** Tables 23 and 24 show the HRs from applying exposure at baseline and the time-varying exposure analysis instead of the 2010 exposure in the main model. In Appendix 3, Figure A14 compares exposures at baseline and the 2010 exposure, documenting generally higher exposure at baseline. The difference was

**Table 22.** Hazard Ratios for Associations Between PM<sub>2.5</sub> and NO<sub>2</sub> and Natural-Cause Mortality in Six Administrative Cohorts: Indirect Adjustment, Adjustment for Area-Level Disease<sup>a</sup>

Cohort	HR (95% CI), Main Model	HR (95% CI) with Indirect Adjustment	HR (95% CI), Main Model, Reduced Population Area-Level Adjustment	HR (95% CI) with Area-Level Lung Cancer, COPD, Diabetes
<b>PM<sub>2.5</sub></b>				
Belgian	1.023 (1.011, 1.035)	1.049 (1.036, 1.062)	NA <sup>b</sup>	1.029 (1.017, 1.041) <sup>c</sup>
Danish	1.141 (1.118, 1.164)	1.118 (1.095, 1.140)	NA <sup>b</sup>	1.143 (1.120, 1.167) <sup>d</sup>
Dutch	1.021 (0.999, 1.044)	1.015 (0.993, 1.038)	NA <sup>b</sup>	1.007 (0.985, 1.030) <sup>e</sup>
Norwegian	1.076 (1.066, 1.086)	1.055 (1.045, 1.065)	1.095 (1.080, 1.109) <sup>f</sup>	1.081 (1.067, 1.095) <sup>c</sup>
Roman	1.066 (1.033, 1.099)	1.111 (1.080, 1.142)	NA <sup>b</sup>	1.041 (1.009, 1.075) <sup>e</sup>
Swiss	1.026 (1.015, 1.038)	1.015 (1.003, 1.027)	NA <sup>b</sup>	NA
<b>NO<sub>2</sub></b>				
Belgian	1.001 (0.995, 1.007)	1.012 (1.006, 1.019)	NA <sup>b</sup>	1.004 (0.998, 1.011) <sup>c</sup>
Danish	1.107 (1.096, 1.118)	1.088 (1.077, 1.099)	NA <sup>b</sup>	1.107 (1.096, 1.119) <sup>d</sup>
Dutch	1.030 (1.019, 1.041)	1.020 (1.009, 1.031)	NA <sup>b</sup>	1.024 (1.013, 1.035) <sup>e</sup>
Norwegian	1.062 (1.055, 1.070)	1.051 (1.044, 1.059)	1.080 (1.070, 1.090) <sup>f</sup>	1.073 (1.063, 1.083) <sup>c</sup>
Roman	1.028 (1.018, 1.038)	1.044 (1.035, 1.053)	NA <sup>b</sup>	1.019 (1.009, 1.029) <sup>e</sup>
Swiss	1.050 (1.041, 1.059)	1.034 (1.025, 1.042)	NA <sup>b</sup>	NA

<sup>a</sup> HR (95% CI) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC, 0.5 × 10<sup>-5</sup>/m; O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for age, sex, year of baseline visit, and cohort-specific individual- and area-level SES variables.

<sup>b</sup> NA implies HR of full population applies.

<sup>c</sup> Municipality-level diabetes and lung cancer simultaneously.

<sup>d</sup> Municipality-level lung cancer, diabetes, and COPD simultaneously.

<sup>e</sup> Neighborhood scale lung cancer, diabetes, and COPD simultaneously.

<sup>f</sup> Reduced population available for municipality-level diseases adjustment (N = 1,868,397).

largest for PM<sub>2.5</sub>. In Appendix 3, Tables A31 and A32 present HRs by cohort. The combined HRs were almost identical to the HR in the main model for the baseline exposure back-extrapolated with the difference method (Table 23). The HRs for the baseline exposure back-extrapolated with the ratio method were slightly smaller than for the main exposure for PM<sub>2.5</sub> and NO<sub>2</sub> and almost identical for BC and O<sub>3</sub>, probably reflecting the smaller time trends for the latter pollutants. The CIs were smaller for the ratio-baseline exposure,

reflecting the larger variability in exposure and the reduced heterogeneity in effect estimates across cohorts (Table A31). In the Danish cohort, the HR for the ratio method was substantially smaller than the large HR for the main analysis. HRs in models with exposures back-extrapolated to the recruitment year were nearly identical for the Dutch cohort (Table A31), probably because the recruitment year (2008) was very close to the main ELAPSE exposure model year (2010). For the Belgian,

**Table 23.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause Mortality in Time-Varying Analyses: Baseline Exposure; Meta-analysis of Seven Administrative Cohorts<sup>a</sup>

Pollutant	Main Model HR (95% CI)	Back-Extrapolated Baseline Exposure (difference method) HR (95% CI)	Back-Extrapolated Baseline Exposure (ratio method) HR (95% CI)
PM <sub>2.5</sub>	1.053 (1.021, 1.085)	1.051 (1.018, 1.085)	1.039 (1.018, 1.061)
NO <sub>2</sub>	1.044 (1.019, 1.069)	1.044 (1.020, 1.069)	1.036 (1.018, 1.055)
BC	1.039 (1.018, 1.059)	1.039 (1.019, 1.059)	1.038 (1.017, 1.058)
O <sub>3</sub>	0.953 (0.929, 0.979)	0.953 (0.929, 0.978)	0.953 (0.930, 0.976)

<sup>a</sup>  $N = 28,121,173$ . HR (95% CI) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC,  $0.5 \times 10^{-5}$ /m; O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for age, sex, year of baseline visit, and cohort-specific individual- and area-level SES variables.

**Table 24.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause Mortality in Time Varying Analyses: Time-Varying Exposure Including Residential Mobility; Meta-analysis<sup>a</sup> of Administrative Cohorts

Pollutant	Main Model	Strata for 1 Year <sup>b</sup> (difference method) HR (95% CI)	Strata for 1 Year <sup>b</sup> (ratio method) HR (95% CI)	Strata for 5 Year <sup>b</sup> (difference method) HR (95% CI)	Strata for 5 Year <sup>b</sup> (ratio method) HR (95% CI)
PM <sub>2.5</sub>	1.053 (1.021, 1.085)	1.047 (1.012, 1.083)	1.039 (1.012, 1.068)	1.070 (0.969, 1.182)	1.061 (0.975, 1.154)
NO <sub>2</sub>	1.044 (1.019, 1.069)	1.039 (1.014, 1.063)	1.037 (1.014, 1.060)	1.044 (1.007, 1.082)	1.049 (1.007, 1.093)
BC	1.039 (1.018, 1.059)	1.033 (1.014, 1.052)	1.034 (1.015, 1.054)	1.027 (0.988, 1.068)	1.015 (0.945, 1.090)
O <sub>3</sub>	0.953 (0.929, 0.979)	0.957 (0.932, 0.982)	0.958 (0.934, 0.983)	1.009 (0.944, 1.078)	1.014 (0.952, 1.080)

<sup>a</sup> Meta-analyses of seven cohorts (1-year strata) and of five cohorts (5-year strata, all but Dutch and English because of short follow-up time).

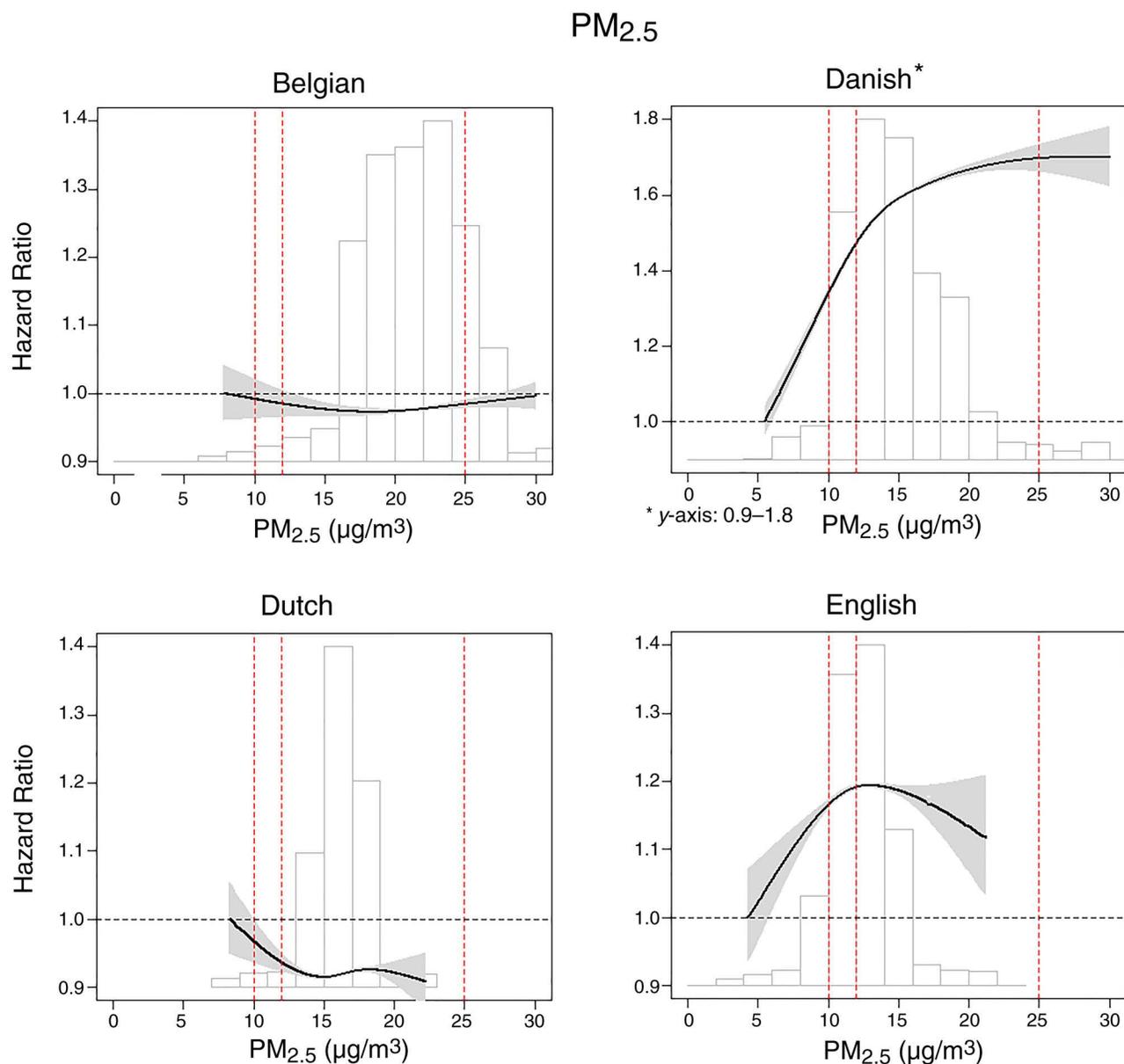
<sup>b</sup> To adjust for time trends in mortality and pollution.

Swiss, Norwegian, and Roman cohorts — all with a start of follow-up around 2000 — HRs were slightly smaller in the baseline exposure analysis than in the main model. In the English cohort, this analysis was not conducted, as the recruitment year was 2011.

Combined HRs in time-varying exposure analyses were slightly lower than the HRs for the main ELAPSE exposure when time trends were adjusted for with 1-year strata (Table 24). HRs for time-varying exposures with 5-year strata (to adjust for time trends) were larger for PM<sub>2.5</sub>, similar for NO<sub>2</sub>, and smaller for BC. The O<sub>3</sub> estimates were positive, albeit nonsignificant. HRs for 5-year strata may be more difficult to interpret in analyses of approximately 15-year follow-up periods. Time-varying exposure analyses in the Swiss cohort showed similar HRs compared with the main ELAPSE exposure of 2010 when adjustments for time trends were made with 1-year strata (Table A32). In time-varying analyses with 5-year time trends, HRs were sub-

stantially larger than in the main analyses. In the Roman cohort, HRs were slightly smaller compared with the main analysis with 1-year strata adjustment. With 5-year strata, HRs below unity were observed. In the Belgian cohort, HRs were smaller in the time-varying analysis for 1-year strata, with HRs below 1 for NO<sub>2</sub> and BC. HRs for PM<sub>2.5</sub> were higher in the 5-year strata analysis. In the Norwegian cohort, HRs were modestly larger in the time-varying analysis, both for the 1- and 5-year strata analyses.

Additional specification of natural splines in the time-varying exposure analysis further supported that the associations with mortality were observed at low pollutant levels (Figure 17 for PM<sub>2.5</sub> and NO<sub>2</sub> and Appendix 3, Figure A15 for BC and O<sub>3</sub>). The time-varying exposure analysis is important because in our main model we used the 2010 exposure, which represented relatively low levels during follow-up.

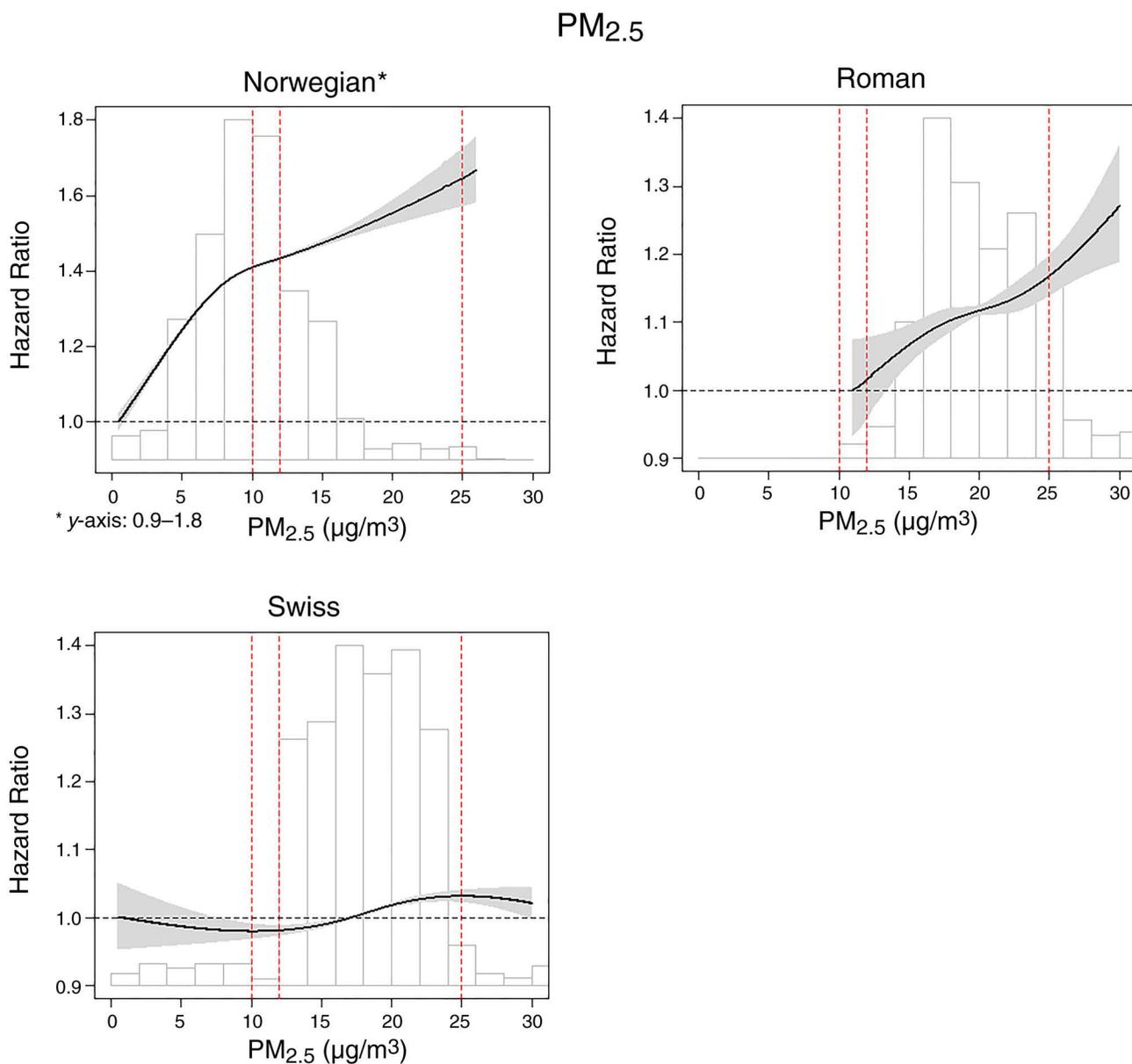


**Figure 17. Natural spline (three degrees of freedom) of the association between natural-cause mortality and PM<sub>2.5</sub> and NO<sub>2</sub>; time-varying exposure analysis.** Red dotted lines are air quality limit and guideline values. Hazard ratios expressed relative to minimum exposure. Histograms are of exposure distributions (Adapted from Stafoggia et al. [In review]).

(Figure 17 continues next 3 pages.)

**Comparison with Local Exposure Models** In Appendix 3, Table A33 shows the correlations between national dispersion models and ELAPSE exposure models for PM<sub>2.5</sub>, NO<sub>2</sub>, and BC. Correlations with local dispersion models were moderate for PM<sub>2.5</sub>, ranging from 0.55 to 0.65. Correlations with local dispersion models were high for NO<sub>2</sub>,

ranging from 0.78 to 0.83. In the Dutch and Roman cohorts, correlation with local ESCAPE models was assessed. In both cohorts, the correlation was higher for NO<sub>2</sub> than for PM<sub>2.5</sub>. In three cohorts the correlation with BC was assessed and it ranged from 0.65 to 0.83.

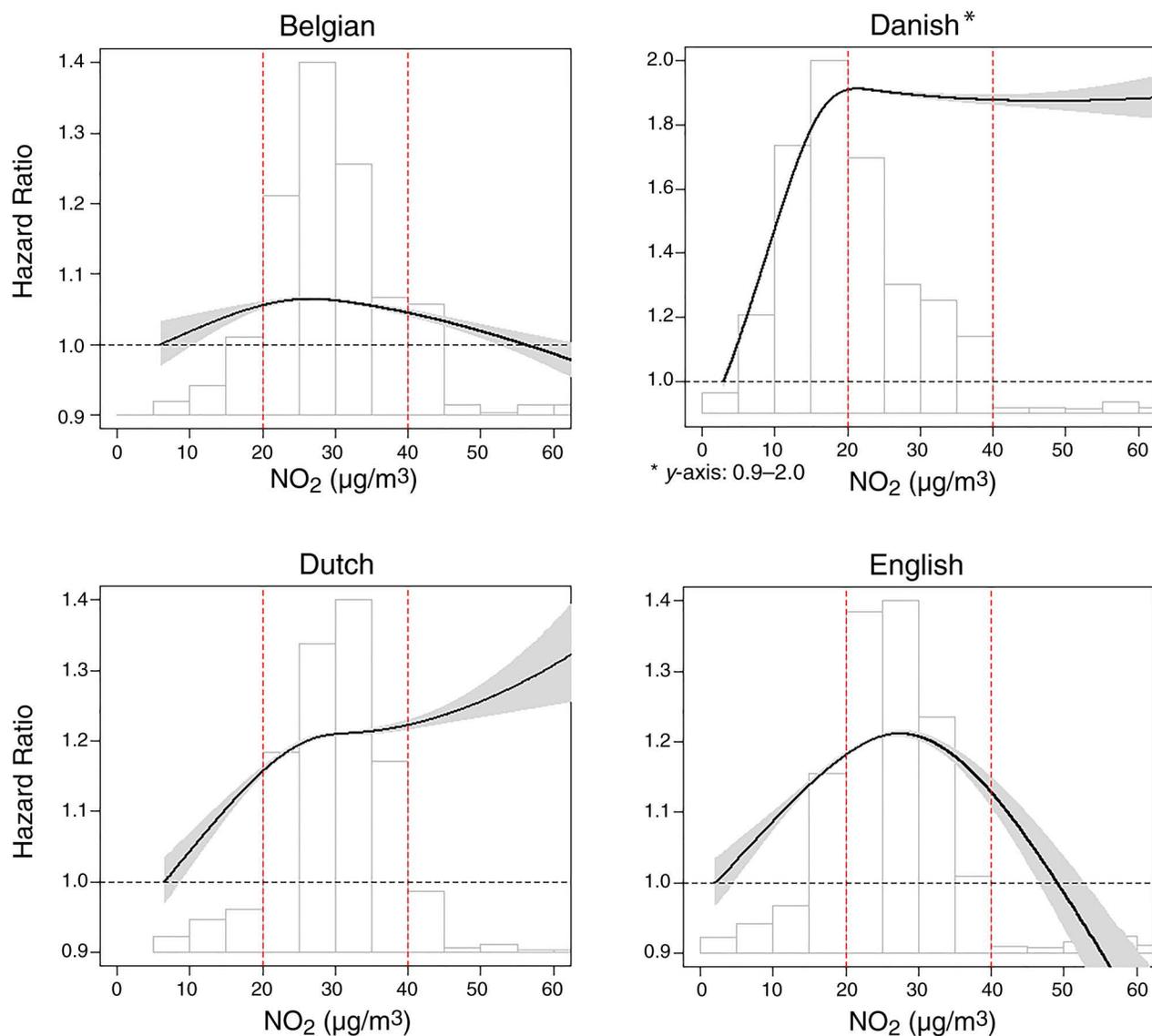


(Figure 17 Continued.)

Table 25 compares HRs for the ELAPSE model and local exposure dispersion models for natural-cause mortality. Appendix 3, Tables A34–A36 show HRs for cardiovascular, respiratory, and lung cancer mortality.

PM<sub>2.5</sub> HRs for natural-cause mortality did not show a consistent difference between the ELAPSE and local model across cohorts. HRs were smaller for the local model compared with the ELAPSE model for the Belgian, Swiss,

and Roman cohorts. For the Roman cohort, HRs were significant for both exposure models and the magnitude differed, probably because of much smaller variation of predicted exposure on an absolute scale by the ELAPSE model. When the same health outcome variation is related to a larger range of exposure, smaller regression slopes result. In contrast, PM<sub>2.5</sub> HRs were higher for the local model for the Dutch, Danish, and Norwegian cohorts. In

NO<sub>2</sub>

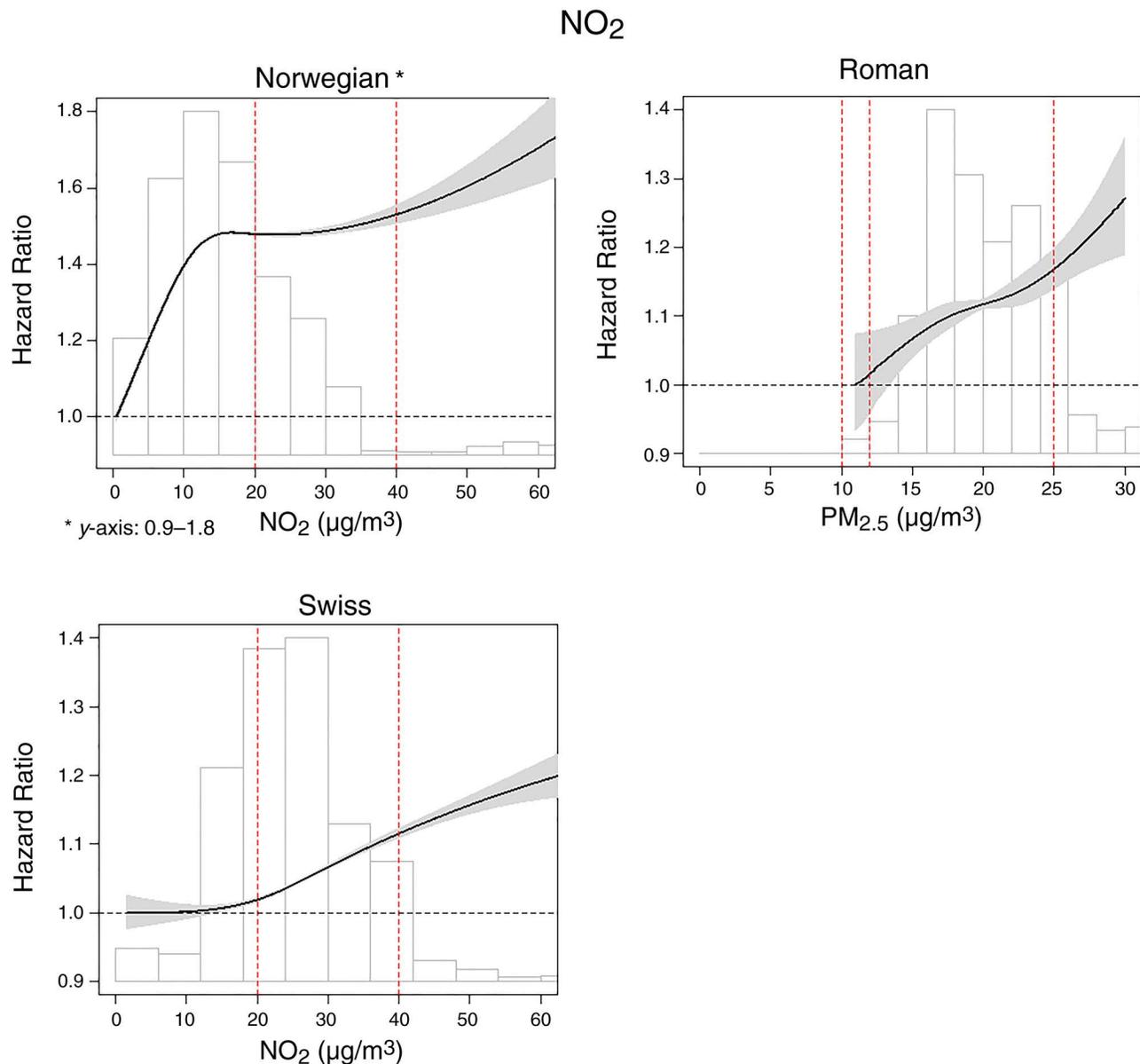
(Figure 17 Continued.)

the Norwegian cohort a dispersion model was available but only for the more urban areas.

For NO<sub>2</sub>, HRs were more similar for the local and ELAPSE model, consistent with the higher correlation compared with PM<sub>2.5</sub>. For BC, in most cohorts HRs were smaller and less significant for the local models.

For cardiovascular, respiratory, and lung cancer mortality, we also observed different HRs for the ELAPSE

versus the local dispersion models (Tables A34–A36). In some analyses, higher HRs were found for the ELAPSE model (e.g., PM<sub>2.5</sub> and respiratory mortality in the Belgian cohort). In other analyses, higher HRs were found for the local model (e.g., PM<sub>2.5</sub> and lung cancer mortality in the Belgian cohort). In the Roman cohort, the magnitude of HRs differed, but the pattern of associations was similar.



(Figure 17 Continued.)

Overall, no clear difference in associations was found comparing the ELAPSE and local models. Associations differed less in the cohorts with relatively strong associations (e.g., Danish and Norwegian cohort PM<sub>2.5</sub> findings).

**Additional Adjustment for Area of the Country** Figure A16 in Appendix 3 shows the results of additional analyses to assess impact of adjustment for region of the country for natural-cause mortality. In the Swiss, Norwegian, and Danish cohorts, HRs were robust to further

adjustment for large area indicator variables for PM<sub>2.5</sub>, NO<sub>2</sub>, and BC. In the Dutch cohort, the HRs for NO<sub>2</sub> and BC were robust, but the PM<sub>2.5</sub> HRs were reduced to unity in all models with a large area indicator. In the Belgian cohort, HRs increased with further adjustment for large area. CIs in the models with cluster option for neighborhood were modestly wider than for the simple Cox model. In the Roman cohort, no adjustment for large area was needed. HRs and CIs were almost identical in the models with and without a cluster option for neighborhood.

**Traffic Noise Adjustment** Further adjustment of the cardiovascular mortality associations for traffic noise did not result in changes of the air pollution effect estimates in the Dutch and Roman cohorts (Table A37). In the Norwegian cohort, HRs were slightly attenuated upon noise adjustment. In the Swiss cohort, associations were attenuated substantially, with only a borderline significant association for BC remaining. In the Belgian cohort, HRs were not affected by further adjustment for noise, but noise exposure assessment was only possible for the Brussels capital area. HRs in this strongly reduced population differed from the full cohort, so the Belgian cohort noise adjustment results need to be interpreted with caution.

**Associations with PM<sub>2.5</sub> Exposure from MAPLE** PM<sub>2.5</sub> exposures estimated by the MAPLE model were slightly lower and less variable than estimated by the ELAPSE model for the Dutch and Norwegian cohorts, lower and more variable in the Swiss, Danish, and Belgian cohort and almost identical for the Roman cohort (Table A38).

The correlation between MAPLE and ELAPSE PM<sub>2.5</sub> modeled concentrations was moderately high for all cohorts ( $R = 0.53$ – $0.65$ , Table A39), except for the Norwegian cohort. In the Norwegian cohort the correlation was very low ( $R = 0.17$ ).

**Table 25.** Hazard Ratios for Associations Between PM<sub>2.5</sub>, NO<sub>2</sub>, and BC and Natural-Cause Mortality in Seven Administrative Cohorts: ELAPSE Versus Local Exposure Models<sup>a</sup>

Cohort	PM <sub>2.5</sub> HR (95% CI)		NO <sub>2</sub> HR (95% CI)		BC HR (95% CI)		O <sub>3</sub> HR (95% CI)	
	ELAPSE	Local	ELAPSE	Local	ELAPSE	Local	ELAPSE	Local
Belgian (N = 5,474,470)	1.023 (1.011, 1.035)	0.989 (0.977, 1.001)	1.001 (0.995, 1.007)	0.998 (0.992, 1.004)	1.002 (0.995, 1.008)	1.002 (0.998, 1.005)	1.007 (0.998, 1.017)	0.998 (0.991, 1.006)
Danish (N = 3,082,500)	1.141 (1.118, 1.165)	1.316 (1.254, 1.380)	1.108 (1.097, 1.119)	1.148 (1.124, 1.172)	1.085 (1.074, 1.095)	1.076 (1.037, 1.116)	0.910 (0.895, 0.926)	0.872 (0.851, 0.893)
Dutch (N = 10,401,335)	1.021 (0.999, 1.044)	1.035 (1.018, 1.052)	1.030 (1.019, 1.041)	1.015 (1.005, 1.024)	1.030 (1.019, 1.041)	1.018 (1.009, 1.027) <sup>b</sup>	0.983 (0.972, 0.995)	NA
English (N = 1,368,740)	1.023 (1.001, 1.045)	NA	1.031 (1.021, 1.041)	1.012 (1.005, 1.020)	1.017 (1.007, 1.028)	NA	0.932 (0.917, 0.948)	NA
Norwegian (N = 1,001,579)	1.039 (1.020, 1.060)	1.045 (1.018, 1.073)	1.035 (1.022, 1.048)	1.031 (1.019, 1.043)	1.033 (1.019, 1.046)	NA	0.957 (0.934, 0.980)	NA
Roman (N = 1,263,712)	1.066 (1.033, 1.099)	1.017 (1.011, 1.022) <sup>c</sup>	1.028 (1.018, 1.038)	1.024 (1.016, 1.033)	1.031 (1.022, 1.041)	1.004 (0.994, 1.014)	0.964 (0.948, 0.981)	NA
Swiss (N = 4,188,175)	1.026 (1.015, 1.038)	0.998 (0.983, 1.012)	1.050 (1.041, 1.059)	1.057 (1.047, 1.068)	1.057 (1.048, 1.067)	NA	0.946 (0.939, 0.954)	NA

NA = not applicable.

<sup>a</sup> HR (95% confidence interval) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>; NO<sub>2</sub>, 10 µg/m<sup>3</sup>; BC, 0.5 × 10<sup>-5</sup>/m; O<sub>3</sub>, 10 µg/m<sup>3</sup>. Main model adjusted for age, sex, year of baseline visit, and cohort-specific individual- and area-level SES variables. Local is a national dispersion model in all cohorts. Danish local exposure model results pending.

<sup>b</sup> EC (elemental carbon) instead of BC.

<sup>c</sup> “Local” is a local LUR 2007 for NO<sub>2</sub>, ESCAPE for BC, a local dispersion model for PM<sub>2.5</sub>.

Table 26 compares linear associations between PM<sub>2.5</sub> estimated with the ELAPSE model and the MAPLE model for natural-cause and cause-specific mortality. The meta-analytical summary estimates were similar for the two exposure models, with only slightly smaller HRs for the PM<sub>2.5</sub> estimated with the MAPLE model. There were larger differences in both directions for individual cohorts. In the Belgian cohort, HRs were smaller for MAPLE PM<sub>2.5</sub> for natural-cause mortality and similar for cause-specific mortality. In the Danish cohort, HRs were similar across the different causes. In the Dutch cohort, HRs were generally larger for PM<sub>2.5</sub> MAPLE. In the Norwegian cohort, no association was found with MAPLE PM<sub>2.5</sub> and a strong association with ELAPSE PM<sub>2.5</sub>. In the Swiss cohort, associations with MAPLE PM<sub>2.5</sub> were weaker than with ELAPSE PM<sub>2.5</sub> and nonsignificant for natural-cause, cardiovascular, and respiratory mortality. In the Roman cohort, HRs were larger for MAPLE PM<sub>2.5</sub>. Only for lung cancer were associations nonsignificant for the ELAPSE model and highly significant for the MAPLE model.

Natural splines (with three degrees of freedom) for natural mortality and MAPLE exposures are shown in Figure 18. The shapes differ from the PM<sub>2.5</sub> ELAPSE models, consistent with the pattern described for the linear model. SCHIFs for natural-cause mortality and natural splines (with three degrees of freedom) for cardiovascular, respiratory, and lung cancer mortality are shown in Appendix 3, Figures A17–A20.

**Associations with Large Spatial Scale Ozone** O<sub>3</sub> exposures from the DEHM model at a larger spatial scale were modestly lower in all cohorts and generally slightly more variable (per the interquartile range) than the 100 × 100 m ELAPSE estimates (Appendix 3, Table A38). O<sub>3</sub> exposures from the MACC model were lower because they are an annual average. The variability for the MACC model was lower in all cohorts.

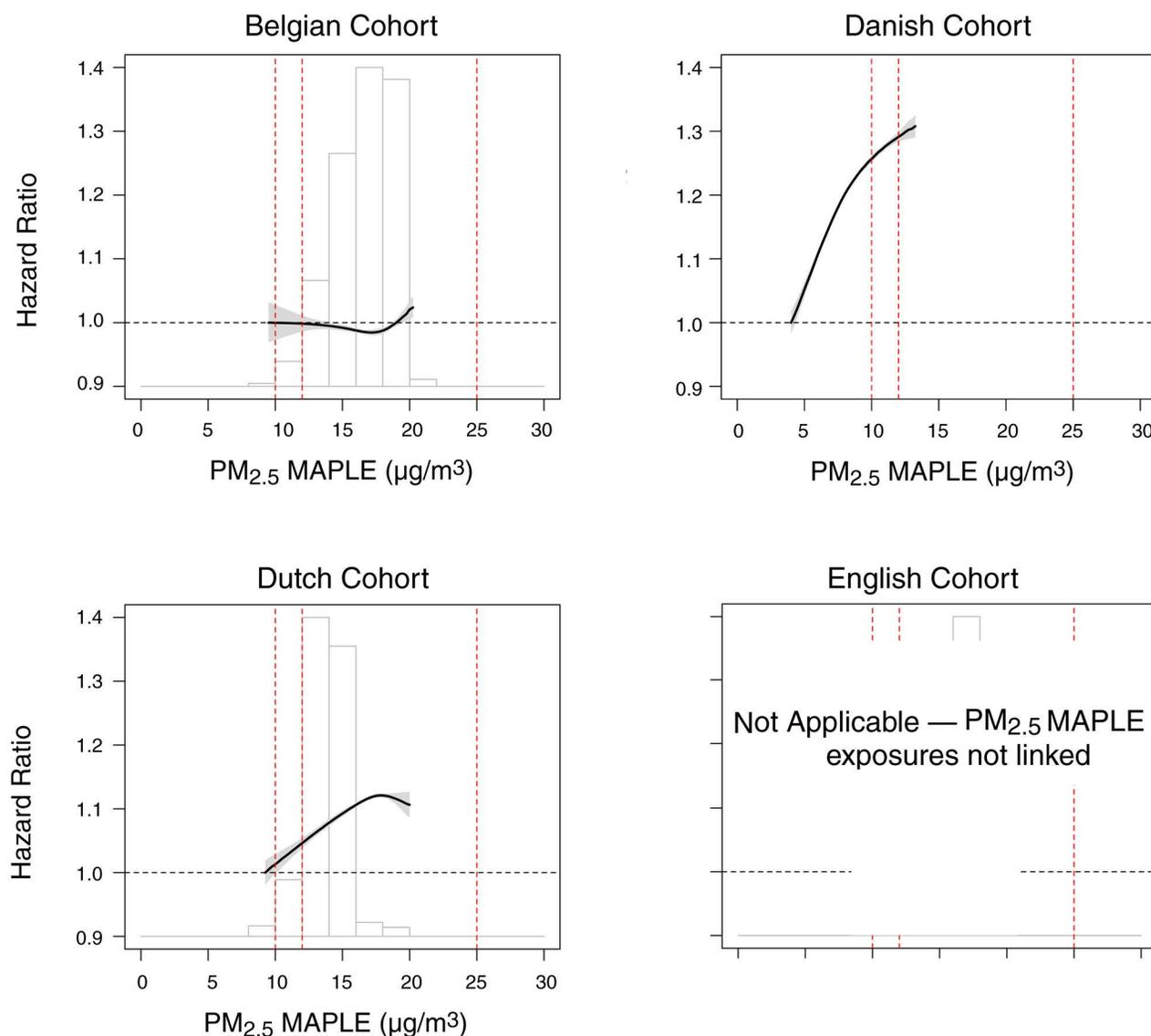
The correlation between the fine-scale ELAPSE estimates and larger scale model estimates was moderate to high in all cohorts (Table A38). The lowest correlation was found in the

**Table 26.** Hazard Ratios for Associations Between PM<sub>2.5</sub> from ELAPSE and MAPLE Exposure Model and Natural-Cause and Cause-Specific Mortality in Administrative Cohorts<sup>a</sup>

	Cohort	Natural Cause HR (95% CI)	Cardiovascular HR (95% CI)	Respiratory HR (95% CI)	Lung Cancer HR (95% CI)
PM <sub>2.5</sub> ELAPSE	Belgian	1.023 (1.011, 1.035)	1.001 (0.983, 1.020)	1.046 (1.016, 1.077)	0.983 (0.950, 1.017)
PM <sub>2.5</sub> MAPLE		1.005 (0.996, 1.015)	0.981 (0.966, 0.996)	1.058 (1.032, 1.084)	1.001 (0.973, 1.030)
PM <sub>2.5</sub> ELAPSE	Danish	1.140 (1.117, 1.163)	1.104 (1.076, 1.133)	1.127 (1.087, 1.169)	1.219 (1.169, 1.272)
PM <sub>2.5</sub> MAPLE		1.128 (1.110, 1.146)	1.068 (1.046, 1.090)	1.143 (1.107, 1.180)	1.202 (1.159, 1.248)
PM <sub>2.5</sub> ELAPSE	Dutch	1.021 (0.999, 1.044)	1.015 (0.988, 1.042)	1.058 (1.009, 1.111)	1.169 (1.121, 1.218)
PM <sub>2.5</sub> MAPLE		1.052 (1.036, 1.068)	1.037 (1.016, 1.059)	1.118 (1.082, 1.155)	1.140 (1.105, 1.175)
PM <sub>2.5</sub> ELAPSE	Norwegian	1.074 (1.063, 1.085)	1.020 (1.005, 1.035)	1.158 (1.127, 1.191)	1.186 (1.146, 1.228)
PM <sub>2.5</sub> MAPLE		0.997 (0.984, 1.010)	1.001 (0.985, 1.018)	0.990 (0.961, 1.020)	0.976 (0.938, 1.016)
PM <sub>2.5</sub> ELAPSE	Roman	1.066 (1.033, 1.099)	1.112 (1.058, 1.168)	1.064 (0.931, 1.215)	1.013 (0.940, 1.091)
PM <sub>2.5</sub> MAPLE		1.113 (1.065, 1.163)	1.202 (1.136, 1.273)	1.111 (0.966, 1.278)	1.157 (1.031, 1.299)
PM <sub>2.5</sub> ELAPSE	Swiss	1.026 (1.015, 1.038)	1.026 (1.008, 1.044)	0.981 (0.953, 1.010)	1.113 (1.078, 1.150)
PM <sub>2.5</sub> MAPLE		1.006 (0.998, 1.014)	0.992 (0.979, 1.005)	1.004 (0.983, 1.026)	1.039 (1.017, 1.062)
PM <sub>2.5</sub> ELAPSE <sup>b</sup>	Meta-analysis	1.058 (1.022, 1.095)	1.043 (1.007, 1.080)	1.073 (1.016, 1.133)	1.110 (1.036, 1.190)
PM <sub>2.5</sub> MAPLE	Meta-analysis	1.047 (1.003, 1.094)	1.041 (0.984, 1.101)	1.064 (1.011, 1.120)	1.079 (1.006, 1.156)

<sup>a</sup> HR (95% CI) presented for the following increments: PM<sub>2.5</sub>, 5 µg/m<sup>3</sup>. Main model adjusted for age, sex, year of baseline visit, and cohort-specific individual- and area-level confounders.

<sup>b</sup> Meta-analysis conducted on six cohorts only (all except English).



**Figure 18.** Natural spline (three degrees of freedom) of the association between natural-cause mortality and PM<sub>2.5</sub> from MAPLE in six administrative cohorts. Red dotted lines are air quality limit and guideline values. Hazard ratios expressed relative to minimum exposure. Histograms show exposure distributions.

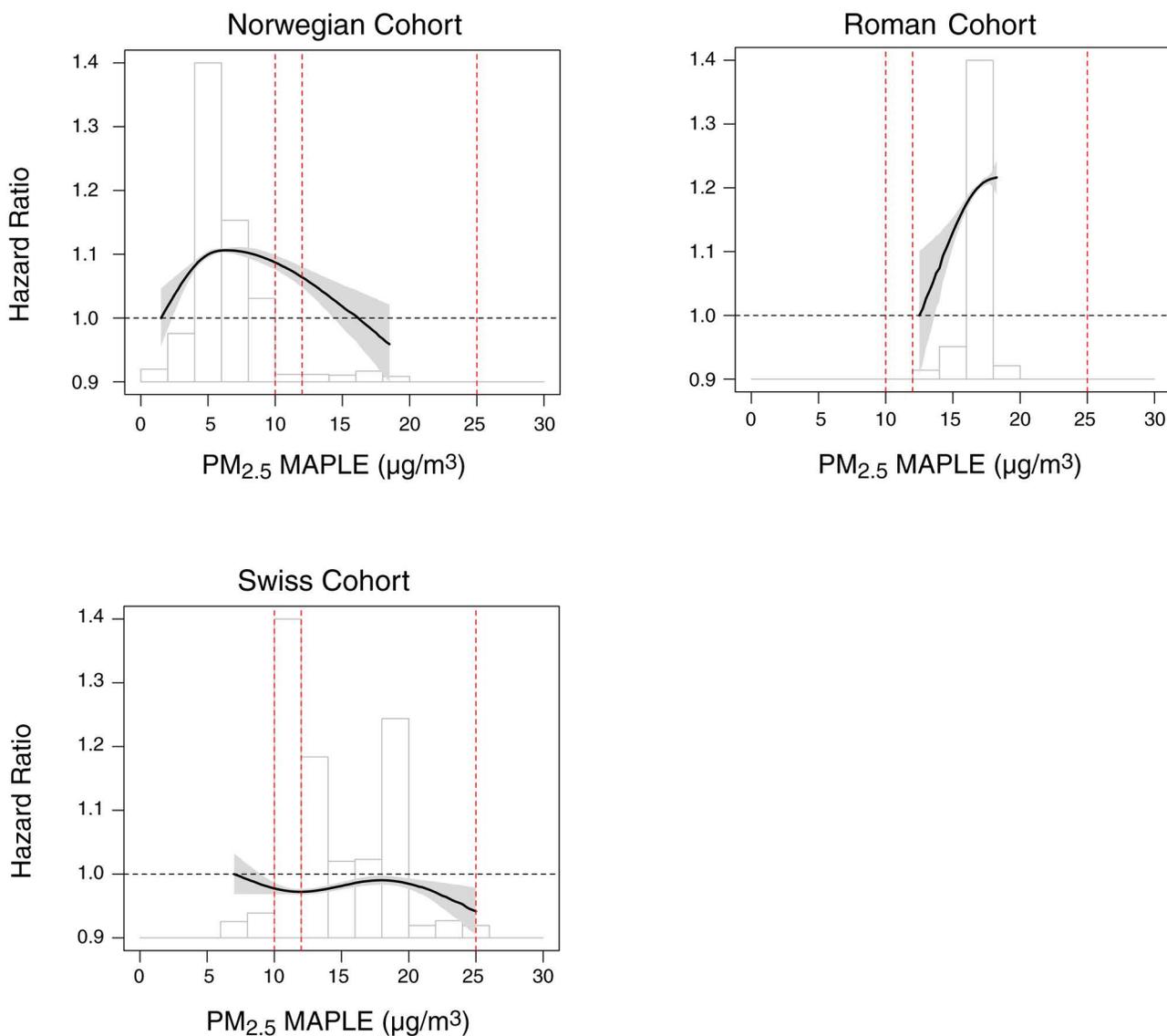
(Figure 18 continues next page.)

Roman cohort, probably because of the small size of the study area compared with the other national cohorts.

The modest difference in estimated exposures and the moderately high correlation between the fine- and large-scale O<sub>3</sub> exposure estimates were consistent with the major contribution of large-scale predictor variables in the ELAPSE model (de Hoogh et al. 2018). Although in the fine-scale model local factors resulted in lower exposures for some subjects, the impact on the overall variability was modest compared with the larger scale predictors. In the

ELAPSE model, the DEHM warm-season prediction was in fact the main predictor.

In Appendix 3, Table A40 compares the effect estimates for O<sub>3</sub> from the three exposure models. Combined effect estimates using meta-analysis were slightly less negative for the large spatial scale O<sub>3</sub> models compared with the fine spatial scale ELAPSE model. In most cohorts, HRs were modestly closer to unity in the large-scale O<sub>3</sub> exposure variables compared with the fine-scale ELAPSE model for natural-cause mortality. In the Dutch cohort,



(Figure 18 Continued.)

associations were more negative for the large-scale O<sub>3</sub> exposures. In the Belgian cohort, HRs were small (both above and below unity) with no clear differences between fine and large spatial scale O<sub>3</sub>.

**Effect Modification** In all cohorts, effect estimates were much stronger for the nonelderly (Table 27). In the Belgian, Dutch, English, and Swiss cohorts, HRs were substantially increased among the nonelderly (HR for PM<sub>2.5</sub> between 1.07 and 1.10) and HRs of essentially unity for PM<sub>2.5</sub> in the elderly. In the Roman, Danish and Norwegian cohorts, HRs in the elderly were statistically higher than unity. The het-

erogeneity of effect estimates across cohorts within the nonelderly was much smaller than in the full population (Figure 14). In the Swiss cohort, HRs for PM<sub>2.5</sub> and NO<sub>2</sub> were stronger for the higher educated and divorced subjects (Appendix 3, Table A41).

#### Summary of Mortality Findings in Administrative Cohorts

Virtually all subjects had 2010 PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below the EU limit values. More than 3.9 million and 1.9 million subjects had a residential PM<sub>2.5</sub> exposure below the U.S. EPA NAAQS (12 µg/m<sup>3</sup>)

and the WHO guideline (10  $\mu\text{g}/\text{m}^3$ ), respectively. We found significant positive associations between  $\text{PM}_{2.5}$ ,  $\text{NO}_2$ , and BC and natural-cause, respiratory, cardiovascular, and lung cancer mortality, with moderate to high heterogeneity between cohorts. In our main model meta-analysis, the HRs (95% CI) were 1.05 (CI = 1.02, 1.09) for an increase of 5  $\mu\text{g}/\text{m}^3$   $\text{PM}_{2.5}$ , 1.04 (CI = 1.02, 1.07) for an increase of 10  $\mu\text{g}/\text{m}^3$   $\text{NO}_2$ , 1.04 (CI = 1.02, 1.06) for an increase of  $0.5 \times 10^{-5}/\text{m}$  BC, and 0.95 (CI = 0.93, 0.98) for an increase of 10  $\mu\text{g}/\text{m}^3$   $\text{O}_3$  for natural-cause mortality.

The shape of the concentration–response functions differed between cohorts, though the associations were generally linear to supralinear, with no indication of a level below which no associations were found. Subset analyses documented that these associations remained at low

levels: below 10  $\mu\text{g}/\text{m}^3$  for  $\text{PM}_{2.5}$  and 20  $\mu\text{g}/\text{m}^3$  for  $\text{NO}_2$ . BC and  $\text{NO}_2$  remained significantly associated with mortality in two-pollutant models with  $\text{PM}_{2.5}$  and  $\text{O}_3$ . The  $\text{PM}_{2.5}$  HR reduced to unity in a two-pollutant model with  $\text{NO}_2$ . The negative  $\text{O}_3$  association was attenuated to unity and became nonsignificant. The mortality associations were robust to alternative specifications of exposure, including time-varying exposure analyses. Effect estimates obtained with the ELAPSE  $\text{PM}_{2.5}$  model did not differ from the MAPLE  $\text{PM}_{2.5}$  model on average, but in some individual cohorts substantial differences were found. Time-varying exposure natural spline analyses confirmed associations at low pollution levels. Effect estimates in the younger participants (<65 years at baseline) were much larger than in the elderly who were >65 years at baseline.

**Table 27.** Hazard Ratios for Associations Between Air Pollution and Natural-Cause Mortality: Effect Modification with Age in Seven Administrative Cohorts

Age Group (years)	HR (95% CI) $\text{PM}_{2.5}$	HR (95% CI) $\text{NO}_2$
<b>Belgian Cohort</b>		
<65	1.068 (1.043, 1.094)	1.063 (1.052, 1.073)
>65	1.009 (0.995, 1.022)	0.984 (0.978, 0.991)
<b>Danish Cohort</b>		
<65	1.265 (1.228, 1.303)	1.169 (1.154, 1.184)
>65	1.080 (1.059, 1.103)	1.078 (1.067, 1.090)
<b>Dutch Cohort</b>		
<65	1.064 (1.033, 1.096)	1.049 (1.034, 1.064)
>65	1.016 (0.991, 1.042)	1.024 (1.012, 1.036)
<b>English Cohort</b>		
<65	1.080 (1.034, 1.127)	1.068 (1.049, 1.088)
>65	1.010 (0.987, 1.034)	1.023 (1.012, 1.034)
<b>Norwegian Cohort</b>		
<65	1.095 (1.077, 1.114)	1.084 (1.071, 1.097)
>65	1.070 (1.059, 1.081)	1.057 (1.048, 1.065)
<b>Roman Cohort</b>		
<65	1.099 (1.034, 1.169)	1.041 (1.022, 1.060)
>65	1.054 (1.019, 1.091)	1.023 (1.013, 1.033)
<b>Swiss Cohort</b>		
<65	1.084 (1.062, 1.106)	1.125 (1.112, 1.138)
>65	1.009 (0.997, 1.021)	1.028 (1.019, 1.036)

**OVERVIEW OF EVIDENCE FROM THE POOLED AND ADMINISTRATIVE COHORTS**

We judged that providing a combined quantitative analysis of effect estimates from the pooled cohort and the administrative cohorts was not informative, given the qualitative and quantitative differences between the pooled cohort on the one hand, and the administrative cohorts on

the other hand. A comparison of the robustness of the main findings in the two sets of cohorts was thought to be more useful. Table 28 provides a summary of the main design features and the mortality findings of the pooled cohort and administrative cohorts. Table 29 gives an overview of the incidence findings. The tables are further interpreted in the Discussion chapter.

**Table 28.** Overview of Main Design Features and Findings in Pooled Cohort and Administrative Cohorts' Mortality Analyses

Outcome	Administrative Cohorts Mortality	Pooled Cohort Mortality
Studies	Six large national cohorts in Belgium, Denmark, England, Norway, the Netherlands, Switzerland and a cohort in Rome	Eight cohorts in Stockholm area, Denmark, the Netherlands, German Ruhr area and Augsburg area, France, Vorarlberg (Austria)
Population (N)	28,153,138 (3,593,741 deaths)	325,367 (47,131 deaths)
Follow-up	Variable, but most 2001–2013 (most recent started in 2011)	Variable, but most 1995–2013 (most recent started in 2001)
Exposures	Substantial fraction with PM <sub>2.5</sub> exposure below 10 µg/m <sup>3</sup> , especially in Norway	Small fraction with PM <sub>2.5</sub> exposure below 10 µg/m <sup>3</sup>
Spatial scale	Most nationwide, exposure contrast due to local, urban- and regional scale	Most smaller regions of a country, exposure contrast due to local and urban scale
Confounders	No individual lifestyle <sup>a</sup> , more detailed individual-level SES and area-level SES	Individual lifestyle and area-level SES. No individual SES in main analysis
Analysis	Per cohort and meta-analysis of effect estimates	Pooled cohort analyses
Linear models (HR 95% CI)	PM <sub>2.5</sub> 1.05 (1.02, 1.09) <sup>b</sup> NO <sub>2</sub> 1.04 (1.02, 1.07) BC 1.04 (1.02, 1.06) O <sub>3</sub> 0.95 (0.93, 0.98)	PM <sub>2.5</sub> 1.13 (1.11, 1.16) NO <sub>2</sub> 1.09 (1.07, 1.10) BC 1.08 (1.06, 1.10) O <sub>3</sub> 0.90 (0.88, 0.91)
Shape of concentration–response function	Variable across cohorts but generally supralinear with no indication of a level below which no association was found	Generally supralinear with no indication of a level below which no association was found
Subset analysis	Clear evidence of an association below 10 µg/m <sup>3</sup> for PM <sub>2.5</sub> and 20 µg/m <sup>3</sup> for NO <sub>2</sub>	Clear evidence of an association below 10 µg/m <sup>3</sup> for PM <sub>2.5</sub> and 20 µg/m <sup>3</sup> for NO <sub>2</sub>
Two-pollutant model	BC and NO <sub>2</sub> remained significantly associated with mortality. PM <sub>2.5</sub> HR reduced to unity with NO <sub>2</sub> . The negative O <sub>3</sub> association attenuated to unity.	PM <sub>2.5</sub> , BC, and NO <sub>2</sub> remained significantly associated with mortality. The negative O <sub>3</sub> association attenuated but remained significant.
Diabetes mortality	Weak positive statistically nonsignificant association, with lower HRs compared to most other causes of death	Strong positive association, with higher HRs compared to other causes of death
MAPLE PM <sub>2.5</sub>	HRs differed from ELAPSE exposure model, though general pattern of associations broadly similar	HRs mostly consistent with the ELAPSE exposure model
Large spatial scale O <sub>3</sub> exposure	HRs not consistently different from fine scale ELAPSE exposures	HRs not consistently different from fine scale ELAPSE exposures

<sup>a</sup> Except the English cohort.

<sup>b</sup> Natural-cause mortality.

**Table 29.** Overview of Main Design Features and Findings in Pooled Cohort Incidence Analyses

Outcomes	Acute Coronary Heart Disease	Stroke	Lung Cancer	Asthma	COPD
Studies	Six cohorts in Stockholm area, Denmark, the Netherlands, German Ruhr and Augsburg areas	Six cohorts in Stockholm area, Denmark, the Netherlands, German Ruhr and Augsburg areas	Seven cohorts in Stockholm area, Denmark, the Netherlands, German Ruhr area, France, Vorarlberg (Austria)	Three cohorts in Stockholm area and Denmark	Three cohorts in Stockholm area and Denmark
Population (N)	137,148 (10,071 events)	137,148 (6,950 events)	307,550 (3,965 events)	98,326 (1,965 events)	98,508 (4,928 events)
Follow-up	Recruitment: 1992–2004 Follow-up: 2011–2015	Recruitment: 1992–2004 Follow-up: 2011–2015	Variable, but most 1995–2013 (most recent start 2001)	Variable, but most 1995–2013 (most recent start 2001)	Variable, but most 1995–2013 (most recent start 2001)
Confounders	Individual lifestyle, individual SES, and area-level SES	Individual lifestyle, individual SES, and area-level SES	Individual lifestyle and area-level SES	Individual lifestyle, individual SES, and area-level SES	Individual lifestyle, individual SES, and area-level SES
Linear models (HR 95% CI)	PM <sub>2.5</sub> 1.02 (0.95, 1.10) NO <sub>2</sub> 1.04 (1.01, 1.07) BC 1.02 (0.99, 1.06) O <sub>3</sub> 0.94 (0.90, 0.98)	PM <sub>2.5</sub> 1.10 (1.01, 1.21) NO <sub>2</sub> 1.08 (1.04, 1.12) BC 1.06 (1.02, 1.10) O <sub>3</sub> 0.96 (0.91, 1.01)	PM <sub>2.5</sub> 1.13 (1.05, 1.23) NO <sub>2</sub> 1.02 (0.97, 1.07) BC 1.02 (0.97, 1.07) O <sub>3</sub> 0.95 (0.89, 1.02)	PM <sub>2.5</sub> 1.22 (1.04, 1.43) NO <sub>2</sub> 1.17 (1.10, 1.25) BC 1.15 (1.08, 1.23) O <sub>3</sub> 0.90 (0.81, 0.99)	PM <sub>2.5</sub> 1.17 (1.06, 1.29) NO <sub>2</sub> 1.11 (1.06, 1.16) BC 1.11 (1.06, 1.15) O <sub>3</sub> 0.99 (0.93, 1.05)
Shape of concentration response function	Sublinear for NO <sub>2</sub> with no association below 20 µg/m <sup>3</sup> ; supralinear for PM <sub>2.5</sub>	Supralinear for PM <sub>2.5</sub> , BC, and NO <sub>2</sub>	Linear to supralinear for PM <sub>2.5</sub> , BC, and NO <sub>2</sub>	Linear to supralinear for PM <sub>2.5</sub> , BC, and NO <sub>2</sub>	Linear to supralinear for PM <sub>2.5</sub> , BC, and NO <sub>2</sub>
Subset analysis	Evidence of an association below 30 and 20 µg/m <sup>3</sup> for NO <sub>2</sub>	Evidence of an association below 15 and possibly 12 µg/m <sup>3</sup> for PM <sub>2.5</sub> and 30 and possibly 20 µg/m <sup>3</sup> for NO <sub>2</sub>	Evidence of an association below 20 and possibly 12 µg/m <sup>3</sup> for PM <sub>2.5</sub>	Evidence of an association below 20 and possibly 12 µg/m <sup>3</sup> for PM <sub>2.5</sub> and below 30 and possibly 20 µg/m <sup>3</sup> for NO <sub>2</sub>	Evidence of an association below 15 and possibly 12 µg/m <sup>3</sup> for PM <sub>2.5</sub> and below 30 µg/m <sup>3</sup> for NO <sub>2</sub>
Two-pollutant model	NO <sub>2</sub> remained significant with PM <sub>2.5</sub> and BC, but not with O <sub>3</sub> . The negative O <sub>3</sub> association attenuated and just turned non-significant.	PM <sub>2.5</sub> reduced to unity with NO <sub>2</sub> . NO <sub>2</sub> remained significant in all models. The negative O <sub>3</sub> association changed to a nonsignificant positive association with NO <sub>2</sub> .	PM <sub>2.5</sub> remained significant in all models. The negative O <sub>3</sub> association changed to unity with PM <sub>2.5</sub> .	PM <sub>2.5</sub> HR reduced to unity with NO <sub>2</sub> and BC. NO <sub>2</sub> and BC remain significant with PM <sub>2.5</sub> and O <sub>3</sub> . The negative O <sub>3</sub> association changed to a nonsignificant positive association with NO <sub>2</sub> and BC.	PM <sub>2.5</sub> HR reduced to unity with NO <sub>2</sub> and BC. NO <sub>2</sub> and BC remain significant with PM <sub>2.5</sub> and O <sub>3</sub> . The negative O <sub>3</sub> association changed to a significant positive association in models with NO <sub>2</sub> and BC.
MAPLE PM <sub>2.5</sub>	Associations stronger for MAPLE PM <sub>2.5</sub>	Associations stronger for MAPLE PM <sub>2.5</sub>	Associations weaker for MAPLE PM <sub>2.5</sub>	Associations consistent with MAPLE PM <sub>2.5</sub>	Associations consistent with MAPLE PM <sub>2.5</sub>
Large spatial scale O <sub>3</sub> exposure	Associations null similar with large-scale model	Associations positive with large-scale model	Associations more negative with large-scale model	Associations positive with large-scale model	Associations positive with large-scale model

---

## DISCUSSION AND CONCLUSIONS

---

### MAIN FINDINGS

In the pooled cohort, virtually all subjects in 2010 had PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below the annual EU limit values of 25 and 40 µg/m<sup>3</sup>, respectively. More than 50,000 and 25,000 subjects had a residential PM<sub>2.5</sub> exposure below the U.S. EPA NAAQS (12 µg/m<sup>3</sup>) and the WHO guideline (10 µg/m<sup>3</sup>), respectively. We found significant positive associations between PM<sub>2.5</sub>, NO<sub>2</sub>, and BC and natural-cause, respiratory, cardiovascular, and diabetes mortality. For these pollutants, we generally observed a supralinear association with steeper slopes at low exposures and no evidence of a threshold below which no association was found. Subset analyses documented that these associations remained at low levels — below 10 µg/m<sup>3</sup> for PM<sub>2.5</sub> and 20 µg/m<sup>3</sup> for NO<sub>2</sub> — and HRs were similar to the HRs in the full cohort and statistically significant. The mortality associations were robust to other specifications of exposure, including different time periods, the PM<sub>2.5</sub> exposure estimated from the MAPLE project, and the estimates from the local ESCAPE model. The highest HRs were found for diabetes mortality. In two-pollutant models, associations of PM<sub>2.5</sub>, NO<sub>2</sub>, and BC were attenuated but remained statistically significant. Associations with O<sub>3</sub> were negative, both in the fine spatial scale of main ELAPSE model and in large-scale exposure models.

We found significant positive associations between PM<sub>2.5</sub>, NO<sub>2</sub>, and BC and incidence of stroke, asthma, and COPD hospital admissions. NO<sub>2</sub> was furthermore significantly related to acute coronary heart disease and PM<sub>2.5</sub> with lung cancer incidence. We generally observed linear to supralinear associations with no evidence of a threshold, with the exception of the association between NO<sub>2</sub> and acute coronary heart disease, which was sub-linear. Subset analyses documented that PM<sub>2.5</sub> associations remained below 20 µg/m<sup>3</sup> and in some cases 12 µg/m<sup>3</sup>. NO<sub>2</sub> associations remained below 30 and possibly 20 µg/m<sup>3</sup>. In two-pollutant models, NO<sub>2</sub> was most consistently associated with acute coronary heart disease, stroke, asthma, and COPD hospital admissions. PM<sub>2.5</sub> was not associated with these outcomes in two-pollutant models. PM<sub>2.5</sub> was the only pollutant that was associated with lung cancer incidence in two-pollutant models. Associations with O<sub>3</sub> were negative, though not statistically significant in two-pollutant models.

In the administrative cohorts, virtually all subjects had 2010 PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below the EU limit values. More than 3.9 million and 1.9 million subjects had a residential PM<sub>2.5</sub> exposure below the U.S. EPA

NAAQS (12 µg/m<sup>3</sup>) and the WHO guideline (10 µg/m<sup>3</sup>), respectively. We found significant positive associations between PM<sub>2.5</sub>, NO<sub>2</sub>, and BC and natural-cause, respiratory, cardiovascular, and lung cancer mortality, with moderate to high heterogeneity of effect estimates between cohorts. We found weak associations with diabetes and stroke mortality. The shape of the association differed between cohorts, though associations were generally linear to supralinear with no indication of a level below which no associations were found. There was clear evidence of an association below 10 µg/m<sup>3</sup> for PM<sub>2.5</sub> and below 20 µg/m<sup>3</sup> for NO<sub>2</sub>. BC and NO<sub>2</sub> remained significantly associated with mortality in two-pollutant models with PM<sub>2.5</sub> and O<sub>3</sub>. The PM<sub>2.5</sub> HR reduced to unity in a two-pollutant model with NO<sub>2</sub>. The negative O<sub>3</sub> association was attenuated to unity and became nonsignificant after adjustment for NO<sub>2</sub>. Effect estimates in participants who were <65 years at baseline were much larger than in the elderly in all cohorts. Effect estimates obtained with the ELAPSE PM<sub>2.5</sub> model did not differ from the MAPLE PM<sub>2.5</sub> model on average, but in individual cohorts substantial differences were found.

### ASSOCIATIONS AT LOW POLLUTION LEVELS: CONCENTRATION–RESPONSE FUNCTIONS

We combined spline, SCHIF, subset, and threshold analyses to assess the shape of the concentration–response functions. Each method has its merits and problems. In our study, the methods generally resulted in the same conclusion about the shape of the concentration–response function, particularly the absence of evidence for a level below which no associations were found. We did observe that the spline functions tended to be less smooth than the more parametric SCHIF function. Furthermore, the uncertainty of the prediction differed between the spline and SCHIF method: with larger uncertainty at the tails of the distribution for the spline and the lowest uncertainty near the minimum exposure for the SCHIF method. In the HEI commentary on the MAPLE interim report, the small uncertainty of SCHIF in the low end of the exposure distribution was attributed to the choice of the reference value as the minimum exposure instead of the median exposure (Brauer et al. 2019). Because of harmonization with the MAPLE study, we used the same SCHIF script. We consider that the uncertainty in the splines better reflects the distribution of the pollutant incorporated in its estimation. Spline plots incorporate the standard error of the estimated function in the curve uncertainty, which is wider where data are fewer. As we compared with the minimum pollutant concentration, we basically subtracted the predicted HR at each level from that in the minimum level.

A SCHIF plot provided wider uncertainty as we moved further from the minimum pollutant value (used as the reference point) following the estimation of the CI. A SCHIF simulates 1,000 draws from a normal distribution with mean equal to the parametric model's coefficient and variance based on the corresponding standard error. The procedure then calculates HRs (according to the selected monotonic parametric shape) for each draw for pollutant's levels using the minimum as reference. For each pollutant's level then the SCHIF plots the mean of the HR (over 1,000 draws) and estimates the upper CI as the 2.5th percentile of 1,000 HRs and the lower CI as the 97.5th percentile of 1,000 HRs. Hence, where the distance from the minimum is smaller, the HR and the CIs, as defined above, are closer. Although the choice of the minimum exposure as reference value results in unrealistic CIs at the low end of the exposure distribution, the shape of the curve is not affected. The use of the median as reference value would result in HRs below 1 for the lower half of the exposure distribution for most of the functions, which is intuitively more difficult to appreciate.

For assessment of the shape of the concentration–response function, the nonparametric nature of the splines may have an advantage compared with approaches borrowing information from assumptions about the shape of the function. For health impact assessments, the biologically more plausible SCHIF functions are more attractive than the spline functions (Burnett et al. 2018; Nasari et al. 2016).

Patterns were consistent between splines with two, three, or four degrees of freedom. Especially in the pooled cohort, the high uncertainty at the tails of the distribution did not allow statements about the shape of the concentration–response function in the spline analysis at very low levels. Consistently, the threshold models showed no difference in performance between models with, for example, a  $PM_{2.5}$  threshold at 5 and  $7.5 \mu\text{g}/\text{m}^3$  versus a linear model in the pooled cohort.

Penalized splines showed similar patterns as natural splines, but the patterns were less smooth. The patterns of the penalized splines probably indicated some overfitting. This comparison was between natural splines with a fixed degree of freedom and penalized splines with the degrees of freedom determined by the AIC. A larger number of degrees of freedom (typically six) was selected for the penalized splines.

Our main analyses were based upon the modeled exposure of the year 2010, which for the pooled cohort especially is an exposure toward the end of follow-up (which ranged from 2011 to 2017). We note that in the pooled cohort a large number of events occurred in the last years of follow-up. From 2010–2017, 33% of all deaths occurred;

from 2005–2017, 63% of all deaths, and from 2000–2017, 84% of all deaths. In the administrative cohorts, follow-up started from 2000–2001 in five cohorts and 2008 and 2011 in the Dutch and English cohorts, respectively. The main exposure was therefore more in the middle of the follow-up period. The documented stability of the spatial pollution surfaces in our region (de Hoogh et al. 2018) further supports the use of the 2010 exposure metric. Sensitivity analyses using exposure at baseline or time-varying exposure generally corroborated the pattern of findings for  $PM_{2.5}$ , BC, and  $NO_2$  in our main model, although the magnitude of the effect estimates differed depending on the exposure time window. For example, substantially lower mortality significant HRs were found for the baseline exposure than for the 2010 exposure in the pooled cohort. In the time-varying exposure analysis we incorporated residential history, which was available for a large number of subjects. The HRs from the time-varying analyses using a 1-year moving average were similar to the 2010 main model HRs. The findings of an analysis with start of follow-up in 2000, 2005, and 2008 further supported our main findings, albeit with effect estimates that were lower than in our main analysis. We acknowledge that although temporal misalignment of exposure assignment was less in this analysis, selection bias may have played a role as we only analyzed the survivors of the original cohort. Also, with a later start of follow-up, deaths occurred at a later age. We found that HRs generally were smaller in subjects with higher age at baseline, so this is another explanation. Finally, for  $NO_2$  and  $O_3$  we applied models based upon monitoring data for the year 2000 and 2005 and found significant HRs with a magnitude in between the main and baseline exposure analysis. The importance of the analysis based upon 2000 and 2005 models is that they did not rely on back-extrapolation. For  $PM_{2.5}$  this analysis was not possible. In the MAPLE study, time-varying analyses were conducted with 1-year, 3-year, and 8-year moving averages (Brauer et al. 2019). HRs increased somewhat with longer averaging times, but the pattern of associations was the same and the authors noted high correlations between the different time periods (Brauer et al. 2019). Earlier work in the DCH study found small differences between different exposure periods for mortality, attributed to high correlations among different time periods. It is not straightforward to determine which exposure period is most relevant in our study. The assumed latency time is an important determinant. There is some evidence that for (cardiovascular) mortality more recent exposures are most important, whereas for lung cancer exposure over a longer period is most relevant. We did not investigate different time-varying exposure periods, but will do so in an extension of the current project.

The observation of associations at low pollution levels was also found in natural spline analyses in the time-varying exposure model. In this model, we no longer relied on the use of the relatively low 2010 concentrations (compared with earlier time periods) but evaluated associations of mortality and exposure in the year of the event. The small shift of the concentration–response curves toward higher concentrations, despite the secular trend in air pollution, was related to the large number of events toward the end of follow-up.

An important issue in both the spline, SCHIF, and subset analyses is that the populations in the low and high exposure ranges may have differed. In the administrative cohorts, differences may have been related to region of the country, urbanicity, and neighborhood characteristics. In the pooled cohort, differences may have been related to country, but also urbanicity and neighborhood characteristics. The pooled cohort included general population samples but also specific subpopulations such as teachers and nurses. As we used strata for subcohorts in the pooled cohort analyses and analyzed the administrative cohorts individually, confounding was unlikely. Potentially different sensitivity of these populations to air pollution may affect the observed shape of the concentration–response function. Figures 2 and 3 and Figures P1 and P2 in Appendix 2 illustrate that differences in exposure between cohorts occurred especially for PM<sub>2.5</sub>. As we observed supralinear concentration–response curves for PM<sub>2.5</sub> but also for NO<sub>2</sub> and BC, it seems unlikely that the shape of the curve is fully determined by different populations contributing to specific parts of the exposure range. The cohorts with the lowest PM<sub>2.5</sub> exposures were general population samples (Swedish CEANS). The next lowest were general population (the large DCH) and a specific population group (Danish nurse cohort). In future work we intend to evaluate this issue more by using causal modeling methods.

In the pooled cohort, we excluded about 15% of the subjects from the main analysis due to missing covariate data. We documented that linear associations were not affected by selection bias using multiple imputation and a comparison of HRs in the full and complete case study population. As the loss of subjects was even less in the (North European) cohorts with the lowest exposures, the impact on low-level associations is likely very limited. In the administrative cohorts, the fraction of excluded subjects was smaller than in the pooled cohort and our sensitivity analysis found no indication of selection bias. For the interpretation of the low-level analysis, it is particularly important that in the cohort with the lowest exposure levels — the Norwegian cohort — HRs were identical in the full and complete case study population.

#### MORTALITY ASSOCIATIONS IN THE POOLED VERSUS THE ADMINISTRATIVE COHORTS

The combined effect estimates for PM<sub>2.5</sub>, NO<sub>2</sub>, and BC in the administrative cohorts were smaller than the effect estimates from the pooled cohort. Despite the much larger number of subjects and person-years in the administrative cohorts, the widths of the CIs in the meta-analyses were slightly larger than in the pooled cohort analysis. The wider CIs are likely due to the heterogeneity in effect estimates between the seven cohorts, which is incorporated in the random effects meta-analysis. For five of the individual administrative cohorts, the width of the CI of the cohort-specific effect estimate was smaller than the width of the CI for the combined meta-analytic summary.

The smaller effect estimate in the administrative cohorts is interesting because an implicit assumption is that lack of adjustment for individual-level confounders, such as smoking, may lead to an overestimation of air pollution risks. In the Canadian MAPLE study, effect estimates for PM<sub>2.5</sub> and natural-cause mortality were also larger in the Canadian Community Health Survey that had individual lifestyle data available than the effect estimates in the large census cohort without lifestyle data (Brauer et al. 2019). In the recent systematic review of the association between PM<sub>2.5</sub> and mortality by WHO, the overall summary effect estimates were not affected by exclusion of all administrative cohorts that did not have individual lifestyle data available. HRs per 10 µg/m<sup>3</sup> were 1.08 (CI = 1.06, 1.09) for all studies and 1.08 (CI = 1.05, 1.10) for the studies with individual lifestyle data only (Chen and Hoek 2020). For the entire body of evidence there is thus no systematic difference in effect estimates between the two types of studies, which is probably related to the different direction of associations between air pollution exposure and lifestyle risk factors at different sites or countries. In a large Dutch survey, higher air pollution exposures were associated weakly with more unfavorable lifestyles, such as smoking (Strak et al. 2017), whereas in the national Canadian population the opposite was found (Shin et al. 2014). In the ESCAPE study, we observed that adjustment for smoking attenuated effect estimates (Beelen et al. 2014a). In a recent review, additional adjustment for area-level SES tended to increase the effect estimates (Vodonos et al. 2018). The different direction of the association between air pollution and lifestyle is likely related to socioeconomic factors and the historic development of cities and countries. As an example, in Rome the (more polluted) historic city center is predominantly inhabited by wealthier people because of the high housing costs.

We note that the effect estimates in the administrative cohorts for PM<sub>2.5</sub> and NO<sub>2</sub> are more consistent with recent

meta-analyses (Chen and Hoek 2020; Pope et al. 2020). The smaller effect estimates in ELAPSE for the administrative cohorts could be due to the more recent recruitment date. The 2010 exposure is lower toward the end of follow-up for the administrative cohorts than for the pooled cohort. In the pooled cohort we observed that using back-extrapolated exposure at baseline resulted in significant associations with effect estimates more similar to the administrative cohorts HRs. In the administrative cohorts, the difference between the 2010 exposure and start of follow-up exposure made less of a difference. The subjects in the administrative cohorts were on average 2 to 9 years older at baseline, compared with the participants in the pooled cohort. In the administrative cohort analyses, effect estimates in subjects <65 years old were substantially larger than in subjects  $\geq 65$  years old. The spatial scale of the study area of the administrative cohorts was larger than those for the individual cohorts making up the pooled cohort. In the pooled cohort primary analysis, we used strata for (sub)cohort and hence did not exploit contrast in exposure between cohorts. In the administrative cohorts, exposure contrast within countries was exploited. If between-city exposure contrasts resulted in smaller risks than within-city contrasts, HRs would be lower in the administrative cohort analysis. Negative confounding by missing lifestyle factors would also explain the lower effect estimates. However, in the different cohorts we observed both increased and decreased HRs after more complete adjustment for socioeconomic factors. Indirect adjustment resulted in both increases and decreases of the HRs. The modestly lower average pollution levels and the more representative study populations would both lead to higher effect estimates in the administrative cohorts, based on the observed supralinear associations and a potentially higher effect estimate in subjects with lower SES, who tend to participate less in cohort studies.

We do not have a clear explanation why in the pooled cohort in two-pollutant models,  $PM_{2.5}$ , BC, and  $O_3$  remained significant whereas in the administrative cohorts' meta-analysis, the  $PM_{2.5}$  HR was reduced to unity after adjustment for  $NO_2$ .

Associations with diabetes mortality were positive but statistically nonsignificant in the administrative cohorts (HR per  $5 \mu\text{g}/\text{m}^3$  for  $PM_{2.5}$  1.038 [CI = 0.974, 1.106]), whereas in the pooled cohort the HRs were large for diabetes mortality (HR per  $5 \mu\text{g}/\text{m}^3$  for  $PM_{2.5}$  1.316 [CI = 1.144, 1.514]). HRs in the pooled cohort were the largest among the evaluated causes. The CI in the pooled cohort was much wider, related to the small number cases (<1,000).

Despite the harmonized methodology, we observed clear heterogeneity in the concentration–response function in the

seven administrative cohorts. Part of the heterogeneity related to differences in recruitment year. The large effect estimate for the Danish cohort in the main analysis was reduced when the 2001 baseline exposure was used.

### COMPARISON OF THE SHAPE OF THE CONCENTRATION–RESPONSE FUNCTION WITH PREVIOUS STUDIES

Our finding of steeper slopes of the  $PM_{2.5}$  mortality associations in the pooled cohort and several administrative cohorts at low pollution levels is consistent with previous administrative cohort studies and a health survey study in Canada and North America (Crouse et al. 2012, 2015; Di et al. 2017; Pinault et al. 2016, 2017). In the interim reports of the Canadian and U.S. HEI-funded low-level air pollution studies, similar findings were obtained (Brauer et al. 2019; Dominici et al. 2019). In some cohort studies with individual lifestyle data, the shape of the  $PM_{2.5}$  function was sublinear (Thurston et al. 2016; Villeneuve et al. 2015) or linear (Jerrett et al. 2013; Pope et al. 2002; Schwartz et al. 2008). In a comprehensive analysis using the SCHIF methodology in a large number of cohorts, a supralinear association was observed as well (Burnett et al. 2018). Our study has added the observation of a supralinear association without a threshold in a large, well-characterized cohort. The administrative cohorts, especially the observations in the Norwegian cohort, provide strong evidence of associations with mortality at pollution levels well below the current limit values and WHO  $PM_{2.5}$  guidelines.

Our finding of linear or supralinear  $PM_{2.5}$  associations with lung cancer incidence is in agreement with two large studies on lung cancer mortality (Bowe et al. 2019; Burnett et al. 2018). Other studies, including studies on lung cancer incidence, suggested sublinear associations with little evidence of an association at the lowest levels (Bai et al. 2020; Pinault et al. 2017).

Our finding of supralinear  $PM_{2.5}$  associations with stroke incidence is in agreement with one large study on hospital admissions of Medicare participants in the Southeast United States (Danesh Yazdi et al. 2019). The authors also reported a steep slope for  $PM_{2.5}$  and acute myocardial infarction, which could not be confirmed for acute coronary heart disease incidence in ELAPSE. A linear association was reported for  $PM_{2.5}$  and acute myocardial infarction across 5.1 million Canadian adults living in Ontario, whereas the slopes were supralinear for  $NO_2$  (flattening above 5–10 ppb) and sub-linear for  $O_3$  (increasing above 45 ppb) (Bai et al. 2019). Also, positive slopes for  $O_3$  and stroke and acute myocardial infarction hospital admissions (though rather wiggly for myocardial infarction)

were reported for the U.S. Medicare study (Danesh Yazdi et al. 2019) but could not be confirmed.

We can only speculate about the reasons for a supra-linear association. The lack of evidence for a threshold at low pollution levels has been attributed to the large difference in individual sensitivity within the population and the absence of a sharp threshold within individuals. A steeper slope at low levels could be explained by a saturation effect. This may apply to the global scale but seems quite unlikely within the North American context of generally low-to-moderate concentrations. Differential measurement error may affect the shape of the observed concentration–response function as well.

### **ASSOCIATIONS WITH PM<sub>2.5</sub> FROM ELAPSE AND MAPLE MODELS**

We evaluated the sensitivity of our findings to applying the modeling method developed in the MAPLE study. The MAPLE method assesses PM<sub>2.5</sub> at a 1 × 1 km spatial resolution primarily based upon satellite AOD (Brauer et al. 2019). The method uses a global chemical transport model (GEOS-Chem) to calculate surface concentrations. A geographically weighted regression approach is used to calibrate the surface concentrations to ground monitoring, incorporating data on land use (van Donkelaar et al. 2015, 2016). The MAPLE project has further refined the method by long-term AOD measurements from the ground at selected stations (Brauer et al. 2019). We used the refined PM<sub>2.5</sub> estimates at 1 × 1 km resolution. In ELAPSE we also used satellite AOD as a predictor in the LUR model. The satellite data were made available to us by van Donkelaar and Martin, who were also responsible for the satellite data in MAPLE. The model included 10 × 10 km surface concentrations for 2009–2011 without the land use calibration (de Hoogh et al. 2018). The new 1 × 1 km satellite data for 2010 were less predictive for PM<sub>2.5</sub> surface monitoring data than the 10 × 10 km data for 2009–2011. The ELAPSE model further includes an ensemble chemical transport model, land use, and traffic variables. The ELAPSE model did not include the refinements added by geographically weighted regression and the ground-based AOD monitoring, which have been shown to improve PM<sub>2.5</sub> exposure estimation significantly in the Canadian setting (Brauer et al. 2019). The two models differ both in spatial scale and structure. The spatial scale difference is likely not too critical as PM<sub>2.5</sub> is dominated by larger scale regional trends. Consistently, we observed that in the ELAPSE model, the large-scale predictors (chemical transport model and satellite AOD) contributed most to the model (de Hoogh et al. 2018).

Considering the difference in method, the two approaches resulted in overall reasonable agreement in absolute levels for 2010. In Denmark and Germany, the MAPLE method predicted lower levels than the ELAPSE method. Routine measurements from AirBase and ESCAPE suggested the MAPLE method likely underestimated the levels.

The correlation between the two models was high across Europe but moderate within the relatively small study areas of the cohorts comprising the pooled cohort and the administrative cohorts. Relatively small variability within study areas likely has contributed to the moderate correlation, in addition to differences in model structure and spatial scale.

The 2001 MAPLE predictions did not differ much from the 2010 predictions. In some areas, the 2001 predictions were lower than for 2010, whereas in the Roman cohort the 2001 exposures were higher than the 2010 exposures. In ELAPSE, we applied downward trends in PM<sub>2.5</sub> from the 1990s to 2010 and further based on the DEHM model over Europe. Routine monitoring data in a small number of countries documented a downward trend in PM<sub>2.5</sub> and, for a longer period, PM<sub>10</sub>. The difference is partly explained by the year 2001 being a year with relatively low pollution levels and by the very sparse historic PM<sub>2.5</sub> monitoring data across Europe.

Despite the only moderate agreement within cohorts, we observed nearly identical associations with mortality in the pooled cohort for the two methods. In the mortality analysis of the administrative cohorts and the morbidity analysis of the pooled cohort there were larger discrepancies, but without a consistent pattern of larger effect estimates for the ELAPSE or MAPLE exposure model.

### **ASSOCIATIONS WITH OZONE**

For O<sub>3</sub>, we did not find the previously reported positive associations with mortality and morbidity in studies conducted in North America (Di et al. 2017; Jerrett et al. 2009). In the MAPLE project, consistent associations with O<sub>3</sub> were found. Associations in single-pollutant models in our study were negative, most of them statistically significant. This may reflect the high negative correlation between O<sub>3</sub> and especially NO<sub>2</sub> and BC. In two-pollutant models, the negative associations were no longer statistically significant in the administrative cohort mortality analyses and all incidence analyses. However, in the mortality analysis in the pooled cohort, the negative associations with O<sub>3</sub> remained statistically significant in two-pollutant models. These negative associations were also found when we represented O<sub>3</sub> exposure at a larger spatial scale, more comparable to previous North American studies. We note that the associations with O<sub>3</sub> were observed over a narrow range of

exposure (98% of the pooled cohort between 60 and 100  $\mu\text{g}/\text{m}^3$ ; 68% between 80 and 100  $\mu\text{g}/\text{m}^3$ ). In comparison, the standard deviation of the  $\text{O}_3$  distribution in the Canadian MAPLE cohort was about 15  $\mu\text{g}/\text{m}^3$ , with 68% of observations between 57 and 87  $\mu\text{g}/\text{m}^3$ . The mean  $\text{O}_3$  concentration in the U.S. Medicare study was 92  $\mu\text{g}/\text{m}^3$  with a range of about 60–120  $\mu\text{g}/\text{m}^3$ . Our study was therefore less well suited to assess health effects of ambient  $\text{O}_3$ . We did not interpret the remaining negative association in two-pollutant models in one of the assessed outcomes (mortality) in one group of cohorts (pooled only), as indicating a protective effect of  $\text{O}_3$ . Rather, it may reflect the difficulty of assessing independent effects in two-pollutant models in a setting of moderately high correlations between pollutants. Another difference with the U.S. Medicare study and the Canadian MAPLE study is the correlation between  $\text{PM}_{2.5}$  and  $\text{O}_3$ . In both North American studies, a positive correlation was found (+0.24 in the Medicare cohort) while within our European cohorts the correlation was highly negative. This difference is likely explained by the study area of the respective studies: very large areas in the North American studies versus smaller individual countries or parts of countries in ELAPSE. An analysis over entire Europe would also likely show a more positive correlation between warm season  $\text{O}_3$  and  $\text{PM}_{2.5}$ , as for both pollutants concentrations are high in Southern Europe and low in Northern Europe. We did not exploit these large-scale exposure patterns in our analysis, because we could not pool the large administrative cohorts for privacy reasons and we were not comfortable with comparing health across countries that differ in many covariates that are difficult to control.

## TWO-POLLUTANT MODELS

In our main analysis we evaluated four key pollutants. We assessed exposure to  $\text{PM}_{2.5}$ ,  $\text{NO}_2$ , BC, and  $\text{O}_3$  with the same method and at the same fine spatial scale. We applied single- and two-pollutant models to assess potential independent effects of the four pollutants. Two-pollutant models can be difficult to interpret, particularly if mutual correlations are high (Coull et al. 2015; Davalos et al. 2017; Stafoggia et al. 2017). Differential measurement error may also complicate the interpretation of two-pollutant models (Butland et al. 2019; Evangelopoulos et al. 2020). The pollutant with the lowest measurement error may show the most consistent association in two-pollutant models. We have not been able to fully analyze the complex measurement error structure (see the section on the impact of measurement error below).

In some analyses, we found that after adjustment for  $\text{NO}_2$ , a significant single-pollutant HR for  $\text{PM}_{2.5}$  was

reduced to unity, whereas the association with  $\text{NO}_2$  remained robust. Because the correlation between  $\text{PM}_{2.5}$  and  $\text{NO}_2$  was moderate and the width of the CI was only modestly increased in two-pollutant models, the reduction of the  $\text{PM}_{2.5}$  HR cannot be interpreted as an artefact related to multicollinearity. The  $\text{NO}_2$  association may reflect direct effects of  $\text{NO}_2$  or of correlated combustion-related particles, such as ultrafine particles. We did not interpret the reduction of the  $\text{PM}_{2.5}$  HR as implying that in our setting, particles had no effect, as adjustment for  $\text{NO}_2$  also adjusted for particles from the sources shared with  $\text{NO}_2$ , including motorized traffic and other sources of fossil fuel combustion. The  $\text{PM}_{2.5}$  HR was reduced in some analyses (stroke, asthma, and COPD in the pooled cohort and mortality in the administrative cohorts) but remained significant in other analyses (lung cancer incidence and mortality pooled cohort).

$\text{NO}_2$  and BC were highly correlated within cohorts, resulting in substantially wider CIs in two-pollutant models. It is remarkable that in all two-pollutant models of  $\text{NO}_2$  and BC, the association with  $\text{NO}_2$  remained (pooled and administrative cohorts as well as mortality and various morbidity outcomes). This provides some support that the associations we found for  $\text{NO}_2$  at least partly reflects a direct  $\text{NO}_2$  effect. Analyses with the lasso method, which is less affected by high correlations than the standard linear regression model, supported this. The high correlation between the two-pollutants clearly limits this interpretation. We furthermore cannot fully exclude that differential measurement error explains the more robust association for  $\text{NO}_2$ .  $\text{NO}_2$  models were based upon more sites, but the performance of the  $\text{NO}_2$  model was only mildly better than the BC model (hold-out validation  $R^2$  58% vs. 51%). In our measurement error analysis, we found little impact on HRs, but we evaluated only the ambient part of all possible sources of measurement error.

## FURTHER METHODOLOGICAL ISSUES

### Exposure Assessment

We estimated exposure to all air pollutants at a fine spatial scale (100 × 100 m) using Western Europe-wide hybrid LUR models. We used satellite observations, chemical transport models, land use data, and traffic data as predictors in models developed based on AirBase routine monitoring ( $\text{PM}_{2.5}$ ,  $\text{NO}_2$ , and  $\text{O}_3$ ) or ESCAPE monitoring data (BC and PM composition). Our main exposure was the modeled exposure for the year 2010. We assessed annual average concentrations for every year until 1990 using back-extrapolation based upon the Danish DEHM model.

The models had moderate-to-good performance in hold-out validation and validation with external data (de Hoogh et al. 2018). Although  $100 \times 100$  m is a detailed spatial scale — especially for NO<sub>2</sub> and BC, which are components with strong local sources (HEI Panel on the Health Effects of Traffic-Related Air Pollution 2010) — very local traffic impacts (including street configuration) are not sufficiently captured in the model. We furthermore did not have traffic intensity data available across Europe to include in the models. Model performance measured by  $R^2$  but not RMSE was lower for the lowest levels, especially for O<sub>3</sub>.

Correlations with local dispersion models in the administrative cohorts were moderate for PM<sub>2.5</sub> (and high for NO<sub>2</sub>). This pattern agrees well with a comparison made between study-area specific ESCAPE LUR models with local dispersion models (de Hoogh et al. 2014). The difference was attributed to better developed dispersion models and more robust NO<sub>2</sub> LUR models as they were based upon twice as many sites as the PM<sub>2.5</sub> sites (40 vs. 20 per area in ESCAPE). The latter explanation is less likely in ELAPSE as the robustness does not increase much beyond 80 monitoring locations (Wang et al. 2012). The correlations between ELAPSE and ESCAPE model predictions were low to moderate for PM<sub>2.5</sub> and moderate to high for NO<sub>2</sub> in the individual cohorts of the pooled cohort, with the exception of Stockholm, where the correlation was near zero. Remarkably, the correlation of the ELAPSE model predictions with the local dispersion model in Stockholm was much higher, especially for BC. The lower correlation for the SNACK subcohort may be related to the much smaller study area for this specific subcohort (an area in central Stockholm versus Stockholm County, including surround smaller towns). In such a small area, the European model may have more limited validity. The lower within-area correlations within the pooled cohort compared with the administrative cohorts may be due to the lower robustness of the ESCAPE models, which were based upon a small number of sites especially for PM<sub>2.5</sub>. Alternatively, the ELAPSE model may miss specific local variables. Finally, the study areas of the cohorts making up the pooled cohort were generally smaller than the national administrative cohorts, reducing the within-area variability. This will likely reduce the correlation between two exposure assessment methods.

Back-extrapolation has resulted in additional uncertainty about the concentrations in years before 2010. Back-extrapolation assumes that the spatial contrast is stable in previous years (constant with the difference method and larger with the ratio back-extrapolation method). In our primary exposure paper (De Hoogh et al. 2016), we documented the stability of measured and modeled spatial

surfaces to at least 2000 for NO<sub>2</sub> and O<sub>3</sub>. For PM<sub>2.5</sub>, measurements have not been performed long enough to test the stability assumption. The stability of spatial contrasts is in agreement with previous studies in the Netherlands, Rome, the United Kingdom, and Canada (Cesaroni et al. 2012; Eeftens et al. 2011; Gulliver et al. 2011; Wang et al. 2013). The DEHM used for back-extrapolation has been extensively validated (Brandt et al. 2012). Comparisons of trends of DEHM predictions with measured trends in countries with a long time series of monitoring data generally showed good agreement. The DEHM is based on some of the best available gridded emission databases back in time to 1990 and country-based totals back to 1980. Naturally there are some uncertainties in the emissions that also might increase as one moves further into the past. Especially in 1980s, we can expect greater uncertainties, because both emission data and measurement data were sparse. From 1990 onward, the emission data are of high quality. The meteorological data are not expected to be less accurate back in time, because there have been good measurements and data for the whole period. Global meteorological models have been run again in re-analysis mode to provide high-quality analysis data. Our regional weather model (WRF) (Grell et al. 1995, Janjic 1994) has also been rerun whenever global data were updated. There might be differences in the uncertainties between pollutants. For O<sub>3</sub>, NO<sub>x</sub>, NO<sub>2</sub>, and SO<sub>2</sub> we have more good quality and longer time series for validation of the modeling system both in terms of trends in concentrations and underlying emissions. For these species and other secondary species, that is, the secondary inorganic aerosols, we don't see a significant decrease in performance over time — at least from 1990 and onward. For the 1980s, the measurement data are so sparse that it is hard to draw conclusions on the development in the performance of this decade. For total PM<sub>2.5</sub> and even more so for BC, high-quality and long-term measurements are quite rare, giving us only little chance to evaluate the system. However, we have carried out evaluation of the model from the 1990s and onward for the secondary inorganic components (sulfate, nitrate, and ammonium) showing very good performance — and also back in time. For secondary organic aerosols there are only few measurements. For sea salt, we expect that the performance is equally good back in time, since this component only depends on meteorology. To our best knowledge and based on all available long-term evaluations with measurement data in Europe, we believe that the long-term trends predicted by DEHM for the pollutants O<sub>3</sub>, NO<sub>2</sub>, NO<sub>x</sub>, SO<sub>2</sub>, SO<sub>4</sub>, NO<sub>3</sub>, NH<sub>4</sub>, total PM<sub>2.5</sub>, and BC perform well from the 1990s and that there is no large difference in performance between these pollutants.

We had consistent residential history data available for all administrative cohorts and a large fraction of the participants in the pooled cohort. Residential history data were used in our time-varying analysis.

### Impact of Measurement Error in Exposure

We documented moderate-to-good performance of our exposure models for assessing ambient concentrations using cross-validation and external validation using the ESCAPE data (PM<sub>2.5</sub> and NO<sub>2</sub>). Although several studies have documented associations between long-term average ambient concentrations and long-term average personal exposure (Hoek 2017), these studies are not useful for formal measurement error correction. The main reason is the difficulty to characterize long-term average personal exposure to outdoor-generated pollution. This difficulty not only applies to ELAPSE, but to other epidemiological studies on long-term exposure as well. We did not have the data to formally evaluate the complex measurement error structure as has been laid out in previous work, which includes type of error (Berkson or classical) and correlations of errors (Samoli and Butland 2017). We therefore relied on two simpler methods focusing on how robustly the model explained outdoor concentrations. Both the regression calibration and the application of multiple models derived from the five-fold cross-validation suggested robust associations. We note, however, that the measurement error impact may vary temporally (likely higher in earlier years) and spatially (higher in some countries). Furthermore, as pollutant concentrations are correlated and the same predictor input data are used to develop models, the measurement error is likely correlated between pollutants. The measurement error impact differs by modeling approach and exposure used (time-varying vs. fixed year). In a multipollutant model, the direction of bias due to measurement error is unpredictable in terms of direction and magnitude, especially when the pollutants are highly correlated. In principle the pollutant measured with less measurement error will provide a less biased estimate (Samoli and Butland 2017). Estimates from both pollutants are affected in two-pollutant models, but the amount of impact due to confounding or measurement error is not straightforward to distinguish and also may be interlinked. We could not address these complexities in ELAPSE.

### Health Data

Health data were obtained by linking to national mortality and morbidity registries in the administrative cohorts and most of the pooled cohort. Loss to follow-up was minimal, related to obtaining our health outcome data

by linking to routine mortality and hospital admission registries. There may be differences in coding of causes of mortality or morbidity between countries. Potential differences in coding would unlikely have biased our associations, as we analyzed all administrative cohorts separately and used strata for subcohort in the analysis of the pooled cohort.

In the mortality analyses, we only had access to the underlying cause of death. In Europe, contributing cause of death is not consistently recorded on death certificates. The lack of contributing cause of death data has especially affected our diabetes mortality analysis, since most people with diabetes predominantly die from cardiovascular disease. We were therefore unable to conduct the analysis in the Canadian cohort showing increased HRs for PM<sub>2.5</sub> and cardiovascular mortality if diabetes is listed as a contributing cause (Pinault et al. 2018).

Because of the uncertainty of the date of onset of asthma and COPD, the Cox proportional hazard model may not have been the optimal model for analysis. Because of coherence with other endpoints and the loss of any information on timing of disease onset, we did not conduct the alternative logistic regression analyses.

Adult-onset asthma is a chronic disease with complex phenotype and recurring symptoms that makes it difficult to diagnose and identify a precise time of onset. Asthma definitions based on self-reports from respiratory disease surveys are subject to recall bias, resulting in less strict definitions and likely an overestimation of the true burden. In this study we benefited from objective definitions based on hospital-discharge diagnoses from nationwide hospital registers in Denmark and Sweden. Asthma incidence rates defined by hospital discharge diagnoses may underestimate true asthma burden, as not all asthma patients require hospital contact, and thus, an asthma hospital-discharge diagnosis typically represents a point of disease progression to a more severe stage or exacerbation. It is appealing as it presents a well characterized asthma definition, typically confirmed by objective measurements of lung function and reversible airflow obstruction as standard procedures in Danish and Swedish hospitals. The specificity of asthma diagnoses in the Danish Hospital Discharge Register was found to be as high as 0.98, validating their use in epidemiological studies.

The use of the first-ever hospital discharge diagnosis for COPD in hospital registers confers some degree of objective assessment, and studies have demonstrated the high specificity of COPD diagnoses using hospital registers in both Denmark (92% overall positive predictive rate) (Thomsen et al. 2011) and Sweden (<10% misclassification rate) (Inghammar et al. 2012), validating their use in epidemiology research. However, COPD is often undiagnosed for many

years and not all patients with COPD are hospitalized, therefore first-ever hospital discharge diagnoses may underrepresent the true COPD incidence. Additionally, it is difficult to separate COPD from asthma diagnosis and a part of COPD cases thus may in fact be asthma cases or mixed cases, which is another factor leading to underrepresentation of true COPD incidence. However, in our study, censoring occurred at the time of diagnosis of asthma, which could offset parts of the misclassification for first-ever COPD diagnosis. Rather, COPD hospital discharge diagnosis should not be regarded as onset of a disease but a hallmark of the progression of pre-existing COPD into a more severe stage. Thus, our study examines the onset of severe COPD, and likely represents older patients with a severe disease. One limitation of our study is the lack of lung function measurements in the cohorts for verifying COPD diagnoses.

### **Confounding and Effect Modification**

We have chosen our confounder models based on prior studies (ESCAPE and more recent cohort studies) and data availability. The philosophy was to have a comprehensive adjustment for known risk factors on an individual- and area-level scale. Therefore, we preferred to adjust for a comprehensive set of area-level SES variables rather than simple indicators of region of the country in the main model. We did not use directed acyclic graphs to identify the confounder models, as was conducted in the MAPLE study (Brauer et al. 2019). Although transparent, DAG models require strong assumptions and often result in additional debate (Brauer et al. 2019). In MAPLE, HRs did not differ substantially between DAG models and fully adjusted models. We performed systematic analyses of the robustness of our findings to differences in confounder adjustment. HRs were generally robust in a series of sensitivity analyses, including additional confounder variables in the pooled cohort, indirect adjustment in the administrative cohorts, adjustment for traffic noise for the cardiovascular and diabetes outcomes, and imputation of missing values in confounders in the pooled cohort.

In a review study, it was found that effect estimates for PM<sub>2.5</sub> and natural-cause mortality were higher at low concentrations than at high concentrations; and higher with better adjustment for area-level SES (Vodonos et al. 2018).

We acknowledge that some over-adjustment may be possible. Especially in the pooled cohort, effect estimates were not affected much by additional confounder control, however. In our main confounder model of the pooled cohort, we adjusted for BMI. There is an emerging literature (de Bont et al. 2019; Kim et al. 2018) suggesting that air pollution exposure may affect BMI, implying a risk of over-

adjustment, particularly in analyses involving cardio-metabolic outcomes. As BMI is affected probably substantially more by lifestyle factors, such as physical activity and dietary habits, than by air pollution, inclusion in the main model was considered justified. Analyses with and without BMI in the confounder model, showed only marginal differences in air pollution effect estimates. In the pooled cohort, we adjusted for a large number of individual-level confounders, but some variables were lacking, including physical activity, occupational risk factors, and passive smoking. It is unlikely that physical activity has biased effect estimates, as adjustment for BMI likely has accounted for most of the potential confounding by physical activity. Residual confounding is possible if physical activity affects health by mechanisms other than through BMI. In general population studies, exposure to occupational risk factors is also unlikely to be an important confounder. A sensitivity analysis further adjusting for white versus blue collar workers did not affect our effect estimates.

In the administrative cohorts, we did not have information on individual lifestyle factors, except for the English cohort, in which smoking status and BMI were available. We addressed this limitation by indirect adjustment and additional adjustment for area-level proxies for smoking and BMI. HRs remained elevated after these additional adjustments. Importantly, in the administrative cohorts we were able to adjust for multiple individual- and area-level SES variables. We made a large effort to include multiple SES variables in the analysis, at both the neighborhood and regional spatial scale. Since associations between air pollution and lifestyle are affected in a major way by SES, the inclusion of detailed SES information likely has reduced confounding significantly. We were able to include more detailed area-level SES information in the administrative cohorts than in the pooled cohort, as availability of data differed across countries.

We additionally applied the indirect adjustment approaches for missing lifestyle factors (Shin 2014) and adjustment for area-level lung cancer, COPD, or diabetes as proxies of missing lifestyle. These further adjustments did not explain the associations we found between air pollution and mortality. Before applying the indirect adjustment, we compared characteristics of the survey and cohort population, as the relationship between air pollution and lifestyle may differ by population group (Brauer et al. 2019; Shin et al. 2014; Strak et al. 2017). In particular, the Dutch and Norwegian survey was mostly comparable to the full cohort, for the other cohorts differences were somewhat larger. For all cohorts, the quantitative results should be interpreted with caution. In the MAPLE study, different weighting schemes were applied to increase the

comparability between the survey and cohort. The indirect adjustment impact on the HRs differed only mildly between the different weighting schemes (Brauer et al. 2019).

We used the same covariate model for the different outcomes. In the administrative cohorts, models were identical for all causes of death. The confounder models in the pooled cohort differed between mortality and lung cancer incidence on the one hand and stroke, coronary heart disease, asthma, and COPD on the other hand. For the latter outcomes, we were able to include education and smoking intensity and duration for former smokers, as the Vorarlberg cohort did not contribute to these endpoints. The rationale for the same covariate model is that all (administrative cohorts) or most (pooled cohort) covariates are general variables that affect both respiratory and cardio-metabolic morbidity and mortality. BMI might not have been necessary to adjust for in respiratory mortality analyses, so some over-adjustment may have occurred.

We had a small fraction of subjects with missing data on exposure and confounders. Selection bias related to missing values is unlikely because of the small fraction of missing data, the similar effect estimates when applying multiple imputation for covariates, and the very small difference in HRs from model 1 between the full population and the full case study population, which was used for our main analyses.

### Noise Adjustment

A strength of ELAPSE is that we were able to adjust for road-traffic noise using state of the art noise models in the Swedish and Danish cohorts in the pooled cohort and the Belgian, Dutch, Norwegian, Roman, and Swiss administrative cohorts. Noise has been associated with cardiovascular morbidity and diabetes (incidence) in previous studies, resulting in updated WHO guidelines for noise (Eze et al. 2017; Sorensen et al. 2013; WHO 2018). Road-traffic noise and air pollution are typically correlated spatially, because motorized road traffic is an important source of both noise and air pollution. The correlation between noise and air pollution varies by study area, but is generally moderate (for example, see Fecht et al. 2016). Mutual confounding is therefore possible, but the moderate correlation suggests two-exposure models will be effective to address potential confounding (Klompmaaker et al. 2019). In our study, we also found moderate correlations between noise and air pollution. The air pollution associations were robust to noise adjustment in the pooled cohort and the Belgian, Dutch, Norwegian, and Roman administrative cohorts. In the Swiss cohort, air pollution effect estimates were attenuated substantially by noise adjustment. The different impact of adjustment for noise

in different cohorts agrees with previous observations. Differential measurement error, populations, correlations, and the magnitude of noise effects may contribute to these differences. Reporting noise effect estimates is beyond the scope of ELAPSE.

### Multiple Testing

We performed a large number of epidemiological analyses. We do not think a formal adjustment for multiple testing is needed, as the large number of analyses primarily results from sensitivity analyses of the same exposure–response relationship. Furthermore, we prefer to assess the consistency of patterns rather than rely primarily on statistical significance. We did assess a large number of outcomes (especially cause-specific mortality) in relation to the four main pollutants.

### COMPARISON WITH THE CANADIAN AND U.S. HEI-FUNDED STUDIES, ESCAPE, AND RECENT SYSTEMATIC REVIEWS

The ELAPSE study is one of three studies funded by HEI to investigate effects of low levels of air pollution on mortality and morbidity. Table 30 provides a comparison of linear effect estimates for natural-cause mortality.

Table 30 shows that for PM<sub>2.5</sub>, the ELAPSE pooled cohort estimate was larger than the estimates in the other two HEI-funded low-level studies, the ESCAPE study, and recent reviews. The ELAPSE effect estimate was much more precise than the effect estimates in the ESCAPE cohorts. ESCAPE cohorts contributed the large majority of subjects to the ELAPSE pooled cohort. The effect estimate in the ELAPSE pooled cohort was much smaller when expressed over the higher baseline exposure contrast: 1.096 to 1.116, depending on back-extrapolation method (data from Table 9, re-expressed per 10 µg/m<sup>3</sup>). The baseline exposure estimates are more in line with the original ESCAPE estimates, and with the North American HEI-funded studies and several recent meta-analyses (Chen and Hoek 2020; Pope et al. 2020; Vodonos et al. 2018). The time-varying exposure analyses showed effect estimates that were similar to the main ELAPSE findings.

All systematic reviews documented substantial heterogeneity in effect estimates of PM<sub>2.5</sub> across studies. The effect estimate in the pooled cohort was identical to the originally published HR for the Canadian Community Health Survey of 1.26 (1.19–1.34) per 10 µg/m<sup>3</sup> (Pinault et al., 2016). This was the highest HR in the systematic review by WHO (Chen and Hoek 2020). Multiple studies showed HRs of between 1.14 and 1.18 in the Chen and Hoek review. The reasons for this heterogeneity have not been fully elucidated but include degree of confounder

**Table 30.** Comparison of Findings for Natural-Cause Mortality with the Canadian and U.S. HEI-Funded Studies, ESCAPE, and Recent Systematic Reviews

Study	HR (95% CI) PM <sub>2.5</sub> per 10 µg/m <sup>3</sup>	HR (95% CI) NO <sub>2</sub> per 10 µg/m <sup>3</sup>	HR (95% CI) O <sub>3</sub> per 10 µg/m <sup>3</sup>
ELAPSE pooled cohort	1.28 (1.22, 1.33)	1.09 (1.07, 1.10)	0.90 (0.88, 0.91)
ELAPSE administrative cohorts	1.11 (1.04, 1.17)	1.04 (1.02, 1.07)	0.96 (0.93, 0.98)
MAPLE CanCHEC	1.053 (1.041, 1.065)	1.004 (1.002, 1.006)	1.036 (1.033, 1.038)
MAPLE CCHS	1.11 (1.04, 1.18)	1.024 (1.016, 1.040)	1.025 (1.015, 1.034)
MEDICARE	1.084 (1.081, 1.086)	NA	1.012 (1.011, 1.012)
ESCAPE	1.14 (1.04, 1.26)	1.01 (0.99, 1.03)	
Pope et al. 2020	1.08 (1.06, 1.11)		
Chen and Hoek 2020	1.08 (1.06, 1.09)		
Huangfu and Atkinson 2020		1.02 (1.01, 1.04)	1.01 (1.00, 1.02)

NA = not applicable.

adjustment, the average pollution level, and the age of the population (Vodonos et al. 2018).

For three administrative cohort studies, associations with mortality have been reported previously. Consistent with ELAPSE, all three studies found statistically significant associations between PM<sub>2.5</sub> and natural-cause mortality, however the magnitude of the HRs differed substantially with the previously published results. The Roman cohort study (Cesaroni et al. 2013), involving more than 1 million subjects, estimated PM<sub>2.5</sub> for 1-km<sup>2</sup> grids using a dispersion model and found an HR for all-cause mortality of 1.04 (CI = 1.03, 1.05) per 10 µg/m<sup>3</sup> PM<sub>2.5</sub>. NO<sub>2</sub> estimated with local LUR models was also related to mortality. The English cohort study, including more than 800,000 subjects, estimated PM<sub>2.5</sub> for 1-km<sup>2</sup> grids using a national dispersion model and found a HR for all-cause mortality of 1.13 (CI = 1.00, 1.27) per 10 µg/m<sup>3</sup> PM<sub>2.5</sub> (Carey et al. 2013). The Dutch cohort study estimated PM<sub>10</sub> from a LUR model at a 100 × 100 m grid. An HR for all-cause mortality of 1.08 (CI = 1.07, 1.09) per 10 µg/m<sup>3</sup> PM<sub>10</sub> was found, which, assuming a ratio of PM<sub>2.5</sub>:PM<sub>10</sub> of 0.67 translates into a HR of 1.12 per 10 µg/m<sup>3</sup> PM<sub>2.5</sub> (Fischer et al. 2015). NO<sub>2</sub> was associated independently with mortality. The Roman cohort study was conducted in the same population as ELAPSE. The Dutch and English cohorts

were newly constructed for ELAPSE with a more recent start and follow-up period.

Table 31 compares the ELAPSE findings for incidence of acute coronary heart disease, stroke, and lung cancer with the findings from ESCAPE in strongly overlapping cohorts. The findings in the ELAPSE pooled cohort were much more precise (narrower CIs) compared with the ESCAPE cohorts. HRs for ELAPSE were lower for PM<sub>2.5</sub> and BC and higher for NO<sub>2</sub>. The HR for lung cancer was larger for PM<sub>2.5</sub> than in recent meta-analyses, whereas the HR for NO<sub>2</sub> was lower than in recent meta-analyses.

### STRENGTHS AND LIMITATIONS

Here we summarize the strengths and limitations that have been discussed in more detail above.

The strengths of the ELAPSE study include the following:

- Pooled analysis of a number of well-characterized cohorts with individual-level data on important potential confounders and effect modifiers;
- Highly standardized analysis of seven large administrative cohorts from different European regions and countries;

**Table 31.** Comparison of Findings for Incidence of Acute Coronary Heart Disease, Stroke, and Lung Cancer with ESCAPE and with Recent Systematic Reviews

Study	HR (95% CI) PM <sub>2.5</sub> per 5 µg/m <sup>3</sup>	HR (95% CI) NO <sub>2</sub> per 10 µg/m <sup>3</sup>	HR (95% CI) BC per 10 <sup>-5</sup> /m
ELAPSE pooled cohort acute coronary heart disease	1.02 (0.95, 1.10)	1.04 (1.01, 1.07)	1.02 (0.99, 1.06)
ESCAPE acute coronary heart disease	1.13 (0.98, 1.30)	1.03 (0.97, 1.08)	1.10 (0.98, 1.24)
ELAPSE pooled cohort stroke	1.10 (1.01, 1.21)	1.08 (1.04, 1.12)	1.06 (1.02, 1.10)
ESCAPE stroke	1.19 (0.88, 1.62)	0.99 (0.89, 1.11)	1.08 (0.83, 1.41)
ELAPSE pooled cohort lung cancer	1.13 (1.05, 1.21)	1.02 (0.97, 1.07)	1.02 (0.97, 1.07)
ESCAPE lung cancer	1.18 (0.96, 1.46)	0.99 (0.93, 1.06)	1.12 (0.88, 1.42)
Hamra et al. 2014 lung cancer mortality/incidence	1.04 (1.02, 1.07)	1.04 (1.01, 1.08)	
Cui et al. 2015 lung cancer mortality	1.04 (1.03, 1.06)		
Yang et al. 2016 lung cancer mortality/incidence	1.04 (1.01, 1.06)	1.07 (1.03, 1.10)	

- Comparison of results of pooled cohort and administrative cohorts' analyses;
- Analyses of natural-cause as well as cause-specific mortality in the pooled cohort and in the administrative cohorts, using the exact same ICD outcome definition codes;
- Analyses of incidence of major morbidity outcomes: acute coronary heart disease, cerebrovascular events, lung cancer, asthma, and COPD;
- Uniform exposure assessment using state-of-the-art hybrid exposure assessment models at a fine, 100 × 100 m, scale across Europe;
- Identical assessment of four key pollutants, by adding fine-scale exposure to NO<sub>2</sub>, BC, and O<sub>3</sub> to the more often assessed PM<sub>2.5</sub>;
- Back-extrapolation to year of recruitment using a state of art chemical transport model;
- Adjustment for road-traffic noise for natural-cause mortality, cardiovascular mortality, and morbidity and diabetes mortality; and
- Different approaches to assess the shape of the concentration–response functions; a multitude of sensitivity analyses.

The limitations of the ELAPSE study include the following:

- Relatively few subjects residing in the areas with the very lowest PM<sub>2.5</sub> concentrations (in comparison with the North American studies) of <10 µg/m<sup>3</sup>, the current WHO Air Quality Guideline;
- High correlation between NO<sub>2</sub> and BC, limiting the possibilities to assess the potential direct effects of NO<sub>2</sub> and primary combustion particles;
- Very small contrast in O<sub>3</sub> exposure, especially in the pooled cohort;
- Moderate performance of the exposure models for NO<sub>2</sub> and BC especially at a scale that still misses very local traffic impacts. We furthermore did not have traffic intensity data available across Europe to include in the models;
- Some missing confounder data in the pooled cohort and especially administrative cohorts;
- Limited availability of area-level confounders harmonized across countries in Europe, limiting the pooled cohort analysis; and
- Because of the uncertainty of the date of onset of asthma and COPD, the Cox proportional hazard model may not have been the optimal model for analysis.

---

## IMPLICATIONS OF FINDINGS

---

The ELAPSE study was designed to investigate health effects of air pollution occurring at levels below current annual EU limit values, U.S. EPA National Ambient Air Quality standards, and WHO Air Quality Guidelines for annual average concentrations. For PM<sub>2.5</sub>, the annual average values are 25 µg/m<sup>3</sup>, 12 µg/m<sup>3</sup>, and 10 µg/m<sup>3</sup>, respectively. In ELAPSE significant associations with mortality and morbidity were found with exposures even below the WHO Air Quality Guidelines. We note that in ELAPSE, a small fraction of the population resided in areas (mostly in Scandinavia) with PM<sub>2.5</sub> concentrations below 10 µg/m<sup>3</sup>. Because of the large study population, ELAPSE still included almost 2 million subjects living in areas with PM<sub>2.5</sub> levels below 10 µg/m<sup>3</sup>, allowing for robust assessments.

For NO<sub>2</sub>, the EU limit, EPA standard, and WHO guideline annual average values are 40 µg/m<sup>3</sup>, 100 µg/m<sup>3</sup>, and 40 µg/m<sup>3</sup>, respectively. In ELAPSE, significant associations with mortality and morbidity were found with exposures even below the WHO Air Quality Guidelines. We note that in ELAPSE, almost all subjects resided in areas with NO<sub>2</sub> concentrations below 40 µg/m<sup>3</sup>, so that the generalizability of the findings below the current EU Limit Value and WHO Air Quality Guidelines is very high. This is true even for concentrations below 20 µg/m<sup>3</sup>. Almost 6 million subjects included in ELAPSE had annual average NO<sub>2</sub> exposures below 20 µg/m<sup>3</sup>.

For O<sub>3</sub>, there are no long-term annual standards or guidelines. The ELAPSE results do not support effects of long-term exposure to O<sub>3</sub> on mortality or morbidity, but the contrast in exposure for O<sub>3</sub> was small.

No long-term standards or guidelines exist for BC or other PM components. The positive ELAPSE findings for BC support our hypothesis that primary combustion particles play a role in adverse health effects of air pollution.

Overall, robust evidence was found that major air pollutants were associated with mortality and morbidity well below current air quality standards and guideline values.

---

## ACKNOWLEDGMENTS

---

This report describes the results of a large European study on health effects of low-level air pollution, conducted as per HEI Research Agreement #4954-RFA14-3/16-5-3. The project was conducted in the framework of RFA14-3 Assessing Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution. We thank Marjan Tewis for her construction of the pooled cohort. We thank

Dr. Rick Burnett for supplying the code for the SCHIF functions and commenting on the first application. We thank Dr. Randall Martin and Dr. Aaron van Donkelaar for providing the MAPLE PM<sub>2.5</sub> exposure estimates and the satellite PM<sub>2.5</sub> estimates used in developing the main exposure models.

---

## REFERENCES

---

- Andersen ZJ, Hvidberg M, Jensen SS, Ketzel M, Loft S, Sørensen M, et al. 2011. Chronic obstructive pulmonary disease and long-term exposure to traffic-related air pollution: A cohort study. *Am J Respir Crit Care Med* 183:455–461.
- Atkinson RW, Butland BK, Dimitroulopoulou C, Heal MR, Stedman JR, Carslaw N, et al. 2016. Long-term exposure to ambient ozone and mortality: A quantitative systematic review and meta-analysis of evidence from cohort studies. *BMJ Open* 6:e009493.
- Atkinson RW, Carey IM, Kent AJ, van Staa TP, Anderson HR, Cook DG. 2013. Long-term exposure to outdoor air pollution and incidence of cardiovascular diseases. *Epidemiology* 24:44–53.
- Austin PC. 2017. A tutorial on multilevel survival analysis: Methods, models and applications. *Int Stat Rev* 85:185–203.
- Bai L, Shin S, Burnett RT, Kwong JC, Hystad P, van Donkelaar A, et al. 2020. Exposure to ambient air pollution and the incidence of lung cancer and breast cancer in the Ontario population health and environment cohort. *Int J Cancer* 146:2450–2459.
- Bai L, Weichenthal S, Kwong JC, Burnett RT, Hatzopoulou M, Jerrett M, et al. 2019. Associations of long-term exposure to ultrafine particles and nitrogen dioxide with increased incidence of congestive heart failure and acute myocardial infarction. *Am J Epidemiol* 188:151–159.
- Basagaña X, Rivera M, Aguilera I, Agis D, Bouso L, Elosua R, Foraster M, de Nazelle A, Nieuwenhuijsen M, Vila J, et al. 2012. Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmos Environ*. 54:634–642. doi: 10.1016/j.atmosenv.2012.01.064.
- Bechle MJ, Millet DB, Marshall JD. 2015. National spatio-temporal exposure surface for NO<sub>2</sub>: Monthly scaling of a satellite-derived land-use regression, 2000–2010. *Environ Sci Technol* 49:12297–12305.

- Beelen R, Hoek G, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, et al. 2015. 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* 123:525–533.
- Beelen R, Hoek G, van den Brandt PA, Goldbohm RA, Fischer P, Schouten LJ, et al. 2008. Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-air study). *Environ Health Perspect* 116:196–202.
- Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedeli X, et al. 2013. Development of NO<sub>2</sub> and NO<sub>x</sub> land use regression models for estimating air pollution exposure in 36 study areas in Europe — the ESCAPE project. *Atmos Environ* 72:10–23.
- Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B, et al. 2014a. Effects of long-term exposure to air pollution on natural-cause mortality: An analysis of 22 European cohorts within the multicentre ESCAPE project. *Lancet* 383:785–795.
- Beelen R, Stafoggia M, Raaschou-Nielsen O, Andersen ZJ, Xun WW, Katsouyanni K, et al. 2014b. Long-term exposure to air pollution and cardiovascular mortality: An analysis of 22 European cohorts. *Epidemiology* 25:368–378.
- Bowe B, Xie Y, Yan Y, An-Aly Z. 2019. Burden of cause-specific mortality associated with PM<sub>2.5</sub> air pollution in the United States. *JAMA Network Open* 2:e1915834.
- Brandt J, Silver JD, Frohn LM, Geels C, Gross A, Hansen AB, et al. 2012. An integrated model study for Europe and North America using the Danish Eulerian hemispheric model with focus on intercontinental transport of air pollution. *Atmos Environ* 53:156–176.
- Brauer M, Brook JR, Christidis T, Chu Y, Crouse DL, Erickson A, et al. 2019. Mortality–Air Pollution Associations in Low-Exposure Environments (MAPLE): Phase 1. Research Report 203. Boston, MA:Health Effects Institute.
- Brook RD, Rajagopalan S, Pope CA 3rd, Brook JR, Bhatnagar A, Diez-Roux AV, et al. 2010. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* 121:2331–2378.
- Brunekreef B, Holgate ST. 2002. Air pollution and health. *Lancet* 360:1233–1242.
- Burnett R, Chen H, Szyszkowicz M, Fann N, Hubbell B, Pope CA 3rd, et al. 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc Natl Acad Sci U S A* 115:9592–9597.
- Burnett RT, Pope CA 3rd, Ezzati M, Olives C, Lim SS, Mehta S, et al. 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environ Health Perspect* 122:397–403.
- Butland BK, Samoli E, Atkinson RW, Barratt B, Katsouyanni K. 2019. Measurement error in a multi-level analysis of air pollution and health: A simulation study. *Environ Health* 18:13.
- Carey IM, Atkinson RW, Kent AJ, van Staa T, Cook DG, Anderson HR. 2013. Mortality associations with long-term exposure to outdoor air pollution in a national English cohort. *Am J Respir Crit Care Med* 187:1226–1233.
- Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu LA. 1995. Measurement error in nonlinear models. London: Chapman and Hall.
- Cesaroni G, Badaloni C, Gariazzo C, Stafoggia M, Sozzi R, Davoli M, et al. 2013. Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. *Environ Health Perspect* 121:324–331.
- Cesaroni G, Forastiere F, Stafoggia M, Stafoggia M, Sozzi R, Davoli M, et al. 2014. Long term exposure to ambient air pollution and incidence of acute coronary events: Prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE project. *Br Med J* 348:f7412.
- Cesaroni G, Porta D, Badaloni C, Stafoggia M, Eeftens M, Meliefste K, et al. 2012. Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. *Environ Health* 11:48.
- Chen J, de Hoogh K, Gulliver J, Hoffmann B, Hertel O, Ketzler M, et al. 2019. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. *Environ Int* doi: 10.1016/j.envint.2019.104934.
- Chen J, Hoek G. 2020. Long-term exposure to PM and all-cause and cause-specific mortality: A systematic review and meta-analysis. *Environ Int*:105974.
- Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: An analysis of data from the Global Burden of Diseases Study 2015. *Lancet* 389:1907–1918.

- Coull BA, Bobb JF, Wellenius GA, Kioumourtzoglou M-A, Mittleman MA, Koutrakis P, et al. 2015. Part 1. Statistical Learning Methods for the Effects of Multiple Air Pollution Constituents. In: *Development of Statistical Methods for Multipollutant Research*. Research Report 183. Boston, MA:Health Effects Institute.
- Crouse DL, Peters PA, Hystad P, Brook JR, Donkelaar A, Martin RV, et al. 2015. Ambient PM<sub>2.5</sub>, O<sub>3</sub>, and NO<sub>2</sub> exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment cohort (CANCHEC). *Environ Health Perspect* 123:1180–1186.
- Crouse DL, Peters PA, van Donkelaar A, Goldberg MS, Villeneuve PJ, Brion O, et al. 2012. Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: A Canadian national-level cohort study. *Environ Health Perspect* 120:708–714.
- Cui P, Huang Y, Han J, Song F, Chen K. 2015. Ambient particulate matter and lung cancer incidence and mortality: A meta-analysis of prospective studies. *Eur J Public Health* 25:324–329.
- Cyrys J, Eeftens M, Heinrich J, Ampe C, Armengaud A, Beelem R, et al. 2012. Variation of NO<sub>2</sub> and NO<sub>x</sub> concentrations between and within 36 European study areas: Results from the ESCAPE study. *Atmos Environ* 62:374–390.
- Danesh Yazdi M, Wang Y, Di Q, Zanobetti A, Schwartz J. 2019. Long-term exposure to PM<sub>2.5</sub> and ozone and hospital admissions of Medicare participants in the Southeast USA. *Environ Int* 130:104879.
- Davalos AD, Luben TJ, Herring AH, Sachs JD. 2017. Current approaches used in epidemiologic studies to examine short-term multipollutant air pollution exposures. *Ann Epidemiol* 27:145–153 e141.
- de Bont J, Casas M, Barrera-Gomez J, Cirach M, Rivas I, Valvi D, et al. 2019. Ambient air pollution and overweight and obesity in school-aged children in Barcelona, Spain. *Environ Int* 125:58–64.
- de Hoogh K, Chen J, Gulliver J, Hoffmann B, Hertel O, Ketzler M, et al. 2018. Spatial PM<sub>2.5</sub>, NO<sub>2</sub>, O<sub>3</sub> and BC models for Western Europe — Evaluation of spatiotemporal stability. *Environ Int* 120:81–92.
- de Hoogh K, Gulliver J, Donkelaar AV, Martin RV, Marshall JD, Bechle MJ, et al. 2016. Development of West-European PM<sub>2.5</sub> and NO<sub>2</sub> land use regression models incorporating satellite-derived and chemical transport modelling data. *Environ Res* 151:1–10.
- de Hoogh K, Korek M, Vienneau D, Keuken M, Kukkonen J, Nieuwenhuijsen MJ, et al. 2014. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environ Int* 73:382–392.
- de Hoogh K, Wang M, Adam M, Badaloni C, Beelen R, Birk M, et al. 2013. Development of land use regression models for particle composition in twenty study areas in Europe. *Environ Sci Tech* 47:5778–5786.
- Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Choirat C, et al. 2017. Air pollution and mortality in the Medicare population. *N Engl J Med* 376:2513–2522.
- Dimakopoulou K, Samoli E, Beelen R, Stafoggia M, Andersen ZJ, Hoffmann B, et al. 2014. Air pollution and nonmalignant respiratory mortality in 16 cohorts within the ESCAPE project. *Am J Res Crit Care Med* 189:684–696.
- Dockery DW, Pope CA 3rd, Xu X, Spengler JD, Ware JH, Fay ME, et al. 1993. An association between air pollution and mortality in six U.S. cities. *N Engl J Med* 329 1753–1759.
- Dominici F, Schwartz J, Di Q, Braun D, Choirat C, Zanobetti A. 2019. *Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution*. Research Report 200. Boston, MA:Health Effects Institute.
- Eeftens M, Beelen R, de Hoogh K, Brunekreef B, Meliefste K, Hoek G. 2012a. Development of land use regression models for PM<sub>2.5</sub>, PM<sub>2.5</sub> absorbance, PM<sub>10</sub> and PM<sub>coarse</sub> in 20 European study areas: Results of the ESCAPE project. *Environ Sci Tech* 46:11195–11205.
- Eeftens M, Beelen R, Fischer P, Brunekreef B, Meliefste K, Hoek G. 2011. Stability of measured and modelled spatial contrasts in NO<sub>2</sub> over time. *Occup Environ Med* 68:765–770.
- Eeftens M, Tsai MY, Ampe C, Anwander B. 2012b. Spatial variation of PM<sub>2.5</sub>, PM<sub>10</sub>, PM<sub>2.5</sub> absorbance and PM<sub>coarse</sub> concentrations between and within 20 European study areas and the relationship with NO<sub>2</sub> — Results of the ESCAPE project. *Atmos Environ* 62:303–317.
- Evangelopoulos D, Katsouyanni K, Keogh RH, Samoli E, Schwartz J, Ben Barratt B, et al. 2020. PM<sub>2.5</sub> and NO<sub>2</sub> exposure errors using proxy measures, including derived personal exposure from outdoor sources: A systematic review and meta-analysis. *Environ Int* 137:105500.
- Eze IC, Foraster M, Schaffner E, Vienneau D, Harris Héritier H, Rudzik F, et al. 2017. Long-term exposure to transportation noise and air pollution in relation to incident

- diabetes in the Sapaldia Study. *Int J Epidemiol* 46:1115–1125.
- Faustini A, Rapp R, Forastiere F. 2014. Nitrogen dioxide and mortality: Review and meta-analysis of long-term studies. *Eur Respir J* 44:744–753.
- Fecht D, Hansell AL, Morley D, Dajnak D, Vienneau D, Beevers S, et al. 2016. Spatial and temporal associations of road traffic noise and air pollution in London: Implications for epidemiological studies. *Environ Int* 88:235–242.
- Fischer P, Marra M, Ameling C, et al. 2015. Air pollution and mortality — the Dutch environmental longitudinal study (DUELS). *Environ Health Perspect* 123(7):697–704.
- Fisher JA, Puett RC, Hart JE, Camargo CA Jr, Varraso R, Yanosky JD, et al. 2016. Particulate matter exposures and adult-onset asthma and COPD in the Nurses' Health Study. *Eur Respir J* 48:921–924.
- Grell GA, Dudhia J, Stauffer DR. 1995. A Description of the Fifth-generation PennState/NCAR Mesoscale Model (MM5). NCAR Technical Note NCAR/TN-398þSTR. National Center for Atmospheric Research, Boulder, Colorado, USA.
- Gulliver J, Morris C, Lee K, Vienneau D, Briggs D, Hansell A. 2011. Land use regression modeling to estimate historic (1962–1991) concentrations of black smoke and sulfur dioxide for Great Britain. *Environ Sci Tech* 45:3526–3532.
- Hamra GB, Guha N, Cohen A, Laden F, Raaschou-Nielsen O, Samet JM, et al. 2014. Outdoor particulate matter exposure and lung cancer: A systematic review and meta-analysis. *Environ Health Perspect* 122:906–911.
- Hastie TJ. 1992. Generalized additive models. In: *Statistical models in S*. Edited By Chambers JM & Hastie TJ, Routledge, Boca Raton, FL, USA.
- HEI Panel on the Health Effects of Traffic-Related Air Pollution. 2010. *Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects*. Special Report 17. Boston, MA:Health Effects Institute.
- Hoek G. 2017. Methods for assessing long-term exposures to outdoor air pollutants. *Curr Environ Health Rep* 4:450–462.
- Huangfu P, Atkinson R. 2020. Long-term exposure to NO<sub>2</sub> and O<sub>3</sub> and all-cause and respiratory mortality: A systematic review and meta-analysis. *Environ Int* 144:105998.
- Hvitfeldt UA, Severi G, Andersen ZJ, Atkinson R, Bauwelinck M, Bellander T, et al. 2021. Long-term low-level ambient air pollution exposure and risk of lung cancer: A pooled analysis of 7 European cohorts. *Env Int* 146:106249.
- Inghammar M, Engstrom G, Lofdahl CG, Eggesten A. 2012. Validation of a COPD diagnosis from the Swedish inpatient registry. *Scand J Pub Health* 40:773–776.
- Inness A, Baier F, Benedetti A, Bouarar I, Chabrilat S, Clark H, et al. 2013. The macc reanalysis: An 8 yr data set of atmospheric composition. *Atmos Chem Phys* 13:4073–4109.
- Janjic ZI. 1994. The step-mountain Eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. *Monthly Weather Rev* 122:927–945.
- Jerrett M, Burnett RT, Beckerman BS, Turner MC, Krewski D, Thurston G, et al. 2013. Spatial analysis of air pollution and mortality in California. *Am J Resp Crit Care Med* 188:593–599.
- Jerrett M, Burnett RT, Pope CA III, Ito K, Thurston G, Krewski D, et al. 2009. Long-term ozone exposure and mortality. *N Engl J Med* 360:1085–1095.
- Keogh RH, White IR. 2013. Using full-cohort data in nested case-control and case-cohort studies by multiple imputation. *Stat Med* 32:4021–4043.
- Kim JS, Alderete TL, Chen Z, Lurmann F, Rappaport E, Habre R, et al. 2018. Longitudinal associations of in utero and early life near-roadway air pollution with trajectories of childhood body mass index. *Environ Health* 17:64.
- Klomp maker JO, Janssen N, Andersen ZJ, Atkinson R, Bauwelinck M, Chen J, et al. 2021. Comparison of associations between mortality and air pollution exposure estimated with a hybrid, a land-use regression and a dispersion model. *Environment Int* 146:106306.
- Klomp maker JO, Janssen NAH, Bloemsma LD, Gehring U, Wijga AH, van den Brink C, et al. 2019. Associations of combined exposures to surrounding green, air pollution, and road traffic noise with cardiometabolic diseases. *Environ Health Perspect* 127:87003.
- Knapp G, Hartung J. 2003 Improved tests for a random effects meta-regression with a single covariate. *Stat Med*. 2003 Sep 15;22(17):2693–710. doi: 10.1002/sim.1482.
- Lippmann M, Chen L-C, Gordon T, Ito K, Thurston GD. 2013. *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.

- Luepker RV, Apple FS, Christenson RH, Crow RS, Fortmann SP, Goff D, et al. 2003. Case definitions for acute coronary heart disease in epidemiology and clinical research studies: A statement from the AHA Council on Epidemiology and Prevention; AHA Statistics Committee; World Heart Federation Council on Epidemiology and Prevention; The European Society of Cardiology Working Group on Epidemiology and Prevention; Centers for Disease Control and Prevention; and the National Heart, Lung, and Blood Institute. *Circulation* 108:2543–2549.
- Nasari MM, Szyszkowicz M, Chen H, Crouse D, Turner MC, Jerrett M, et al. 2016. A class of nonlinear exposure-response models suitable for health impact assessment applicable to large cohort studies of ambient air pollution. *Air Qual Atmos Health* 9:961–972.
- Perperoglou A, Sauerbrei W, Abrahamowicz M, Schmidt M. 2019. A review of spline function procedures in R. *BMC Med Res Methodol* 19:46.
- Pinault L, Brauer M, Crouse DL, Weichenthal S, Erickson A, van Donkelaar A, et al. 2018. Diabetes status and susceptibility to the effects of PM<sub>2.5</sub> exposure on cardiovascular mortality in a national Canadian cohort. *Epidemiology* 29:784–794.
- Pinault L, Tjepkema M, Crouse DL, Weichenthal S, van Donkelaar A, Martin RV, et al. 2016. Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the Canadian Community Health Survey cohort. *Environ Health* 15:18.
- Pinault LL, Weichenthal S, Crouse DL, Brauer M, Erickson A, van Donkelaar A, et al. 2017. Associations between fine particulate matter and mortality in the 2001 Canadian Census Health and Environment Cohort. *Environ Res* 159:406–415.
- Pope CA 3rd, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, et al. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* 287:1132–1141.
- Pope CA 3rd, Coleman N, Pond ZA, Burnett RT. 2020. Fine particulate air pollution and human mortality: 25+ years of cohort studies. *Environ Res* 183:108924.
- Raaschou-Nielsen O, Andersen ZJ, Beelen R, Samoli E, Stafoggia M, Weinmayr G, et al. 2013. 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* 14:813–822.
- Raaschou-Nielsen O, Beelen R, Wang M, Hoek G, Andersen ZJ, Hoffmann B, et al. 2016. Particulate matter air pollution components and risk for lung cancer. *Environ Int* 87:66–73.
- Samoli E, Butland BK. 2017. Incorporating measurement error from modeled air pollution exposures into epidemiological analyses. *Curr Environ Health Rep* 4:472–480.
- Samoli E, Rodopoulou S, Hvidtfeldt UA, Wolf K, Stafoggia M, Brunekreef B et al. 2021. Modeling multi-level survival data in multi-center epidemiological cohort studies: Applications from the ELAPSE project. *Environ Int* 147:106371. doi: 10.1016/j.envint.2020.106371. Epub 2021 Jan 12.
- Schwartz J, Coull B, Laden F, Ryan L. 2008. The effect of dose and timing of dose on the association between airborne particles and survival. *Environ Health Perspect* 116:64–69.
- Shin HH, Cakmak S, Brion O, Villeneuve P, Turner MC, Goldberg MS, et al. 2014. Indirect adjustment for multiple missing variables applicable to environmental epidemiology. *Environ Res* 134:482–487.
- Sorensen M, Andersen ZJ, Nordsborg RB, Becker T, Tjønneland A, Overvad K, et al. 2013. Long-term exposure to road traffic noise and incident diabetes: A cohort study. *Environ Health Perspect* 121:217–222.
- Stafoggia M, Breitner S, Hampel R, Basagaña X. 2017. Statistical approaches to address multi-pollutant mixtures and multiple exposures: The state of the science. *Curr Environ Health Rep* 4:481–490.
- Stafoggia M, Cesaroni G, Peters A, Andersen ZJ, Badaloni C, Beelen R, et al. 2014. Long-term exposure to ambient air pollution and incidence of cerebrovascular events: Results from 11 European cohorts within the ESCAPE project. *Environ Health Perspect* 122:919–925.
- Stafoggia M, Oftedal B, Chen J, Rodopoulou S, Renzi M, Atkinson R, et al. In review. Long-term exposure to low ambient air pollution concentrations and mortality among 28 million subjects: Results from seven European large cohorts within the ELAPSE project. *Lancet Planet Health*.
- Strak M, Janssen N, Beelen R, Schmitz O, Karssenbergh D, Danny Houthuijs D. et al. 2017. Associations between lifestyle and air pollution exposure: Potential for confounding in large administrative data cohorts. *Environ Res* 156:364–373.
- Strak M, Weinmayr G, Rodopoulou S, Chen J, de Hoogh K, Andersen Z, et al. In press. Long-term exposure to low-level air pollution and natural and cause-specific mortality: A pooled analysis of eight European cohorts within the ELAPSE project. *Br Med J*.

- Therneau TM, Grambsch PM. 2000. Modeling survival data: Extending the Cox model. New York:Springer Verlag.
- Thiebaut ACM, Benichou J. 2004. Choice of time-scale in Cox's model analysis of epidemiologic cohort data: A simulation study. *Stat Med* 23:3803–3820.
- Thomsen RW, Lange P, Hellquist B, Frausing E, Bartels PD, Krog BR, et al. 2011. Validity and underrecording of diagnosis of COPD in the Danish National Patient Registry. *Respir Med* 105:1063–1068.
- Thurston GD, Ahn J, Cromar KR, Shao Y, Reynolds HR, Jerrett M, et al. 2016. Ambient particulate matter air pollution exposure and mortality in the NIH-AARP diet and health cohort. *Environ Health Perspect* 124:484–490.
- Tsai MY, Hoek G, Eeftens M, de Hoogh K, Beelen R, Beregszászi T, et al. 2015. Spatial variation of PM elemental composition between and within 20 European study areas — results of the ESCAPE project. *Environ Int* 84:181–192.
- Turner MC, Jerrett M, Pope CA 3rd, Krewski D, Gapstur SM, Diver WR, et al. 2016. Long-term ozone exposure and mortality in a large prospective study. *Amer J Respir Crit Care Med* 193:1134–1142.
- van Donkelaar A, Martin RV, Brauer M, Hsu NC, Kahn RA, Levy RC, et al. 2016. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. *Environ Sci Tech* 50:3762–3772.
- van Donkelaar A, Martin RV, Spurr RJ, Burnett RT. 2015. High-resolution satellite-derived PM<sub>2.5</sub> from optimal estimation and geographically weighted regression over North America. *Environ Sci Tech* 49:10482–10491.
- Vedal S, Campen MJ, McDonald JD, Kaufman JD, Larson TV, Sampson PD, et al. 2013. National Particle Component Toxicity (NPACT) Initiative Report on Cardiovascular Effects. Research Report 178. Boston, MA:Health Effects Institute.
- Villeneuve PJ, Weichenthal SA, Crouse D, Miller AB, To T, Martin RV et al. 2015. Long-term exposure to fine particulate matter air pollution and mortality among Canadian women. *Epidemiology* 26:536–545.
- Vodonos A, Awad YA, Schwartz J. 2018. The concentration–response between long-term PM<sub>2.5</sub> exposure and mortality; A meta-regression approach. *Environ Res* 166:677–689.
- Wang M, Beelen R, Eeftens M, Meliefste K, Hoek G, Brunekreef B. 2012. Systematic evaluation of land use regression models for NO<sub>2</sub>. *Environ Sci Tech* 46:4481–4489.
- Wang M, Beelen R, Stafoggia M, Raaschou-Nielsen O, Andersen ZI, Hoffmann B, et al. 2014. 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* 66:97–106.
- Wang R, Henderson SB, Sbihi H, Allen RW, Brauer M. 2013. Temporal stability of land use regression models for traffic-related air pollution. *Atmos Environ* 64:312–319.
- WHO (World Health Organization). 2005. Air quality guidelines — a global update. Copenhagen, Denmark: WHO.
- WHO (World Health Organization). 2013. Review of evidence on health aspects of air pollution — REVIHAAP project: Technical report. Copenhagen, Denmark: WHO.
- WHO (World Health Organization). 2018. Environmental noise guidelines for the European region (2018). Copenhagen, Denmark: WHO.
- Wolf K, Hoffmann B, Andersen Z, Atkinson R, Bauwelinck M, Bellander T, et al. In press. Long-term exposure to low-level ambient air pollution and incidence of stroke and acute coronary heart disease: Results from six European cohorts within the ELAPSE Project. *Lancet Planet Health*.
- Wolf K, Stafoggia M, Cesaroni G, Andersen ZJ, Beelen R, Galassi C, et al. 2015. Long-term exposure to particulate matter constituents and the incidence of coronary events in 11 European cohorts. *Epidemiology* 26:565–574.
- Yang WS, Zhao H, Wang X, Deng Q, Fan WY, Wang L. 2016. An evidence-based assessment for the association between long-term exposure to outdoor air pollution and the risk of lung cancer. *Eur J Cancer* P25:163–172.

---

#### HEI QUALITY ASSURANCE STATEMENT

---

The conduct of this study was subjected to independent audits by RTI International staff members Dr. Linda Brown and Dr. Prakash Doraiswamy. These staff members are experienced in quality assurance (QA) oversight for air quality monitoring, emission inventories and modeling, data analysis, and epidemiological methods and analysis. The RTI QA oversight team also included statistician Dr. David Wilson who reviewed the exposure and epidemiological statistical model codes.

The QA oversight program consisted of an initial on-site audit of the research study at Utrecht University for con-

formance to the study protocol and standard operating procedures and a final remote audit of the final report and the data processing steps. The onsite audit was performed by Drs. Brown and Doraiswamy. The final remote audit was performed by Drs. Brown, Doraiswamy, and Wilson. Key details are presented below.

#### **BBB On-Site Audit, November 1–3, 2017**

The auditors conducted an on-site audit in the Netherlands at the Institute for Risk Assessment Sciences (IRAS), Utrecht University, and at the National Institute for Public Health and the Environment (RIVM). The audit at IRAS reviewed the following study components: progress reports, personnel and staff, adequacy of equipment and facilities, internal quality assurance procedures, air quality data processing and documentation, health data processing and quality checks, and backup procedures. Program codes were inspected to verify proper documentation. Analytic datasets and codebooks were examined. The audit included an observation of the demonstration of the code executions, file tree structure on the server, and model diagnostics. No errors were noted, but recommendations were made for updating the study plan and expanding the quality plan, documenting codes, and documenting procedures and assumptions related to model development and QA/QC (quality assurance/quality control). The audit at RIVM included a review of security procedures and the process for submitting data requests to Statistics Netherlands. A sample of de-identified data from a related study was viewed, and QC procedures for assignment of residences and of air pollution to each residential grid were discussed.

#### **BBB Final Remote Audit, June–August 2021**

(Audit phase 1, IRAS, Utrecht University, and November 2, afternoon at RIVM)

The final remote audit consisted of two parts: (a) review of the final project report, and (b) audit of data processing steps. The final report review focused on ensuring that it is well documented and easy to understand and highlighted key study findings and limitations. This review also provided guidance on specific aspects of the data processing sequence that could be reviewed remotely. The data audit included a remote live demonstration of selected data processing code and the review of the codes for data reduction, processing and analysis, model development, and visualization. This specific portion of the audit was restricted to the key components of the study and associated findings. Selected scripts (in R/SAS) for final model development, parts of data reduction, and generation of key plots were sent to RTI. No data were sent to RTI due to

data confidentiality restrictions. Therefore, data inputs to the codes were not available.

The codes for the Pooled and Administrative Cohorts were reviewed at RTI to verify, to the extent feasible, linkages between the various scripts, confirmation of the models reported, and verification of key plots. The codes appear to be consistent with the models and key plots described in the report and followed the overall model development procedure described; the values themselves could not be verified due to unavailability of the input data.

The remote live demonstration at Utrecht University included a real-time execution of selected codes generating key tables and figures in the report. Tables and plots related to the pooled cohorts were observed and verified. Data for the Administrative Cohorts were not available at Utrecht University and hence codes related to analysis of Administrative Cohort findings could not be demonstrated.

Tables and figures generated during the real-time demonstration matched perfectly with the values in the report. No major quality-related issues were identified from the review of the codes and the report. Minor editorial comments and recommendations were made for improved clarity.

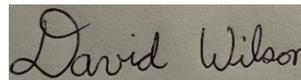
Written report was provided to HEI. The QA oversight audit demonstrated that the study was conducted according to the study protocol. The final report, except as noted in the comments, appears to be representative of the study conducted.



Linda Morris Brown, MPH, DrPH, Epidemiologist, Quality Assurance Auditor



Prakash Doraiswamy, PhD, Air Quality Specialist, Quality Assurance Auditor



David Wilson, PhD, Statistician, Quality Assurance Auditor

August 20, 2021

---

**MATERIALS AVAILABLE ON THE HEI WEBSITE**


---

Additional Materials 1 through 4 (available on the HEI website, [www.healtheffects.org/publications](http://www.healtheffects.org/publications)) contain supplemental material not included in the printed report.

Additional Materials 1. Appendix 1: ELAPSE Methods

Additional Materials 2. Appendix 2: Additional Results from the Pooled Cohort

Additional Materials 3. Appendix 3: Additional Results from the Administrative Cohorts

Additional Materials 4. Appendix 4: Methods and Results from the Particle Composition Analyses

---

**ABOUT THE AUTHORS**


---

**Bert Brunekreef** has a PhD from Wageningen University, the Netherlands. He is a professor in the Institute for Risk Assessment Sciences, Utrecht University, the Netherlands.

**Maciej Strak** has a PhD from Utrecht University, the Netherlands. Until mid-2019 he was a researcher in the Institute for Risk Assessment Sciences, Utrecht University, the Netherlands; currently he is a senior scientist in the National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands.

**Jie Chen** is a PhD student at the Institute for Risk Assessment Sciences, Utrecht University, the Netherlands.

**Zorana Andersen** has a PhD from University of Copenhagen, Denmark, where she is a professor in Environmental Epidemiology and leader of the Environmental Epidemiology Group, at the Section of Environment and Health, Department of Public Health. Andersen has mainly worked on studies of health effects of short- and long-term exposures to outdoor air pollution. She is currently leading several projects on health effects related to exposure to long-term exposure air pollution, road-traffic noise, wind turbine noise, green space, and light at night in the Danish Nurse Cohort.

**Richard Atkinson** has a PhD from University of London, UK. He is a professor of epidemiology in the Population Health Research Institute, St George's, University of London, UK. Atkinson has worked on studies of health effects of both short- and long-term exposures to outdoor air pollution as well as undertaking systematic reviews and meta-analyses of air pollution studies.

**Mariska Bauwelinck** is a PhD student in the research group Interface Demography in the Department of

Sociology, Vrije Universiteit Brussel, Belgium. Her doctoral dissertation is about the relationships among air pollution, green spaces, and different cardiovascular health outcomes in adults.

**Tom Bellander** has a PhD from the Lund University, Sweden, and is a professor at the Institute of Environmental Medicine, Karolinska Institutet, Sweden. His research focuses on environmental epidemiology and in particular the use of population exposure data.

**Marie-Christine Boutron-Ruault**, MD, PhD, is a senior Inserm research director in epidemiology and past head (currently co-head) of the Exposome, Heredity, Cancer, and Health team at Gustave Roussy, Villejuif, France. She has been PI (currently co-PI) of a 100,000-women cohort — the E3N study — and is the author of over 670 peer-reviewed papers in epidemiology of noncommunicable diseases, including cancer, cardiovascular diseases, asthma, inflammatory bowel diseases, and rheumatology.

**Jørgen Brandt** is professor of atmospheric modeling and Climate Centre director in the Department of Environmental Science, Aarhus University, Denmark. He has a PhD from the University of Copenhagen, Denmark. He is an expert in multiscale atmospheric modeling and health impact assessment and has authored more than 140 reviewed journal papers.

**Iain Carey** has a PhD from University of London, UK. He is a senior lecturer in epidemiology in the Population Health Research Institute, St George's, University of London, UK. Carey has worked on a wide range of studies utilizing primary care databases in the UK, including studies of links between health effects and long-term exposures to outdoor air pollution.

**Giulia Cesaroni**, MSc in epidemiology (London School of Hygiene & Tropical Medicine) and math (Sapienza University, Rome), is a senior researcher at the Population Health Unit, Department of Epidemiology of the Lazio Region in Rome, Italy. She is primarily involved in the implementation and analysis of the Rome Longitudinal Study for investigating the health impact of socioeconomic factors and long-term exposure to air pollution.

**Francesco Forastiere**, MD (occupational medicine) and PhD (epidemiology), is an associate researcher at the National Research Council (CNR) in Italy and a visiting professor at King's College, London. He has served as a public health officer for the Department of Epidemiology of the Lazio Region, Rome, focusing on the application of scientific research findings to a wide range of public health issues, including air pollution.

**Daniela Fecht**, PhD, is a lecturer in geospatial health at the Medical Research Council Centre for Environment and Health, School of Public Health, Imperial College London. Her research focuses on geographical variations in population health, with particular focus on environmental and health inequalities. She has more than 15 years' experience in exposure science and has authored more than 50 publications.

**John Gulliver**, PhD, is professor of environmental and exposure sciences and deputy director of the Centre for Environmental Health and Sustainability at the University of Leicester, UK. He has 25 years of experience in environmental modeling and exposure assessment for epidemiological studies with a focus on long-term effects of noise and air pollution.

**Ole Hertel** has a DSc in environmental chemistry from Copenhagen University, Denmark, and a DSc in geophysics from Bergen University, Norway. He is a professor and deputy head of the Department of Environmental Science, Aarhus University, Denmark. He has authored more than 110 peer-reviewed journal papers.

**Barbara Hoffmann** has an MD from the University of Aachen and master's of public health from the University of Bielefeld, Germany. She is a professor of environmental epidemiology at the Institute for Occupational, Social, and Environmental Medicine at the Heinrich Heine University of Düsseldorf, Germany.

**Kees de Hoogh**, PhD, is scientific project leader at the Environmental Exposures and Health Unit, Swiss Tropical and Public Health Institute, Basel, Switzerland. He specializes in spatial modeling and exposure assessment for studies on environment and health and has been (BioSHARE, ESCAPE) and continues to be (ELAPSE, EXPANSE) involved in large EU projects, as a coleader in air pollution exposure assessment.

**Danny Houthuijs**, MSc, is senior researcher at the National Institute for Public Health and the Environment (RIVM), the Netherlands, in the field of effects of environmental stressors on health and well-being.

**Ulla Hvidtfeldt**, PhD, is a senior research scientist at the Danish Cancer Society Research Center.

**Nicole Janssen** is senior researcher at the National Institute for Public Health and the Environment (RIVM). She has over 25 years' experience in performing research in environmental exposure assessment and epidemiology.

She has been the principal investigator for several nationally funded studies in the field of environmental health, including studies using the Dutch administrative cohort (DUELS). She has (co)authored over 100 scientific articles in peer-reviewed journals.

**Jeanette Jørgensen** is a PhD student at the University of Copenhagen, Section of Environmental Health, where she is studying night shift work and risk of dementia, Parkinson's disease, and psychological disorders in the Danish Nurse Cohort. Jørgensen has also been exploring health effects related to long-term exposure to air pollution and road-traffic noise in the Danish Nurse Cohort.

**Klea Katsouyanni**, PhD, was the coordinator of the EU network Agency for Public Health Education Accreditation, which provides Europe-wide results on the short-term effects of air pollution. She has participated in many projects investigating air pollution and other environmental exposure effects (including ESCAPE). She has been a member of several international advisory committees (European Commission and World Health Organization). In 2006 she received the International Society for Environmental Epidemiology John Goldsmith award.

**Matthias Ketzel** is a senior scientist at the Department of Environmental Science, Aarhus University, Roskilde, Denmark, and visiting professor at the Global Centre for Clean Air Research, University of Surrey, UK. He received his PhD from Lund University, Sweden. His main research topics are atmospheric dispersion, aerosol dynamics modeling, and human exposure estimation at regional, urban, and street scale.

**Jochem Klompaker** has a PhD from Utrecht University, the Netherlands. From 2015 to 2020, he worked at the Dutch National Institute for Public Health and the Environment and the Institute for Risk Assessment Sciences, Utrecht University. In 2020 he became a postdoctoral researcher at the Harvard T.H. Chan School of Public Health.

**Norun Hjertager Krog** has a PhD in environmental epidemiology from the University of Oslo, Norway. She is a senior researcher at the Norwegian Institute of Public Health, Oslo, Norway.

**Shuo Liu** is a PhD student in the Department of Public Health, University of Copenhagen, Denmark. He is working on the association of long-term exposure to air pollution with COPD and adult-onset asthma.

**Petter Ljungman**, MD, PhD, is associate professor of epidemiology and a senior consultant and specialist in cardiology. He defended his PhD at Karolinska Institute on air pollution effects on cardiovascular diseases and conducted postdoctoral research at the Harvard T.H. Chan School of Public Health. He is principal investigator for one of the participating cohorts (Cardiovascular Effects of Air pollution and Noise Study).

**Amar Mehta**, SD, is an associate professor at the Section of Epidemiology, Department of Public Health, University of Copenhagen, with joint appointment at Statistics Denmark. His research focuses on the cardiovascular and respiratory health effects of air pollution and occupational exposures.

**Gabriele Nagel**, PhMD, is associate professor of epidemiology at the Ulm University (Germany) and senior consultant at the Agency for Preventive and Social Medicine, Bregenz, Austria. Her research focuses on population studies and cancer epidemiology.

**Bente Oftedal** has a PhD in environmental epidemiology from the Oslo University, Norway. She is a senior researcher at the Norwegian Institute of Public Health, Oslo, Norway.

**Göran Pershagen**, MD, has a PhD from Karolinska Institutet, Stockholm, where he is professor at the Institute of Environmental Medicine. He is also a senior physician at the Centre for Occupational and Environmental Medicine, Region Stockholm, Sweden.

**Annette Peters**, PhD, is director of the Institute of Epidemiology at Helmholtz Zentrum München and full professor at Ludwig Maximilians Universität München, Germany. She studied biology and mathematics in Germany and epidemiology at the Harvard School of Public Health in Boston. She is an environmental and molecular epidemiologist directing the KORA Cohort Study and chairing the board of the German National Cohort.

**Ole Raaschou-Nielsen**, MSc, PhD, is head of the research group Work, Environment and Cancer, at the Danish Cancer Society and is a professor at Aarhus University. His research focuses on environmental exposures such as air pollution, radon, noise, cadmium, and persistent organic pollutants in relation to risk for cancer and other health endpoints.

**Matteo Renzi**, MSc, PhD, is currently involved in a post-graduate specialization course in medical statistics and biometry at the University of Rome (“La Sapienza”), Italy. He is a researcher at the Department of Epidemiology of the Lazio Region in Rome, Italy. In ELAPSE he has been primarily involved in the analysis of the Rome Longitudinal Study and in the meta-analysis of the administrative cohorts’ results.

**Sophia Rodopoulou** has a PhD from National and Kapodistrian University of Athens (NKUA), Greece. She is a research associate at the Department of Hygiene, Epidemiology and Medical Statistics, Medical School, NKUA, and her research focuses on the health effects of environmental stressors, mainly outdoor air pollution and temperature, as well as on statistical methodological issues concerning environmental epidemiology.

**Evi Samoli** is an assistant professor of epidemiology and medical statistics at the National and Kapodistrian University of Athens (NKUA). She has a PhD from the Medical School, NKUA and her research focuses on health effects from air pollution and statistical methods for epidemiological data analysis. She has published more than 125 articles and participated in more than 25 research projects.

**Per Schwarze** is a specialist director for research and health analysis at the Division of Infection Control and Environmental Health, Norwegian Institute of Public Health Oslo, Norway. He is involved in research projects on infectious diseases and the effects of environmental factors on health, using both experimental and epidemiological methods.

**Torben Sigsgaard** MD, has a PhD from Aarhus University, Denmark. He is a professor of environmental & occupational medicine, Dept of Public Health, Aarhus University. He has worked on studies of the health effects from both short- and long-term exposures to occupational, indoor, and outdoor air pollution.

**Massimo Stafoggia** is currently in a PhD program at Karolinska Institutet, Stockholm. He is a senior researcher at the Department of Epidemiology of the Lazio Region in Rome, Italy. In ELAPSE he has been mostly involved in the analysis of the Rome Longitudinal Study and the meta-analysis of the administrative cohorts’ results. Also, he has contributed to the statistical working group.

**Danielle Vienneau**, PhD, is scientific project leader in the Environmental Exposures and Health Unit, Swiss Tropical and Public Health Institute, Basel, Switzerland. Working at the interface between environmental exposures and health effects, she is a key investigator using the Swiss National Cohort to study mortality in relation to environmental exposures including air pollution and transportation noise.

**Gudrun Weinmayr** is a senior researcher with a background in general biology and an MPH. She has been working in air pollution epidemiology for 15 years on various health outcomes. Before the current study, she participated in the ESCAPE and TRANSPHORM studies with the large Vorarlberg cohort VHM&PP and served as member of the statistical working group.

**Kathrin Wolf** has a PhD from Ludwig-Maximilians University, Germany. She is a researcher at the Institute of Epidemiology, Helmholtz Zentrum München Germany. Her main research interests include the impact of short- and long-term effects of air pollution, noise, and meteorology on cardiometabolic outcomes.

**Gerard Hoek** has a PhD from Wageningen University, the Netherlands. He is an associate professor at the Institute for Risk Assessment Sciences, Utrecht University, the Netherlands.

---

#### OTHER PUBLICATIONS RESULTING FROM THIS RESEARCH

---

Hvidtfeldt UA, Severi G, Andersen ZJ, Atkinson R, Bauwelinck M, Bellander T, et al. 2021a. Long-term low-level ambient air pollution exposure and risk of lung cancer — A pooled analysis of 7 European cohorts. *Env Int* 146:106249.

Hvidtfeldt UA, Chen J, Andersen ZJ, Atkinson R, Bauwelinck M, Bellander T, et al. 2021b. Long-term exposure to fine particle elemental components and lung cancer incidence in the ELAPSE pooled cohort. *Environ Res* 193:110568.

Klompmaaker JO, Janssen N, Andersen ZJ, Atkinson R, Bauwelinck M, Chen J, et al. 2021. Comparison of associations between mortality and air pollution exposure estimated with a hybrid, a land-use regression and a dispersion model. *Environ Int* 146:106306.

Liu S, Jørgensen JT, Ljungman P, Pershagen G, Bellander T, Leander K, et al. 2021. Long-term exposure to low-level air pollution and incidence of chronic obstructive pulmonary disease: The ELAPSE project. *Environ Int* 146:106267.

Samoli E, Rodopoulou S, Hvidtfeldt UA, Wolf K, Stafoggia M, Brunekreef B, et al. 2021. Modeling multi-level survival data in multi-center epidemiological cohort studies: Applications from the ELAPSE project. *Environ Int* 147:106371.

Chen J, de Hoogh K, Gulliver J, Hoffmann B, Hertel O, Ketzel M, et al. 2020. Development of Europe-wide models for particle elemental composition using supervised linear regression and random forest. *Environ Sci Technol*, 2020, 54 (24): 15698–15709

Liu S, Jørgensen JT, Ljungman P, Pershagen G, Bellander T, Leander K, et al. 2020. Long-term exposure to low-level air pollution and incidence of asthma: The ELAPSE project. *Eur Respir J* 2003099.

Chen J, de Hoogh K, Gulliver J, Hoffmann B, Hertel O, Ketzel M, et al. 2019. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. *Environ Int* 130:104934; doi:10.1016/j.envint.2019.104934.

de Hoogh K, Chen J, Gulliver J, Hoffmann B, Hertel O, Ketzel M, et al. 2018. Spatial PM<sub>2.5</sub>, NO<sub>2</sub>, O<sub>3</sub> and BC models for Western Europe — Evaluation of spatiotemporal stability. *Environ Int* 120:81–92; doi:10.1016/J.ENVINT.2018.07.036.

Chen J, Rodopoulou S, de Hoogh K, Strak M, Andersen ZJ, Atkinson R, et al. 2021. 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. *Environmental Health Perspectives* 129(4):047009; doi: 10.1289/EHP8368

Research Report 208, *Mortality and Morbidity Effects of Long-Term Exposure to Low-Level PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub>: An Analysis of European Cohorts — Effects of Low-Level Air Pollution: A Study in Europe (ELAPSE)*, Brunekreef et al.

---

## INTRODUCTION

---

Ambient air pollution is an important contributor to the global burden of disease (GBD 2020; HEI 2020). Although levels of air pollution have been declining over the past few decades in many parts of the world, several studies published in the past decade reported associations between risk of mortality and long-term exposures to fine particulate matter (PM<sub>2.5</sub>\*) at even relatively low concentrations (e.g., Beelen et al. 2014a, b; Crouse et al. 2012, 2015; Hales et al. 2012; Pinault et al. 2016). To inform future risk assessment and regulation, it is important to confirm whether associations with adverse health effects continue to be observed as levels of air pollution decline still further. It is also important to understand better the shape of the concentration–response function at those low levels. Both of these issues remain as major uncertainties for setting air quality standards. The growing body of evidence demonstrating effects on health associated with exposures to air pollution at levels below current air quality standards in North America and Europe, the large overall contributions of air pollution to the global burden of disease, and general interests in reducing greenhouse gas emissions suggest that stronger air quality standards and guidelines may be considered in the future.

As described in detail in the Preface to this Report, in 2016 HEI funded three studies under RFA 14-3 to explore this issue of effects on health associated with exposures to low concentrations of air pollution. Dr. Brunekreef’s

ELAPSE study was one of these three studies. Additional information about the RFA and the two other studies conducted in North America is included in the Preface.

This Commentary was prepared by the HEI Low-Exposure Epidemiology Studies Review Panel and members of the HEI Scientific Staff, who were convened to review these three HEI-funded studies. The Commentary includes the scientific and regulatory background for the research, a summary of the study’s approach and key results, and the Panel’s evaluation of the Investigators’ Report (IR) highlighting strengths and weaknesses of the study.

This Commentary is intended to aid the sponsors of HEI and the public by highlighting both the strengths and limitations of the study and by placing the Investigators’ Report into scientific and regulatory context.

---

## SCIENTIFIC AND REGULATORY BACKGROUND

---

The setting of ambient air quality standards — at levels considered adequate to protect public health — is a central component of programs designed to reduce air pollution and improve public health under the U.S. Clean Air Act (U.S. CAA), the European Union (EU) Ambient Air Quality Directives, the World Health Organization (WHO) Air Quality Guidelines, and similar measures around the world. Although the process for setting such standards varies, they all contain several common components:

- identifying, reviewing, and synthesizing the scientific evidence on sources, exposures, and health effects of air pollution;
- conducting risk and policy assessments to estimate what public health effects are likely to be seen at different levels of the standard;
- identifying and setting standards based on the risk analysis;
- air quality monitoring to determine areas that do not meet the standards; and
- implementing air quality control interventions to meet the standards by reducing exposure.

---

Dr. Brunekreef’s 3-year study, “Effects of Low-Level Air Pollution: A Study in Europe (ELAPSE),” began in May 2016. Total expenditures were \$2,089,795. The draft Investigators’ Report from Brunekreef and colleagues was received for review in May 2020. A revised report was submitted in November and reviewed in December by the HEI Low-Exposure Epidemiology Studies Review Panel, who suggested a few more minor revisions. A second revised final report was submitted and accepted for publication in January 2021. During the review process, the Panel and the investigators had the opportunity to exchange comments and to clarify issues in both the Investigators’ Report and the Panel’s Commentary.

This document has not been reviewed by public or private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views of these parties, and no endorsements by them should be inferred.

\* A list of abbreviations and other terms appears at the end of this volume.

## SETTING AIR QUALITY STANDARDS IN THE UNITED STATES

The U.S. CAA requires that in setting the National Ambient Air Quality Standards (NAAQS), the U.S. Environmental Protection Agency (U.S. EPA) Administrator reviews all available science and sets the NAAQS for all major (criteria) pollutants (e.g., PM, nitrogen dioxide [NO<sub>2</sub>], and ozone [O<sub>3</sub>]) at a level “requisite to protect the public health with an adequate margin of safety.” In practice, that review has had two principal steps:

- Synthesis and evaluation of all available science in what is now called an Integrated Science Assessment. This document reviews the widest range of exposure, dosimetry, toxicology, mechanistic, clinical, and epidemiologic evidence. It then — according to a predetermined set of criteria (U.S. EPA 2015) — draws on all lines of evidence to make a determination of whether the exposure is causal, likely to be causal, or suggestive for a series of health outcomes.
- Assessment of the risks based on that science is then conducted in a Risk and Policy Assessment. This further analysis draws on the Integrated Science Assessment to identify the strongest evidence — most often from human clinical and epidemiological studies — of the lowest concentration levels at which health effects are observed, the likely implications of such levels for health across the population, and the degree to which the newest evidence suggests that there are effects observed below the then-current NAAQS for a particular pollutant.

The Risk and Policy Assessment also examines the uncertainties around estimates of health impacts, and the shape of the concentration–response curve, especially at levels near and below the then-current NAAQS. Although a range of possible shapes of the concentration–response curves has been considered, including whether there is a threshold at a level below which effects are not likely, the U.S. EPA’s conclusions in these reviews thus far have not found evidence of a threshold (although studies to date have not always had the power to detect one) (U.S. EPA 2004, 2013). Also, although the standard is set ostensibly to protect public health with an adequate margin of safety, it has been generally understood that there are likely additional, albeit more uncertain, health effects of exposure below the NAAQS.

Both of these documents are subjected to extensive public comments and review by the Clean Air Scientific Advisory Committee, which was established under the U.S. CAA. The Clean Air Scientific Advisory Committee is charged with both peer-reviewing the documents, which includes advising the Administrator on the strength and uncertainties in the science as well as on making the

decision whether to retain or change the NAAQS. The current longer-term NAAQS for PM<sub>2.5</sub>, NO<sub>2</sub>, and O<sub>3</sub> are as follows (<https://www.epa.gov/criteria-air-pollutants/naaqs-table>):

- PM<sub>2.5</sub>: annual mean averaged over three years of 12 µg/m<sup>3</sup>;
- NO<sub>2</sub>: annual mean of 53 ppb (approximately 100 µg/m<sup>3</sup>);
- O<sub>3</sub>: annual fourth-highest daily maximum 8-hour concentration, averaged over three years of 70 ppb (approximately 140 µg/m<sup>3</sup>).

## SETTING AIR QUALITY STANDARDS IN EUROPE

Similar to the United States and around the world, the EU has the overarching goal to protect its citizens and susceptible subpopulations from the adverse effects of major ambient air pollutants. The EU’s clean air policy is based on three main pillars (European Commission 2018): (1) the Ambient Air Quality Directives (European Union 2004, 2008), which set out air quality standards and require Member States to assess air quality in a harmonized and comparable manner and to implement air quality plans to improve or maintain the quality of air; (2) the National Emissions Ceiling Directive (European Union 2016), which establishes national emission reduction commitments; and (3) source-specific legislation establishing specific emission standards for key sources of air pollution.

The directives have established EU air quality standards in the form of limit values and target values. Limit values are not to be exceeded and target values are to be attained where possible. The air quality standards currently in place were established in two complementary Ambient Air Quality Directives in 2004 and 2008 (European Union 2004, 2008).

In the 2008 directive, a limit value for PM<sub>2.5</sub> was added for the first time. The limit value for PM<sub>2.5</sub> is an annual average of 25 µg/m<sup>3</sup>, which is higher than the U.S. annual standard of 12 µg/m<sup>3</sup> and the WHO health-based guideline of 10 µg/m<sup>3</sup>, established in 2005 (World Health Organization 2006). The limit value for NO<sub>2</sub> is an annual average of 40 µg/m<sup>3</sup>, which is more stringent than that in the United States.

Although the U.S. EPA is mandated by law to conduct comprehensive reviews on individual pollutants on a regular basis, there is no such regular process in Europe, and the EU relies on the WHO and others for the rigorous evaluation of the scientific evidence to inform retaining or tightening the air quality standards.

Just recently, in 2019, a fitness check of the EU Ambient Air Quality Directives was published (European Commission 2019a). It assessed whether or not all the directives’ provisions are fit for purpose, looking in particular at the monitoring and assessment methods, the air quality stan-

dards, the provisions for public information, and the extent to which the directives facilitated action to prevent or reduce adverse impacts. The fitness check applied five criteria: relevance, effectiveness, efficiency, coherence, and EU added value. Based on this fitness check, the Commission concluded that the Ambient Air Quality Directives have been partly effective in improving air quality and achieving air quality standards.

Also in 2019, the European Green Deal (European Commission 2019b) was published. This is the European Commission's response to the climate and environmental challenges Europe (and the world) is facing. The European Green Deal has the overarching aim of reducing sources of carbon dioxide sufficiently to make Europe climate neutral by 2050. It also aims for a zero-emission strategy of air pollutants and foresees a revision of the European Air Quality Directive. To reach this target, the Commission will adopt a zero-pollution action plan for air, water, and soil in 2021.

In line with the conclusions from the fitness check, it is expected that the Commission in its plan will propose to revise air quality standards so that they align more closely with the WHO recommendations, which are due to be updated in 2021 (EEA, 2020). More information on the existing legislation in the EU can be found here: [https://ec.europa.eu/environment/air/quality/existing\\_leg.htm](https://ec.europa.eu/environment/air/quality/existing_leg.htm).

#### THE ADVENT OF HEI STUDIES OBSERVING ASSOCIATIONS BELOW CURRENT AIR QUALITY STANDARDS AND GUIDELINES

As the quality and availability of data on levels of PM<sub>2.5</sub> improved over the course of the first decade of this century, results from new studies began to emerge starting in 2012 (e.g., in Canada and New Zealand) suggesting that associations of PM<sub>2.5</sub> and mortality could be observed down to levels well below the NAAQS of 12 µg/m<sup>3</sup> and the EU limit value of 25 µg/m<sup>3</sup> (Crouse et al. 2012; Hales et al. 2012). These studies found robust associations, with some evidence of even steeper slopes of effect at the lowest levels. If replicated in other populations and by other investigators, these findings could change the basis for future determinations of the levels at which to set the NAAQS as well as EU and other air quality standards.

At the same time, these findings posed several questions, for example:

- Would the results be robust to the application of a range of alternative analytic models and their uncertainty?
- Could other important determinants of population health, such as age, socio-economic position, health status, and access to medical care, as well as differ-

ences in air pollution sources and time-activity patterns, modify or confound the associations seen?

- Would the results change if risk estimates were more fully corrected for the effects of important potential confounding variables, such as smoking, in the absence of such data at the individual level?
- What might be the effects of co-occurring pollutants on health effect associations at low ambient concentrations?

As described in the Preface, the advent of these studies and the desire to address these important questions were the basis for the HEI Request for Applications (RFA 14-3) that sought and ultimately supported this study by Dr. Brunekreef and colleagues and the two other studies that make up HEI's Program to Assess Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution.

---

#### STUDY SUMMARY

---

##### STUDY OBJECTIVES

The main purpose of the ELAPSE study was to examine whether exposure to low concentrations of ambient air pollution is associated with adverse effects on human health, where "low" is defined as concentrations lower than the current EU Limit Values, U.S. EPA NAAQS, and/or the WHO Air Quality Guideline values for PM<sub>2.5</sub>, NO<sub>2</sub>, and O<sub>3</sub>. The study had four specific objectives, namely:

1. To estimate long-term average exposure to PM<sub>2.5</sub>, black carbon (BC), NO<sub>2</sub>, O<sub>3</sub>, and PM<sub>2.5</sub> composition by developing new hybrid models that combine monitoring data, land use, satellite observations, and dispersion models for participants in a pooled cohort consisting of participants from 15 existing cohorts from the European Study of Cohorts for Air Pollution Effects (ESCAPE) and in seven large administrative cohorts.
 

To investigate the shape of the relationship between long-term exposure to PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub> and four broad health effect categories — (1) natural- and cause-specific mortality including cardiovascular and nonmalignant as well as malignant respiratory and diabetes mortality; and morbidity measured as (2) coronary and cerebrovascular events; (3) lung cancer incidence; and (4) asthma and chronic obstructive pulmonary disease (COPD) incidence — using a number of different methods to characterize the concentration-response function (linear, nonlinear, or threshold).
2. To investigate, in the context of the second objective, variability of the concentration-response function

across populations and different exposure assessment methods, as well as the impact of different methods for addressing exposure measurement error, the role of co-occurring pollutants, and the effect of indirect approaches for confounder control in administrative cohorts.

3. To compare epidemiological effect estimates between those obtained using the ELAPSE and Mortality–Air Pollution Associations in Low-Exposure Environments (MAPLE) exposure modeling frameworks, respectively (MAPLE is the Canadian companion study funded by HEI within RFA14-3).

Briefly, the ELAPSE study consists of two parallel sets of epidemiologic analysis. As described below, Brunekreef and colleagues created Europe-wide exposure models for all pollutants of interest and assigned estimates of exposure to participants in two sets of cohorts. The first set includes 15 well-characterized cohorts pooled together. Fourteen of these cohorts were analyzed previously as part of the ESCAPE project (see Sidebar). The second set of cohorts includes seven large European administrative cohorts analyzed individually, and with their epidemiological results meta-analyzed to produce summary effect estimates. The purpose of working with these two cohort groups was to address the various strengths and limitations of using each individually. Specifically, the key strength of the pooled cohorts is the rich amount of individual-level information available for participants (despite somewhat smaller sample sizes), whereas the key strength of the administrative cohorts is their large sample sizes and national representativeness (at the expense of having fewer individual-level variables for participants).

Health outcomes in this study included mortality from all-natural causes, cardiovascular disease, diabetes,

cardiometabolic disorders, and respiratory diseases, as well as incidences of lung cancer, coronary and cerebrovascular events, asthma, and COPD. The investigators applied standard Cox proportional hazard models to describe associations between exposures to the pollutants and these outcomes. They considered many sensitivity analyses, related broadly to different approaches to exposure specification, confounder control, and exploration of concentration–response functions. Note that not all analyses nor health outcomes were examined in both cohort groups.

## METHODS AND STUDY DESIGN

### Exposure Modeling

**Main Models** The investigators' main analyses involved the development of Europe-wide, hybrid land use regression (LUR) exposure models based on a consistent modeling approach for the whole area. These models are considered “hybrid” because they included outputs from dispersion models and satellite data along with ground-based observations and land use data. Both LUR and dispersion modeling are alternative approaches for modeling air pollution patterns that have only recently been combined using hybrid approaches (Hoek 2017). Briefly, LUR is a spatial modeling technique that uses observations of pollutant concentrations from point locations as the dependent variable and data describing characteristics such as road density and land use as the independent variables in a multivariate regression model to estimate pollutant concentrations at unsampled locations. Dispersion models estimate air pollution patterns by combining data on emission sources, geophysical characteristics of the area, and meteorological conditions. These models assume deterministic relation-

## WHAT IS ESCAPE?

The ESCAPE project was a collaboration of more than 30 existing European cohort studies. The study took place between 2008 and 2012 and was led by Bert Brunekreef. Briefly, the study sought to quantify health impacts of exposure to air pollution and considered the effects of within-city and within-area contrasts in exposure as well as many different health outcomes. Among the key features was the harmonized exposure assessment, including an extensive air pollution measurement campaign across all areas. Key differences

between the analyses conducted as part of the ELAPSE study compared to the ESCAPE project are summarized in Commentary Table 1. Specifically, in the ELAPSE study, the investigators conducted a pooled analysis (instead of cohort-specific analyses), developed Europe-wide exposure models (instead of local exposure models), incorporated residential history into the exposure assessment, added O<sub>3</sub> to the list of pollutants examined, and incorporated updated mortality and morbidity data (Commentary Table 1).

**Commentary Table 1.** Selected Study Characteristics Between the ESCAPE, ELAPSE Pooled, and ELAPSE Administrative Analyses

Study Characteristics	ESCAPE	ELAPSE Pooled	ELAPSE Administrative
Study design	Cohort-specific analyses and meta-analyses	Pooled cohort analysis	Cohort-specific analyses and meta-analyses
Pollutants	PM <sub>2.5</sub> , BC, NO <sub>2</sub> , PM <sub>10</sub> , coarse PM, NO <sub>x</sub> , and the copper, iron, zinc, and sulfur content of PM <sub>2.5</sub>	PM <sub>2.5</sub> , BC, NO <sub>2</sub> , O <sub>3</sub> , and the copper, iron, zinc, and sulfur content of PM <sub>2.5</sub>	PM <sub>2.5</sub> , BC, NO <sub>2</sub> , O <sub>3</sub> , and the copper, iron, zinc, and sulfur content of PM <sub>2.5</sub>
Monitoring data	ESCAPE network	AirBase (the European air quality database) and ESCAPE network	AirBase (the European air quality database) and ESCAPE network
Exposure models	Within-city, within-area, and within-country	Europewide	Europewide
Total number of cohorts	30	15 <sup>a</sup>	7
Total population	~370,000	~325,000	~28 million
Health endpoints <sup>b</sup>	Pregnancy and birth outcomes Respiratory disease outcomes Cardiovascular disease outcomes Cancer incidence and cause-specific mortality	Natural and cause-specific mortality including cardiovascular, non-malignant respiratory and diabetes mortality Coronary and cerebrovascular events Lung cancer incidence Asthma and COPD incidence	Natural and cause-specific mortality including cardiovascular, nonmalignant and malignant respiratory and diabetes mortality
Typical covariates included (but differed according to cohort and outcome considered)	Information on age, sex, individual and area-level socio-economic status, and health behaviors: smoking status, body mass index	Information on age, sex, individual and area-level socio-economic status, and health behaviors: smoking status, body mass index	Information on age, sex, individual and area-level socio-economic status; the English cohort had health behaviors: smoking status, body mass index

<sup>a</sup> The investigators selected cohorts from the original ESCAPE study that carried the most weight for analyses of the mortality and morbidity endpoints for the current study; were willing to pool data; were at the lower end of the ESCAPE exposure range, and were recruited relatively recently. See Investigators' Report (IR), page 11.

<sup>b</sup> Not all health endpoints were considered in each cohort category; see Investigators' Report Table 1 for a detailed overview.

ships between emissions and concentrations as compared to the empirical nature of LUR models.

Here, the investigators used AirBase routine monitoring data to derive annual average concentrations for PM<sub>2.5</sub>, NO<sub>2</sub>, and O<sub>3</sub>. AirBase is the European air quality database maintained by the European Environment Agency. It contains air quality monitoring data and information submitted by participating countries throughout Europe. The database consists of a multi-annual time series of air quality measurement data and statistics for several air pollutants. The investigators developed their LUR model for BC with monitoring data collected from the network of monitors deployed as part of the ESCAPE project. Spatial data describing roads, land use, elevation, and population density, as well as satellite data, were all used as predictor variables in the LURs. The investigators also made use of pollutant estimates from two long-range chemical transport models, namely the MACC-II ENSEMBLE (Inness et al. 2013) and the Danish Eulerian Hemispheric Model (Brandt et al. 2012). They also applied universal kriging, which is an interpolation technique that considers the distance and direction between sample points (i.e., pollution monitors), to aid in explaining spatial variation in estimated concentrations.

The models were developed with monitoring data from the year 2010 because that was the earliest year with sufficient PM<sub>2.5</sub> data across Europe. For consistency, models for the other pollutants were based on data from this same year. All models were created at a spatial resolution of 100 m × 100 m. Brunekreef and colleagues created exposure models for annual, warm-season, and cold-season O<sub>3</sub>, but correlations were high between the other pollutants and annual and winter season O<sub>3</sub> estimates, and so they focused on warm-season O<sub>3</sub> throughout their study. Other reasons for this were that concentrations and concentration contrasts are higher in summer and people spend more time outdoors in the warm season.

They examined the validity of the final exposure models by comparing indicators of performance and model fit among additional models produced while randomly excluding observations from 20% of the monitoring sites, with subsets of observations at relatively low concentrations, and validating with external ESCAPE monitoring data for PM<sub>2.5</sub> and NO<sub>2</sub>. They also evaluated the performance of 14 algorithms including supervised linear regression to develop LUR models for PM<sub>2.5</sub> and NO<sub>2</sub>. They observed that the performance of most algorithms was similar, with little indication of better performance of more sophisticated algorithms compared with supervised linear regression.

**Additional Models to Examine Spatial Stability Over Time** The investigators developed additional NO<sub>2</sub> and O<sub>3</sub> models for 2000 and 2005, as well as a PM<sub>2.5</sub> model for 2013 to examine spatial stability over time.

**Additional Models to Account for Temporal Changes**

The investigators used estimates from the Danish Eulerian Hemispheric Model, which is an atmospheric chemical transport model developed to study the long-range transport of air pollution across the Northern Hemisphere, to extrapolate pollutant concentrations from 2010 back to 1990 (preceding the earliest baseline year for any of the study cohorts) and forward to 2017 (the latest end of follow-up year for any of the cohorts).

**Additional Area-Specific Models** The investigators made use of previously developed local-, region-, or country-specific pollution models, where available (IR Table 2; Beelen et al. 2013; Eeftens et al. 2012) for comparing results and patterns with the new Europe-wide models.

**MAPLE PM<sub>2.5</sub> Exposure Model** Additionally, the investigators used 2010 PM<sub>2.5</sub> estimates at a 1 × 1 km spatial resolution produced through the MAPLE study (see Preface or Brauer et al. 2019). The MAPLE model is based primarily on satellite aerosol optical depth and uses a global chemical transport model (GEOS-Chem) to calculate surface concentrations. Brauer and colleagues used geographically weighted regression to calibrate the surface concentrations to ground monitoring data (van Donkelaar et al. 2015, 2016). The MAPLE study has further refined the exposure estimation method by incorporating long-term measurements of aerosol optical depth from the ground at selected Canadian stations. Although the ELAPSE model includes satellite PM<sub>2.5</sub> data, it does not include the refinements added by geographically weighted regression and ground-based aerosol optical depth monitoring.

**Study Populations**

This study combines 15 well-characterized cohorts, mostly from the ESCAPE project, pooled together and seven large European administrative cohorts analyzed separately to perform new analyses on the health effects of air pollution at low levels of exposure. Commentary Table 2 presents selected characteristics of all cohorts used in these analyses. An administrative cohort is created through linked “administrative” data (e.g., census records and national mortality databases). Unlike conventional “research” cohorts, where individuals are invited to participate and to respond to questionnaires, the data used to compile administrative cohorts were not originally collected for research purposes nor for linking together.

**Commentary Table 2.** Selected Characteristics of the ELAPSE Cohorts

Cohort	Country	Geographic Coverage	Follow-Up Period	Sample Size
<b>Pooled Cohorts</b>				
CEANS-SDPP	Sweden	Stockholm county	1992–1998 to 2011	7,835
CEANS-SIXTY	Sweden	Stockholm county	1997–1999 to 2014	4,180
CEANS-SALT	Sweden	Stockholm county	1998–2002 to 2011	6,724
CEANS-SNACK	Sweden	Stockholm county	2001–2004 to 2011	3,248
DCH	Denmark	Copenhagen and Aarhus	1993–1997 to 2015	56,308
DNC-1993	Denmark	National	1993 to 2013	19,664
DNC-1999	Denmark	National	1999 to 2013	8,769
E3N	France	National	1993–1996 to 2011	53,521
EPIC-NL-MORGEN	The Netherlands	Four cities	1993–1997 to 2013	20,711
EPIC-NL-PROSPECT	The Netherlands	Four cities	1993–1997 to 2013	16,194
EPIC-VARESE	Italy	City of Varese	1993–1997 to 2014	12,028
HNR	Germany	Ruhr area	2000–2003 to 2015	4,809
KORA-S3	Germany	Augsburg area	1994–1995 to 2011	4,566
KORA-S4	Germany	Augsburg area	1999–2001 to 2014	4,257
VHM&PP	Austria	Vorarlberg Region	1985–2005 to 2014	170,250
<b>Administrative Cohorts</b>				
Belgian	Belgium (BE)	National	2001 to 2011	6,491,801
Danish	Denmark (DK)	National	2000 to 2015	3,409,517
Dutch	The Netherlands (NL)	National	2008 to 2012	10,532,360
English	England (EN)	National	2011 to 2017	1,491,124
Norwegian	Norway (NO)	National	2001 to 2016	2,516,192
Italian	Italy (IT)	City of Rome	2001 to 2015	1,263,712
Swiss	Switzerland (CH)	National	2000 to 2014	4,293,521

The advantage of working with administrative cohorts is that they tend to cover an entire country and are therefore nationally representative. Although the seven administrative cohorts were analyzed separately in this study, meta-analysis was used to produce summary effect estimates. Participants in the 15 conventional cohorts were combined into a single “pooled” dataset for analysis purposes. The investigators also evaluated the consistency of epidemiological results from these two cohort groups.

The purpose of reporting findings from separate analyses based on the two cohort groups was to address the various strengths and limitations of using each individually. For example, although the administrative cohorts are representative of whole populations, have substantially greater statistical power (i.e., approximately 28 million participants compared with approximately 325,000 participants in the pooled cohort), and control effectively for contextual confounders (i.e., area-based neighborhood- or community-level variables), they include relatively few

individual-level confounders. Conversely, the pooled cohorts have detailed information on individual-level characteristics and health behaviors, in addition to contextual confounders, but have relatively smaller sample sizes and may be less representative in their coverage of the full populations (e.g., some of these cohorts were composed exclusively of women).

**Pooled Cohorts** Brunekreef and colleagues chose to pool 15 cohorts, 14 of which were analyzed previously as part of the ESCAPE project, in an effort to gain statistical power for epidemiological analyses and to investigate more efficiently the shapes of concentration–response functions. The specific cohorts used for this study were selected based on willingness to pool data, had populations with exposures to relatively low concentrations of pollution, and were relatively recent in their recruitment. The Danish Nurse Cohort joined the ELAPSE study separately with independent funding. Most of these cohorts are located in a region that included one or multiple large cities and surrounding smaller towns.

**Administrative Cohorts** The administrative cohorts comprised 28 million participants spread across seven large administrative cohorts in seven European countries. These cohorts were formed by linking census data, population registries, and death registries. The administrative cohorts contributed primarily to the mortality analyses. In some previous studies with these cohorts, different exposure assessment and analytical methods were applied, thus making comparisons among results challenging. This work harmonized these cohorts to enhance comparability.

### Exposure Assignment

As noted above, the investigators produced  $100\text{ m} \times 100\text{ m}$  spatial surfaces (i.e., maps) of estimates of annual average concentrations for each pollutant. They used exposure estimates for the year 2010 assigned to participant address at the time of recruitment as the main exposure variable in epidemiological models. Residential history data were available for all administrative cohort participants and for about half of the pooled cohort participants. As such, they assigned exposures for each year of follow-up using the annual, temporally adjusted estimates to all available addresses for use in separate sensitivity analyses, as noted below.

### Main Epidemiological Analyses

Health outcomes in this study included natural cause, cardiovascular, diabetes, cardiometabolic and respiratory mortality, as well as incidences of lung cancer, coronary and cerebrovascular events, asthma, and COPD. The

investigators applied standard Cox proportional hazard models to describe associations between exposures to the pollutants and mortality and morbidity. Main analyses were limited to each pollutant individually. The geographic scale of the contextual variables ranged from neighborhood in some cases to that of the municipality in others. Model specification differed slightly between those conducted with the pooled cohort versus those conducted with the administrative cohorts and according to health outcome considered, as noted below. Broadly, however, the investigators explored increasing levels of confounder control with individual-level and contextual covariates for both the pooled cohort and the administrative cohorts, namely:

- **Model 1:** included age (time axis), sex (as baseline hazard stratification), calendar year of enrollment, and cohort (as baseline hazard stratification) in the pooled cohort analyses;
- **Model 2:** added individual-level variables available to all cohorts in the pooled cohort, and all variables available in each administrative cohort;
- **Model 3:** added contextual variables (as covariates, not as random effects).

Ultimately, the investigators selected the most fully adjusted models (i.e., Model 3) available for use in the main analyses and limited these to only those participants with complete covariate information. In the case of the pooled cohort, the final model was defined by balancing the need to adjust for specific confounders and the desire to include as many cohorts as possible. All cohorts in the pooled cohort included information on age, sex, smoking status, body mass index, and individual-level and contextual indicators of socioeconomic status. In the case of the administrative cohorts, which were analyzed individually and then meta-analyzed using random effects models, the investigators preferred models with maximal adjustment per cohort rather than a common model with fewer covariates. As such, they included all available individual-level and contextual covariates in these models. The administrative cohorts, with the exception of the English cohort, did not include information on smoking or body mass index, which are both important risk factors for mortality. To address this limitation, the investigators applied indirect adjustment methods to adjust the hazard ratios of natural-cause mortality for these missing covariates. Indirect adjustment is a technique to help overcome the issue of unmeasured confounding in epidemiological studies where important risk factor information is missing (e.g., information on smoking or diet). The method allows researchers to adjust hazard ratios by examining the relationship between the missing risk factors and exposure in

an ancillary data set that does include information on those risk factors (along with all other variables also included in the survival models) (Erickson et al. 2019).

### Additional Epidemiological Analyses

The investigators conducted many additional analyses and sensitivity analyses to assess the robustness of the findings reported with the main models, including analyses limited to subsets of participants within selected ranges of exposure concentrations to investigate associations at low levels. Below is a comprehensive enumeration of additional analyses conducted, according to two categories of analyses, namely: **alternative approaches to exposure estimation** and **variations on covariate adjustment**.

Lastly, the investigators also examined the shapes of **concentration–response functions** to investigate associations at low levels. Note that not all of these additional analyses were conducted for all health outcomes, nor necessarily for both cohort groups.

**Alternative Approaches to Exposure Estimation** In addition to using the exposure estimates derived for the year 2010 from the central, Europewide models applied to baseline years, the investigators also explored models using:

- time-varying exposures (i.e., a 1-year moving window of exposure);
- local exposure models (where available);
- estimates for PM<sub>2.5</sub> created for the MAPLE study;
- estimates for O<sub>3</sub> at a larger spatial scale; and,
- estimates for PM<sub>2.5</sub> composition, specifically the copper, iron, zinc, and sulfur content of fine particulate matter. These analyses were not planned originally, and results for the compositional models were included only in Appendix 4. As such, they are not discussed further in this commentary.

**Variations on Covariate Adjustments** In addition to the three levels of model adjustment described above, the investigators also examined models that:

- evaluated effect modification by age at baseline, smoking status, and body mass index;
- were adjusted for every two-pollutant combination of PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub>;
- adjusted for dietary variables (i.e., alcohol and fruit intake);
- adjusted for road traffic noise; and,

- adjusted for indicator variables for “region” to allow for variation in health outcomes not accounted for by the other confounder variables in the models.

**Concentration–Response Functions** The team investigated the shapes of concentration–response functions using natural splines with two, three, and four degrees of freedom, with penalized splines, and with shape-constrained health impact functions (SCHIFs). SCHIFs were developed by Nasari and colleagues (2016) to extend the log-linear model (which relates the logarithm of the hazard ratio to exposure in a linear manner) to nonlinear transformations of exposure (including near-linear, supralinear, and sublinear). The Brunekreef team used the SCHIF method in addition to the more traditional methods to estimate concentration–response curves to harmonize the analytic approaches across the other studies funded under this RFA and conducted in Europe and North America, in particular with the Canadian (MAPLE) team. The Canadian team asserted that a major advantage of the SCHIF over nonparametric smoothing functions is the resulting specific parameter estimates that can be applied in related analyses, for example in cost–benefit evaluations. The SCHIFs are constrained to produce functions that increase monotonically with concentration and forms that are biologically plausible, for example, not allowing multiple upward and downward inflections (Brauer et al. 2019).

---

## SUMMARY OF KEY RESULTS

---

### MODELING AND EXPOSURE ESTIMATION RESULTS

The final, Europewide hybrid LUR exposure models explained 66%, 51%, 58%, and 60% of the variability in concentrations of PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub>, respectively. As noted above, the investigators developed additional NO<sub>2</sub> and O<sub>3</sub> models for the years 2000 and 2005, as well as a PM<sub>2.5</sub> model for 2013 to examine spatial stability over time. Here, squared correlations ( $R^2$ ) between concentrations predicted at random sites across the models for different years overall were generally high (i.e., >78%). The within-country squared correlations were more modest: >49% for PM<sub>2.5</sub>, >80% for NO<sub>2</sub>, and >47% for O<sub>3</sub> (except Italy, i.e., 12%). Additionally, the investigators observed similar patterns for correlations between the Airbase monitoring data in different years. All of this suggested that the main exposure models had good spatial stability (across the full study area) and good temporal stability (throughout the study period).

Key epidemiological results are presented below separately for analyses with the pooled cohort and with the

administrative cohorts. Not all analyses were conducted on both cohort groups.

**POOLED COHORT HEALTH ANALYSIS**

In 2010, almost all participants in the pooled cohort had PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below the EU limit values of 25 and 40 µg/m<sup>3</sup>, respectively, with more than 50,000 and 25,000 participants experiencing residential PM<sub>2.5</sub> exposures below the U.S. EPA NAAQS (12 µg/m<sup>3</sup>) and the WHO guideline (10 µg/m<sup>3</sup>). Mean concentrations (and standard deviations [SD]) in µg/m<sup>3</sup> (BC was measured in 1.5 × 10<sup>-5</sup>/m) for participants in the pooled cohort for PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub> were 15.02 (3.22), 1.52 (0.42), 25.00 (8.05), and 67.46 (6.86), respectively.

The investigators reported significant positive associations between PM<sub>2.5</sub>, BC, and NO<sub>2</sub>, and all causes of death examined in their main models (with the exception of respiratory mortality and PM<sub>2.5</sub>) (Commentary Table 3; See also

IR Table 5). They reported the largest mortality effect estimates for diabetes mortality (e.g., hazard ratio (HR) for PM<sub>2.5</sub> per 5 µg/m<sup>3</sup>: 1.32; 95% confidence interval (CI): 1.14–1.51). They reported inverse associations between O<sub>3</sub> and all causes of death examined.

The HRs for natural-cause mortality remained elevated and significant for PM<sub>2.5</sub> even when all observations higher than 12 µg/m<sup>3</sup> were removed from the analysis (Commentary Table 4 and IR Table 7), although effect estimates tended to be higher than those reported with the full dataset. For NO<sub>2</sub>, HRs remained elevated and significant even when all observations higher than 20 µg/m<sup>3</sup> were removed, and for BC, HRs remain elevated and significant even when all observations higher than 1.5 × 10<sup>-5</sup>/m were removed. HRs for O<sub>3</sub> attenuated toward unity at concentrations below 80 µg/m<sup>3</sup> (IR Table 7).

In the incidence analyses the investigators found significant positive associations between PM<sub>2.5</sub>, BC, and NO<sub>2</sub>

**Commentary Table 3.** Associations Between Air Pollution and Selected Causes of Mortality from Main Epidemiological Models Among Participants in the Pooled and Administrative ELAPSE Cohorts<sup>a</sup>

Cohort (deaths)	Natural Cause		Cardiovascular		Respiratory		Diabetes	
	Pooled <sup>b</sup> (47,131)	Administrative <sup>c</sup> Meta-analysis (3,593,741)	Pooled <sup>b</sup> (15,542)	Administrative <sup>c</sup> Meta-analysis (1,186,101)	Pooled <sup>b</sup> (2,865)	Administrative <sup>c</sup> Meta-analysis (371,990)	Pooled <sup>b</sup> (1,034)	Administrative <sup>c</sup> Meta-analysis (78,622)
<b>PM<sub>2.5</sub></b>	1.130 (1.106, 1.155)	1.053 (1.021, 1.085)	1.135 (1.095, 1.176)	1.041 (1.010, 1.072)	1.054 (0.961, 1.156)	1.064 (1.013, 1.118)	1.316 (1.144, 1.514)	1.038 (0.974, 1.106)
<b>BC</b>	1.081 (1.065, 1.098)	1.039 (1.018, 1.059)	1.085 (1.055, 1.116)	1.022 (1.004, 1.040)	1.084 (1.020, 1.151)	1.053 (1.021, 1.085)	1.240 (1.112, 1.382)	1.015 (0.969, 1.065)
<b>NO<sub>2</sub></b>	1.086 (1.070, 1.102)	1.044 (1.019, 1.069)	1.089 (1.060, 1.120)	1.025 (1.006, 1.044)	1.101 (1.038, 1.168)	1.058 (1.024, 1.093)	1.238 (1.112, 1.378)	1.013 (0.958, 1.070)
<b>O<sub>3</sub></b>	0.896 (0.878, 0.914)	0.953 (0.929, 0.979)	0.887 (0.854, 0.922)	0.976 (0.954, 0.998)	0.890 (0.821, 0.966)	0.948 (0.910, 0.988)	0.744 (0.645, 0.859)	0.984 (0.917, 1.057)

<sup>a</sup> Results are expressed as hazard ratios and 95% confidence intervals. (Source IR: Tables 5 and 19.) HRs presented for the following increments: PM<sub>2.5</sub>: 5 µg/m<sup>3</sup>; BC: 0.5 × 10<sup>-5</sup>/m; NO<sub>2</sub>: 10 µg/m<sup>3</sup>; and O<sub>3</sub>: 10 µg/m<sup>3</sup>.

<sup>b</sup> For the pooled cohort—Model 3: adjusted for cohort id, age, sex, year of baseline visit, smoking (status, duration, intensity, intensity squared), BMI, marital status, employment status, and 2001 neighborhood-level mean income.

<sup>c</sup> For the administrative cohorts—Model 3: adjusted for age, sex, year of baseline visit, and cohort-specific individual and area-level SES variables.

**Commentary Table 4.** Associations Between Air Pollution and Natural-Cause Mortality in Subset Analyses Among Participants in the Pooled Cohort

Pollutant / Subset	N	HR <sup>a</sup> (95% CI)
<b>PM<sub>2.5</sub></b>		
Full dataset	325,367	1.130 (1.106, 1.155)
<12 µg/m <sup>3</sup>	52,528	1.296 (1.140, 1.474)
<10 µg/m <sup>3</sup>	25,422	1.146 (0.931, 1.410)
<b>NO<sub>2</sub></b>		
Full dataset	325,367	1.086 (1.070, 1.102)
<30 µg/m <sup>3</sup>	247,039	1.114 (1.088, 1.140)
<20 µg/m <sup>3</sup>	88,510	1.099 (1.033, 1.170)
<b>BC</b>		
Full dataset	325,367	1.081 (1.065, 1.098)
<1.5 × 10 <sup>-5</sup> /m	142,032	1.125 (1.086, 1.165)
<1 × 10 <sup>-5</sup> /m	35,406	1.041 (0.942, 1.150)

<sup>a</sup> HR (95% confidence interval) presented for the following increments: PM<sub>2.5</sub>: 5 µg/m<sup>3</sup>, NO<sub>2</sub>: 10 µg/m<sup>3</sup>, BC: 0.5 × 10<sup>-5</sup>/m; Model 3 adjusted for cohort id, age, sex, year of baseline visit, smoking (status, duration, intensity, intensity squared), BMI, marital status, employment status, and 2001 neighborhood-level mean income.

and incidence of stroke, asthma, and COPD hospital admissions. Additionally, they reported significant associations between NO<sub>2</sub> and acute coronary heart disease and between PM<sub>2.5</sub> and lung cancer incidence (Commentary Table 5). They reported that results with the full pooled cohort were similar in magnitude in subset analyses limited to participants with exposures below 10 µg/m<sup>3</sup> for PM<sub>2.5</sub> and below 20 µg/m<sup>3</sup> for NO<sub>2</sub>.

In two-pollutant models, the HRs for natural-cause mortality were attenuated, but remained elevated and statistically significant for PM<sub>2.5</sub> and NO<sub>2</sub>. Associations with O<sub>3</sub> were also attenuated but remained negative in the two-pollutant models with PM<sub>2.5</sub>, BC, and NO<sub>2</sub>. It should be noted that BC and NO<sub>2</sub> were highly correlated in all cohorts, PM<sub>2.5</sub> was moderately to highly correlated with BC and NO<sub>2</sub>, and O<sub>3</sub> was negatively correlated with PM<sub>2.5</sub>, NO<sub>2</sub>, and BC.

In spline plots examining the shape of associations between exposure and natural-cause mortality, the investigators observed generally supralinear patterns (i.e., steeper slopes at lower exposures) with no evidence of concentrations below which no associations were found for PM<sub>2.5</sub>, BC, and NO<sub>2</sub>. That is, increased risks for mortality were associated with even the lowest observed concentrations. Commentary Figure 1 shows the results for PM<sub>2.5</sub>; the curves for BC and NO<sub>2</sub> had similar shapes, though were somewhat more u-shaped. Lastly, the investigators pre-

**Commentary Table 5.** Associations Between Air Pollution and Incidence of Acute Coronary Heart Disease, Stroke, Lung Cancer, Asthma, and COPD Among Participants in the Pooled Cohort<sup>a,b</sup>

	Acute Coronary Heart Disease <sup>c</sup>	Stroke <sup>c</sup>	Lung Cancer <sup>d</sup>	Asthma <sup>e</sup>	COPD <sup>f</sup>
<b>PM<sub>2.5</sub></b>	1.02 (0.95, 1.10)	1.10 (1.01, 1.21)	1.13 (1.05, 1.23)	1.22 (1.04, 1.43)	1.17 (1.06, 1.29)
<b>BC</b>	1.02 (0.99, 1.06)	1.06 (1.02, 1.10)	1.02 (0.97, 1.07)	1.15 (1.08, 1.23)	1.11 (1.06, 1.15)
<b>NO<sub>2</sub></b>	1.04 (1.01, 1.07)	1.08 (1.04, 1.12)	1.02 (0.97, 1.07)	1.17 (1.10, 1.25)	1.11 (1.06, 1.16)
<b>O<sub>3</sub></b>	0.94 (0.90, 0.98)	0.96 (0.91, 1.01)	0.95 (0.89, 1.02)	0.90 (0.81, 0.99)	0.99 (0.93, 1.05)

<sup>a</sup> Results are expressed as hazard ratios and 95% confidence intervals. PM<sub>2.5</sub> per 5 µg/m<sup>3</sup>, NO<sub>2</sub> per 10 µg/m<sup>3</sup>, BC per 0.5 × 10<sup>-5</sup>/m, O<sub>3</sub> per 10 µg/m<sup>3</sup>.

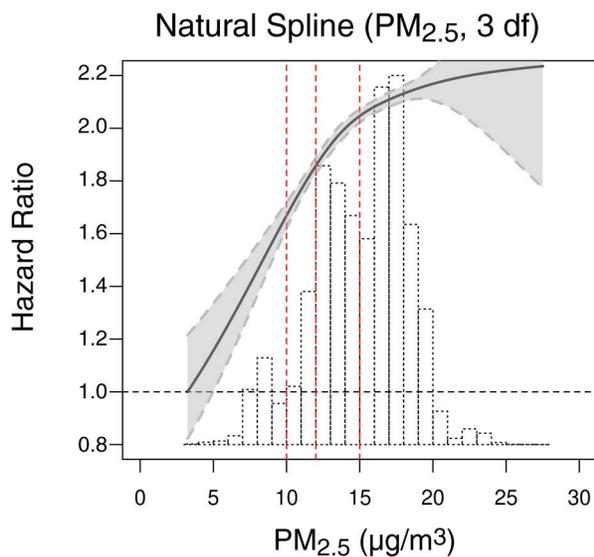
<sup>b</sup> Adjusted for cohort id, age, sex, year of baseline visit, smoking (status, duration, intensity, intensity squared), BMI category, marital status, employment status, and 2001 neighborhood-level mean income (acute coronary heart disease, stroke, asthma, and COPD incidence additionally adjusted for educational status). (Source: IR Table 13.)

<sup>c</sup> Acute coronary heart disease and stroke, *n* = 137,148.

<sup>d</sup> Lung cancer, *n* = 307,550.

<sup>e</sup> Asthma, *n* = 98,326.

<sup>f</sup> COPD, *n* = 98,508.



**Commentary Figure 1. Associations between PM<sub>2.5</sub> and natural-cause mortality among participants in the pooled cohort.** Concentration–response curve shown as a natural cubic spline with three degrees of freedom (df). Shaded area represents 95% confidence interval. Red dotted lines are, from right to left, the “old” U.S. EPA NAAQS, the current U.S. EPA NAAQS, and the 2005 WHO air quality guideline values. X-axis truncated at 30 µg/m<sup>3</sup>. Hazard ratios are expressed relative to minimum exposure (i.e., 3.24 µg/m<sup>3</sup>). Histograms represent exposure distributions. WHO = World Health Organization; NAAQS = U.S. National Ambient Air Quality Standards. (Source: IR Figure 4.)

sented associations between mortality and exposures to PM<sub>2.5</sub> estimated by models from the MAPLE study. The exposure estimates from the MAPLE model were slightly lower than those from the main ELAPSE model, but the two were highly correlated ( $r = 0.7$ ), suggesting similar spatial patterns. Generally, the effect estimates were slightly larger in models that used the ELAPSE exposure model, but the 95% CIs overlapped with those that used the MAPLE exposure model in all cases.

#### ADMINISTRATIVE COHORT HEALTH ANALYSIS

In 2010, almost all participants had PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below the EU limit values, and more than 3.9 million and 1.9 million, respectively, experienced residential PM<sub>2.5</sub> exposures below the U.S. EPA NAAQS and the WHO guideline. Mean concentrations of PM<sub>2.5</sub> ranged from ~12–19 µg/m<sup>3</sup> across all but the Norwegian cohort (8.26 µg/m<sup>3</sup>) (IR Appendix 3, Table A1).

The investigators reported significant positive associations between PM<sub>2.5</sub>, BC, and NO<sub>2</sub> and natural-cause, cardiovascular, respiratory, and lung cancer mortality with moderate to high heterogeneity between cohorts (Commentary Table 3; note results for lung cancer not presented here). Here they found weaker evidence of associations with diabetes and cerebrovascular mortality than were reported with the pooled cohort. Similar to the analyses with the pooled cohort, they reported inverse associations between O<sub>3</sub> and all causes of death examined. Also similar to the results with the pooled cohort, effect estimates for mortality were somewhat larger than those reported with the full datasets when analyses were limited to participants in the lowest concentration ranges for each of PM<sub>2.5</sub>, BC, and NO<sub>2</sub>.

HRs for natural-cause mortality, however, remained significantly associated with BC and NO<sub>2</sub> in two-pollutant models adjusted additionally for either PM<sub>2.5</sub> or O<sub>3</sub>. The HRs for PM<sub>2.5</sub> and for BC were attenuated to unity (i.e., HR = 1.0 indicating no difference) in two-pollutant models with NO<sub>2</sub>. In two-pollutant models for O<sub>3</sub>, associations were attenuated to unity and not statistically significant.

The shape of association between exposure and natural-cause mortality differed among the seven administrative cohorts, although associations were generally linear to supralinear, with effects observed at all pollution concentrations considered.

As with the pooled cohort, the investigators reported comparable associations between mortality and exposures to PM<sub>2.5</sub> produced through the MAPLE study among the administrative cohorts. For example, the meta-analytic HR for natural-cause mortality with ELAPSE exposures was 1.058 (95% CI: 1.022–1.095), compared with HR 1.047 (95% CI: 1.003–1.094) with MAPLE exposures (IR Table 26). The fact that the associations observed with the independently developed exposure model were similar to those observed with the ELAPSE-developed exposures corroborated the robustness of the associations.

---

#### EVALUATION BY THE HEI LOW-EXPOSURE EPIDEMIOLOGY STUDIES REVIEW PANEL

---

In its independent review of the study, the HEI Low-Exposure Epidemiology Studies Review Panel concluded that this was an impressive, well-designed, comprehensively analyzed, and properly interpreted study. The Panel found that having produced exposure models at 100 m × 100 m spatial resolution for all pollutants across Europe was an impressive achievement and that the harmonized exposure models were a strong improvement over the region-specific models used in earlier studies. The Panel

felt that a particular strength of the study was the investigation of so many important health endpoints. The Panel also appreciated the extensive and thorough set of sensitivity analyses examining differences in effect sizes for associations between long-term exposures to several key air pollutants and several health outcomes, exploring different levels of confounder adjustment, exploring different exposure specifications, and exploring different approaches to modeling concentration–response functions.

The findings from this Report contribute to our knowledge of effects on health associated with long-term exposures to relatively lower concentrations of ambient air pollution. As noted previously, almost all participants in both the pooled and administrative cohorts had PM<sub>2.5</sub> and NO<sub>2</sub> annual average exposures below EU limit values. The investigators reported positive associations between exposures to PM<sub>2.5</sub>, BC, and NO<sub>2</sub> and natural-cause, respiratory, and cardiovascular mortality in both cohorts (Commentary Table 3). The shapes of the concentration–response functions between exposures and mortality differed somewhat between cohorts, although associations were generally linear to supralinear with no evidence of thresholds below which no effects were found. These and other aspects of the study design and approach as well as interpretations of the findings and results, are described and discussed in the following sections.

## EVALUATION OF STUDY DESIGN AND APPROACH

### Air Pollution Models and Exposure Estimation

The Panel found that the development of the pollution models used for exposure estimation was a major achievement of this study. Brunekreef and colleagues created the first-ever, Europe-wide, hybrid LUR models for each of PM<sub>2.5</sub>, BC, NO<sub>2</sub>, and O<sub>3</sub> (annual, and warm and cold seasons). The Panel concluded that the investigators' application of a single, harmonized model for each pollutant was an improvement over the approaches applied in earlier analyses, such as those conducted as part of the ESCAPE project, in which exposure models were developed for individual countries and regions in Europe. Importantly, each of the new models was developed at a fine spatial resolution of 100 m × 100 m. The Panel was very satisfied with this level of spatial detail in models for PM<sub>2.5</sub>. PM<sub>2.5</sub> has generally lower spatial variability than BC, NO<sub>2</sub>, and O<sub>3</sub>, which are influenced more by traffic and other local sources. Models for BC, NO<sub>2</sub>, and O<sub>3</sub> were considered acceptable at a 100-m resolution, but the panel felt that that resolution may not have fully captured the finer-scale spatial patterns of these pollutants.

The investigators examined the validity of the exposure models by comparing indicators of model performance

(e.g.,  $R^2$  between observed and predicted values) among subsets of held-out observations at lower concentrations. In all cases of these subsets, the  $R^2$  values were lower than those produced with the full set of observations, suggesting that generally, the exposure models were less effective at estimating concentrations in the lowest pollution areas. There were, however, challenges to evaluating the performance of the models at the lowest concentrations. For example, only 86 of the 543 pollution monitoring sites that measured PM<sub>2.5</sub> were located in areas with concentrations lower than 10 µg/m<sup>3</sup>, and only 841 of 2,399 sites that measured NO<sub>2</sub> were in areas with concentrations lower than 20 µg/m<sup>3</sup>. That is, there were relatively few observations to evaluate and validate model performance for areas with concentrations below these levels. As noted above, the investigators also evaluated the performance of 14 different algorithms to develop LUR models for PM<sub>2.5</sub> and NO<sub>2</sub>. Here, they observed that the performance of most algorithms was similar, with little indication of better performance of more sophisticated algorithms compared with supervised linear regression.

As described above, using exposure data from 2010, at time of recruitment, nearly all participants in the various cohorts were assigned estimates of exposure to annual PM<sub>2.5</sub> in the range of ~12–19 µg/m<sup>3</sup>. Importantly, only participants in cohorts from two countries, namely Sweden (in the pooled cohort) and Norway (among the administrative cohorts) had annual mean concentrations lower than 10 µg/m<sup>3</sup> based on the year 2010. On the one hand, these exposures are all somewhat higher than those reported in a recent U.S.-based cohort study of ~61 million older adults for whom annual average PM<sub>2.5</sub> concentrations for the years 2000–2012 ranged from 6.2 to 15.6 µg/m<sup>3</sup> (5th and 95th percentiles, respectively) (Di et al. 2017). Additionally, a recent national Canadian cohort study (Brauer et al. 2019) reported even lower annual mean exposures of 6.7 µg/m<sup>3</sup> among ~3 million adults for the year 2001. On the other hand, from a global perspective, these values are indeed relatively low: in 2019, over 90% of the world's population was exposed to annual average concentrations of PM<sub>2.5</sub> greater than 10 µg/m<sup>3</sup>, with most exposed to concentrations greater than 20 µg/m<sup>3</sup> (Health Effects Institute 2020). Moreover, many African and Asian countries experience annual mean exposures to PM<sub>2.5</sub> greater than 45 µg/m<sup>3</sup>. That is, “low” concentrations of pollution is a relative term, and the analyses presented in this Report should be considered within that global context.

The investigators reported, in sensitivity analyses with time-varying exposures that incorporated residential histories, associations and effect sizes comparable to those reported in the main models (based on exposure data from 2010 assigned at baseline). The Panel cautioned, however,

that there are important limitations to the approach of back-extrapolating exposure estimates as far back as 1990. Essentially, it should be acknowledged that levels of uncertainty in these estimates increase the further they go back in time. This uncertainty is related in part to the fact that relatively few monitoring stations were in operation during the earlier periods and because most of the model input variables that were fixed in 2010 would likely change in nonlinear, nonspatially uniform ways over that 20-year period (e.g., construction of new roads, changes in urban form and housing density). Additionally, the investigators documented that concentrations of both PM<sub>2.5</sub> and NO<sub>2</sub> have been decreasing over time, but it is not clear the extent to which these decreases have been linear, nor if they have been consistent across regions. Moreover, reductions in pollutant concentrations over time could impact the magnitude of associations reported if the variation in concentrations also shrank (i.e., if the range between high and low exposures shrank), even if correlations between periods were strong. All these details have implications for analyses based on back-extrapolated exposure estimates here and elsewhere.

### Evaluation of Epidemiological Analysis

A key epidemiological contribution of this study is the fact that the investigators reported on associations between exposures and many important health endpoints. Specifically, the investigators considered associations with mortality from all-natural causes, from cardiometabolic diseases (i.e., all cardiovascular diseases and diabetes combined), all cardiovascular diseases, ischemic heart disease, cerebrovascular disease, diabetes, respiratory disease, and from COPD (excluding asthma). Additionally, they examined incidence of lung cancer, acute coronary heart disease, and cerebrovascular, COPD, and asthma events. It is a major achievement for the investigators to have analyzed associations of so many different health endpoints with four different pollutants, being very vigilant about checking for the robustness of their findings to different modeling choices.

In addition to natural-cause mortality summary effect estimates for the full pooled and administrative cohorts, Brunekreef and colleagues presented results for seven of the administrative cohorts individually. Providing these regional results was a valuable contribution that allowed for some regional comparisons between exposure patterns and risk estimates. For example, they showed that annual mean concentrations of PM<sub>2.5</sub> ranged from ~17–19 µg/m<sup>3</sup> in the Roman and Belgian cohorts, to ~12–13 µg/m<sup>3</sup> in the Danish and English cohorts, to ~8 µg/m<sup>3</sup> in the Norwegian cohort. Moreover, they presented the heterogeneity in

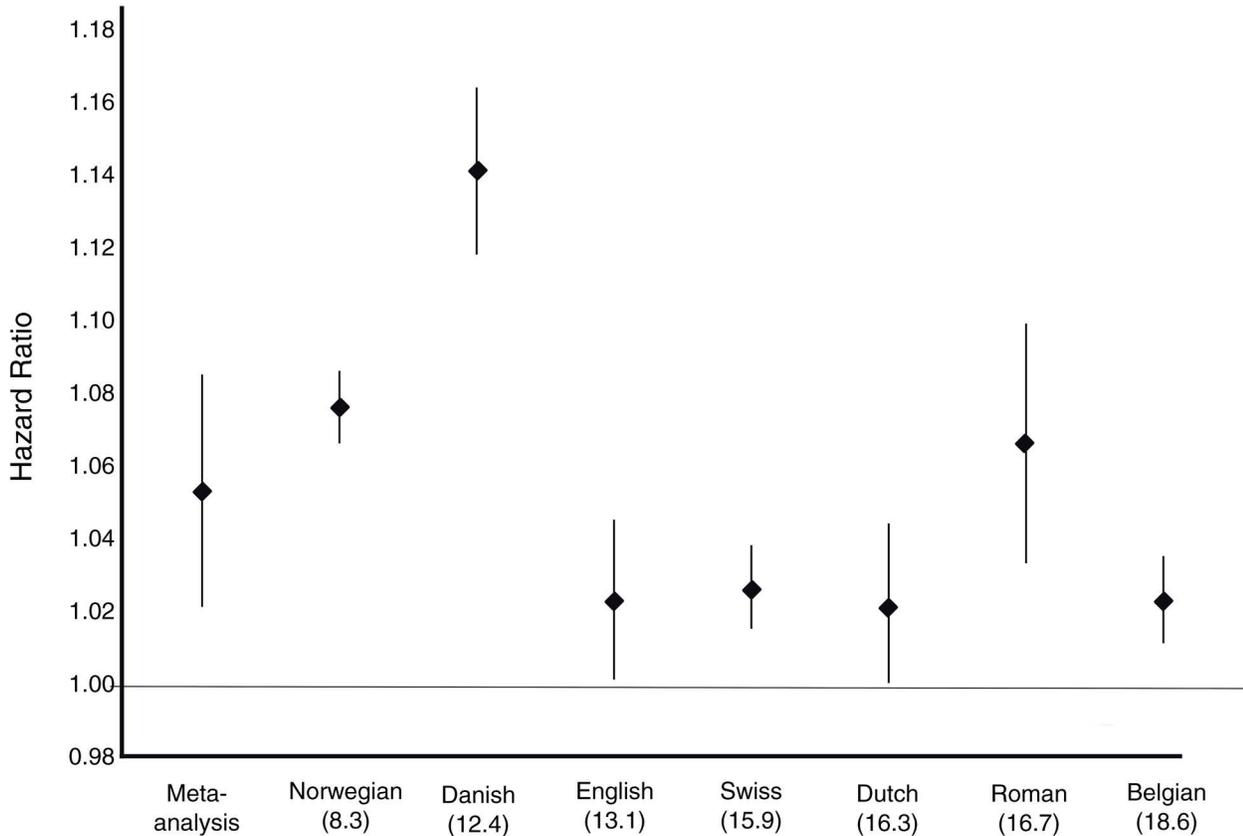
effect estimates for associations between mortality and exposures among the administrative cohorts, namely relatively high effect estimates (across pollutants and causes of death) for the Danish cohort, and generally lower effect estimates among the English, Swiss, Dutch, and Belgian cohorts (e.g., Commentary Figure 2). Note also in that figure, that for Norway, with mean exposures of only 8.3 µg/m<sup>3</sup>, they reported an HR of 1.076; 95% CI: 1.066–1.086. In sum, a key strength of the report is the presentation of evidence of both overall, summary estimates of associations between exposures and health outcomes as well as local and regional results.

Overall, the Panel was impressed with the numerous sensitivity analyses conducted and generally found them helpful in supporting the robustness and interpretation of the findings. Broadly, these sensitivity analyses related to different approaches to exposure specification (e.g., estimating exposures only at baseline versus using time-varying estimates), confounder control (e.g., adjusting for additional confounders and noise), and exploration of concentration–response functions.

The additional adjustments for road traffic noise in some models were deemed by the Panel to provide an especially worthwhile and interesting contribution to the Report. In most cases, the air pollution associations were robust to noise adjustment. In the Swiss cohort, however, effect estimates for cardiovascular mortality were attenuated substantially by this adjustment. Here, the investigators suggested that differences in measurement error, populations, and correlations with the pollutants may have contributed to differences in the effect of noise on these associations across cohorts.

An additional analysis that the Panel appreciated was the comparison of associations between cause-specific mortality and PM<sub>2.5</sub> using the independent exposure model created by the MAPLE team. The fact that the exposure patterns and epidemiological results produced with the MAPLE exposure model were similar to those reported with the main ELAPSE model (except for some countries, notably Norway where the MAPLE HR is fully null) generally reinforced the robustness of the study's main findings.

Although results from the subset analyses restricted to participants with mean exposures below selected concentrations provide additional (and generally interesting) support for evidence of associations between exposure and health at low concentrations, it is important to acknowledge that these analyses are based on notably reduced and less-balanced mixes of cohorts. For the pooled cohort, all cohorts contributed to the PM<sub>2.5</sub> analysis for exposures below 15 µg/m<sup>3</sup>, but the Dutch and German cohorts con-



**Commentary Figure 2. Associations between PM<sub>2.5</sub> and natural-cause mortality: hazard ratios and meta-analysis of seven administrative cohorts, fully adjusted Model 3.** HRs per 5 µg/m<sup>3</sup>. Numbers next to cohort names indicate mean concentrations of PM<sub>2.5</sub> in µg/m<sup>3</sup> assigned to participants. (Source: IR Figure 14 and Appendix 3, Table A1.)

tributed little. Moreover, only the Swedish, Danish, French, and Austrian cohorts contributed to the analyses below 10 and 12 µg/m<sup>3</sup>, and the former analysis was composed primarily of participants in the Swedish cohort. The sample sizes for these latter two analyses were also limited to only 52,528 and 25,422 participants (down from the 325,367 included in the full, pooled cohort; cf. Commentary Table 4). For the administrative cohorts, all cohorts contributed to the PM<sub>2.5</sub> analysis below 15 µg/m<sup>3</sup>. Several cohorts were excluded from analyses below 12 µg/m<sup>3</sup>, and the analysis below 10 µg/m<sup>3</sup> was dominated by the Norwegian cohort. Sample sizes for analyses with the administrative cohorts for PM<sub>2.5</sub> below 10 and 12 µg/m<sup>3</sup> were limited to only ~2 and ~4 million participants, respectively (down from just over 28 million across all administrative cohorts). The limited representation of participants from many cohorts (countries) in the subset analyses therefore

may limit the generalizability of findings from these particular analyses.

An important design issue to acknowledge is that the survival models were stratified by cohort in the pooled cohort analysis. These analyses were stratified by cohort (along with by sex and age) to account for differences in baseline hazards between the cohorts (e.g., the baseline risk of dying from all-natural causes may differ between countries). Although the Panel concurred that this was the appropriate approach given the differences in populations, it is important to highlight the implications for how exposure contrasts are therefore treated in the analyses. Ultimately, this approach entails that exposure contrasts are captured only between participants within each individual cohort but not across them. As such, the risk associated with exposure for a participant with low exposure in one country is not really compared with that of someone with low exposure in another country.

## DISCUSSION OF THE FINDINGS AND INTERPRETATION

Overall, the Panel found that the investigators have carefully reported and discussed the results of multiple analyses in detail, from separate cohorts (administrative or otherwise), together with a comprehensive set of sensitivity analyses. The presentation of this volume of information within a clear and structured narrative presents challenges, namely balancing the focus on overall patterns and trends versus on regional variations and heterogeneity in associations between places or cohorts. In this context, the Panel commends the investigators for preparing a comprehensive report and for the careful, fair, and balanced interpretation of the extensive array of results.

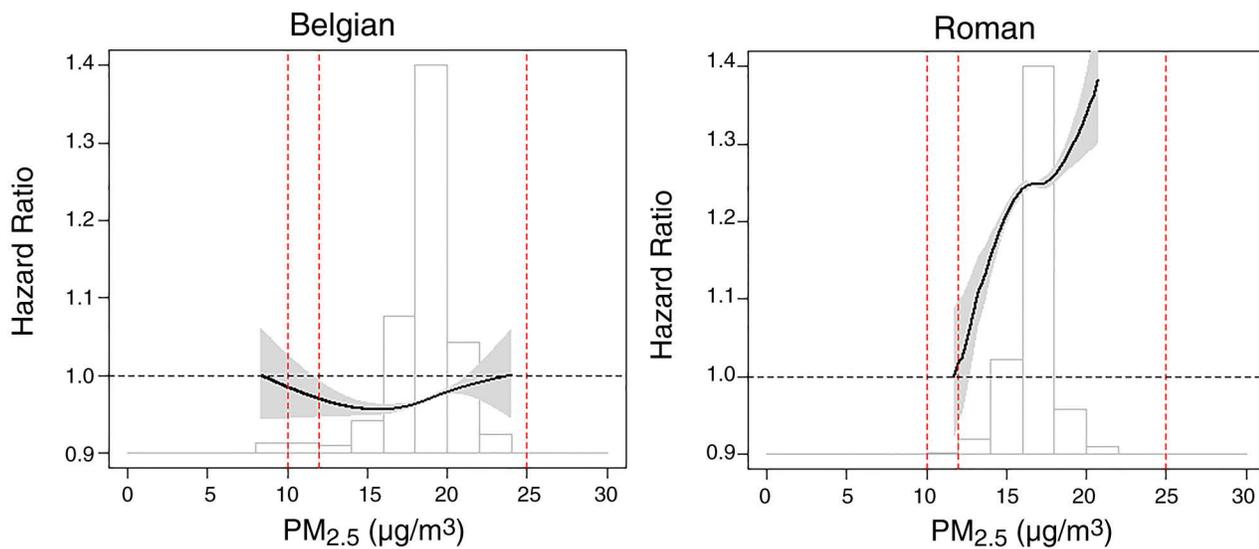
To some extent, the lack of detail provided on the cohorts, including data on recruitment, is a limitation of the Report (although those details have been reported in earlier journal articles). For example, the eligibility criteria for inclusion/exclusion in the various cohorts was not always clear, nor were patterns of participation, nor of attrition or loss of follow-up. It was also unclear whether the given cohorts were population-based (i.e., drawn from all possible people in a given age range and location) nor if they were nationally representative (or if participants tended to be of higher or lower socioeconomic position, for instance). These missing details do not necessarily impact the internal validity of the individual cohorts, unless they obscure the possibility of selection biases, but they may have implications for combining or pooling them and may affect the generalizability of the findings.

Generally, the Panel appreciated that the Report explored several approaches to modeling concentration–response functions and that these models were assessed carefully. They did feel however, that the Report was lacking in interpretation and explanation of the heterogeneity in the shapes reported between the administrative cohorts (beyond acknowledging that the cohorts differed in mean exposures) (e.g., Commentary Figure 3). As a result, readers are left wondering to what extent the heterogeneity in shapes is due to differences in population characteristics, pollutant mixtures, variability in exposure misclassification, or some other unknown factors.

The Panel noted that systematic differences in the exposure distributions for different cohorts/countries in the pooled analyses, however, means that the concentration–response function value at a given exposure is driven primarily by the cohort contributing observations at that exposure level. Most strikingly, as noted above, the results for the lowest PM<sub>2.5</sub> exposures are informed almost exclusively by the Norwegian and Swedish (i.e., Stockholm) cohorts.

Although the Panel thought the careful attention to the potential shape of the concentration–response function and the use of splines was a strength of the work, they found the SCHIF approach to be difficult to interpret. Although the desire to use a function that is constrained to take a plausible functional form is understandable and the Panel appreciated that the investigators sought to harmonize their results with those from the MAPLE team, the Panel felt that the use of the minimum exposure as the reference exposure made the SCHIF uncertainty estimates difficult, if not impossible, to interpret from a practical perspective. In particular, while one can use the fitted curves to estimate the risk under one exposure relative to another exposure (by reading off two hazard ratios and dividing), one can only assess uncertainty for an exposure compared to the reference exposure, which may not be of much practical use when using the minimum as the reference. That is, the choice of the reference causes the seemingly counterintuitive result that uncertainty is lowest at the place on the curve where exposure data are most sparse (i.e., at the lowest concentrations), when this is simply the result of the chosen reference value.

An unexpected finding in this Report is the near consistent inverse associations between warm season O<sub>3</sub> and the risk of the various health outcomes (e.g., Commentary Tables 3 and 5). The investigators reported that O<sub>3</sub> was highly (negatively) correlated with PM<sub>2.5</sub> and NO<sub>2</sub>, and that concentrations did not decrease substantially over time. Moreover, O<sub>3</sub> remained inversely associated with the various health outcomes in multipollutant models, although these were generally attenuated. They also note that the associations with O<sub>3</sub> were observed over a relatively narrow range of exposure (i.e., 98% of the pooled cohort had exposures between 60 and 100 µg/m<sup>3</sup> and 68% were in the range 80–100 µg/m<sup>3</sup>). They concluded that their study was therefore less suited to assessing associations with O<sub>3</sub> than those based in other locations (e.g., Canada or the United States) where greater variability in exposures have been reported. Findings from subsequent analyses carried out by the ELAPSE investigators (not included in the Report, but as reported in a recent publication [Strak et al., in press]) show that the inverse association with O<sub>3</sub> in the pooled cohort was attenuated when the large Austrian cohort (VHM&PP) that experienced the highest O<sub>3</sub> concentrations was not included, but only when coupled with adjustment for any of the copollutants. With additional adjustment for noise, the inverse association was attenuated to unity. Although interesting, interpretation of the impact of this restriction of the study sample population and additional covariate adjustments on the associations with O<sub>3</sub> is complicated and requires further investigation and deliberation.



**Commentary Figure 3. Associations between natural-cause mortality and PM<sub>2.5</sub> in the Belgian and Roman administrative cohorts.** Concentration–response curves are shown as natural cubic splines with three degrees of freedom (df). Hazard ratios expressed relative to minimum exposure. Shaded area represents 95% confidence interval. Red dotted lines are, from right to left, the “old” U.S. EPA NAAQS, the current U.S. EPA NAAQS, and the 2005 WHO air quality guideline values. X-axis truncated at 30 µg/m<sup>3</sup>. Hazard ratios are expressed relative to minimum exposure (i.e., 3.24 µg/m<sup>3</sup>). Histograms represent exposure distributions. (Source: IR Figure 15.)

As noted above, the investigators reported null associations between natural-cause mortality and both PM<sub>2.5</sub> and BC in their two-pollutant models adjusted for NO<sub>2</sub> among participants in the administrative cohort. In the case of the pooled cohort, however, the effect estimates for PM<sub>2.5</sub> and BC, while attenuated, remained positive and significant in such models. Ultimately, the impact of NO<sub>2</sub> adjustment on these estimates is open to interpretation, with no clear answer.

Together, the results of this single Report provide a veritable body of evidence for associations between exposures and health based on different populations and age ranges, and from countries with different cultures, geographies, and health care systems. Despite analytic challenges and limitations associated with combining and pooling various datasets and effect estimates, more can be learned from a multicohort study such as this than one based on only a single, large national cohort.

## CONCLUSIONS

Based on its thorough review, the Panel concluded that this Report provided very good evidence of associations between long-term exposures to relatively low concentrations of ambient air pollution and several important health

endpoints, at concentrations below the current levels of the current EU limit values. They note, however, that there remains limited evidence for associations at the lowest concentrations, with only data from Norway and Stockholm providing the bulk of the evidence in those settings.

As described above, a key highlight of this study, which is not present in most others, is the presentation of results from several different countries (pooled and individually) and the reporting of the cross-country heterogeneity in associations. The large number of cohorts included is, however, both a boon and challenge for readers seeking to synthesize the key messages of the study, to weigh and interpret the heterogeneity in results between places, and to identify overarching conclusions. In sum, this Report provided a comprehensive overview, discussion, and interpretation of the many analyses conducted. The Report includes a wealth of findings that will be of great interest and value to a wide readership of researchers and decision makers.

## ACKNOWLEDGMENTS

The HEI Review Committee is grateful to the Low-Exposure Epidemiology Studies Review Panel for their thorough review of the study. The Committee is also grateful to Hanna Boogaard for oversight of the study, to Eleanne van Vliet for assistance with review of the report, to Dan Crouse and Joanna Keel for assistance with review of the report and in preparing its Commentary, to Mary Brennan for editing of this Report and its Commentary, and to Hope Green, Hilary Polk, and Kristin Eckles for their roles in preparing this Research Report for publication.

## REFERENCES

- Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedeli X, et al. 2013. Development of NO<sub>2</sub> and NO<sub>x</sub> land use regression models for estimating air pollution exposure in 36 study areas in Europe: the ESCAPE project. *Atmos Environ* 72:10–23.
- Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B, et al. 2014a. Effects of long-term exposure to air pollution on natural-cause mortality: An analysis of 22 European cohorts within the multicentre ESCAPE project. *Lancet* 383:785–795.
- Beelen R, Stafoggia M, Raaschou-Nielsen O, Andersen ZJ, Xun WW, Katsouyanni K, et al. 2014b. Long-term exposure to air pollution and cardiovascular mortality: An analysis of 22 European cohorts. *Epidemiology* 25:368–378; doi:10.1097/EDE.0000000000000076.
- Brandt J, Silver JD, Frohn LM, Geels C, Gross A, Hansen AB, et al. 2012. An integrated model study for Europe and North America using the Danish Eulerian Hemispheric Model with focus on intercontinental transport of air pollution. *Atmos Environ* 53:156–176.
- Brauer M, Brook JR, Christidis T, Chu Y, Crouse DL, Erickson A, et al. 2019. Mortality–ONEair pollution associations in low-exposure environments (MAPLE): Phase 1. *Res Rep Health Eff Inst.* 2019:203.
- Crouse DL, Peters PA, Hystad P, Brook JR, van Donkelaar A, Martin RV, et al. 2015. Ambient PM<sub>2.5</sub>, O<sub>3</sub>, and NO<sub>2</sub> exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environ Health Perspect* 123:1180–1186.
- Crouse DL, Peters PA, van Donkelaar A, Goldberg MS, Villeneuve PJ, Brion O, et al. 2012. Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: A Canadian national-level cohort study. *Environ Health Perspect* 120:708–714; doi:10.1289/ehp.1104049.
- Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Dominici F, et al. 2017. Air pollution and mortality in the Medicare population. *N Engl J Med* 376:2513–2522.
- Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. 2012. Development of land use regression models for PM<sub>2.5</sub>, PM<sub>2.5</sub> absorbance, PM<sub>10</sub>, and PM<sub>coarse</sub> in 20 European study areas; results of the ESCAPE project. *Environ Sci Technol* 46:11195–11205.
- Erickson AC, Brauer M, Christidis T, Pinault L, Crouse DL, van Donkelaar A, et al. 2019. Evaluation of a method to indirectly adjust for unmeasured covariates in the association between fine particulate matter and mortality. *Environ Res* 175:108–116.
- European Commission. 2018. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions — A Europe that protects: Clean air for all. Available: [http://ec.europa.eu/environment/air/pdf/clean\\_air\\_for\\_all.pdf](http://ec.europa.eu/environment/air/pdf/clean_air_for_all.pdf) [accessed 24 April 2019].
- European Commission. 2019a. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions — The European Green Deal (COM(2019) 6430 final of 11 December 2019).
- European Commission. 2019b. Commission Staff Working Document — Fitness check of the EU Ambient Air Quality Directives (SWD(2019) 427 final of 28 November 2019).
- European Environment Agency (EEA). 2020. Air Quality in Europe, 2020 Report. Luxembourg:Publications Office of the European Union.
- European Union. 2004. Directive 2004/107/EC of the European Parliament and of the Council of 15 December 2004 relating to arsenic, cadmium, mercury, nickel and polycyclic aromatic hydrocarbons in ambient air. Available: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2005:023:0003:0016:EN:PDF> [accessed 24 April 2019].
- European Union. 2008. Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02008L0050-20150918> [accessed 24 April 2019].

- European Union. 2016. Directive (EU) 2016/2284/EC of the European Parliament and of the Council of 14 December 2016 on the reduction of national emissions of certain atmospheric pollutants, amending Directive 2003/35/EC and repealing Directive 2001/81/EC. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016L2284> [accessed 24 April 2019].
- Global Burden of Disease (GBD) 2019 Risk Factors Collaborators. 2020. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 396:1223–1249; doi:10.1016/S0140-6736(20)30752-2.
- Hales S, Blakely T, Woodward A. 2012. Air pollution and mortality in New Zealand: Cohort study. *J Epidemiol Community Health* 66:468–473; doi:10.1136/jech.2010.112490.
- Health Effects Institute. 2020. State of Global Air 2020. Special Report. Boston, MA:Health Effects Institute.
- Hoek G. 2017. Methods for assessing long-term exposures to outdoor air pollutants. *Curr Environ Health Rep* 4:450–462; doi: 10.1007/s40572-017-0169-5.
- Inness A, Baier F, Benedetti A, Bouarar I, Chabrilat S, Clark H, et al. 2013. The MACC reanalysis: An 8-yr data set of atmospheric composition. *Atmos Chem Phys* 13:4073–4109; doi.org/10.5194/acp-13-4073-2013.
- Nasari MN, Szyszkowicz M, Chen H, Crouse DL, Turner MC, Jerrett M, et al. 2016. A class of non-linear exposure-response models suitable for health impact assessment applicable to large cohort studies of ambient air pollution. *Air Qual Atmos Health* 9:961–972; doi: 10.1007/s11869-016-0398-z.
- Pinault L, Tjepkema M, Crouse DL, Weichenthal S, van Donkelaar A, Martin RV, et al. 2016. Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the Canadian community health survey cohort. *Environ Health* 15:18–31; doi: 10.1186/s12940-016-0111-6.
- Strak M, Weinmayr G, Rodopoulou S, Chen J, de Hoogh K, Andersen ZJ, et al. In press. Long-term exposure to low-level air pollution and natural mortality: A pooled analysis of eight European cohorts within the ELAPSE project. *BMJ*.
- U.S. EPA (Environmental Protection Agency). 2004. Vol I. PM Air Quality Criteria Document. Washington, DC:U.S. Environmental Protection Agency.
- U.S. EPA (Environmental Protection Agency). 2013. Integrated science assessment (ISA) for ozone and related photochemical oxidants. Available: [www.epa.gov/isa/integrated-science-assessment-isa-ozone-and-related-photochemical-oxidants](http://www.epa.gov/isa/integrated-science-assessment-isa-ozone-and-related-photochemical-oxidants) [accessed 24 April 2019].
- U.S. EPA (Environmental Protection Agency). 2015. Preamble to the Integrated Science Assessments (ISA). EPA/600/R-15/067. Washington, DC:U.S. Environmental Protection Agency.
- U.S. EPA (Environmental Protection Agency). 2016. Particulate Matter (PM<sub>2.5</sub>) Trends. Available: [www.epa.gov/air-trends/particulate-matter-pm25-trends](http://www.epa.gov/air-trends/particulate-matter-pm25-trends) [accessed 24 April 2019].
- U.S. EPA (Environmental Protection Agency). 2018. Particulate Matter (PM<sub>2.5</sub>) Trends. Available: [www.epa.gov/air-trends/particulate-matter-pm25-trends](http://www.epa.gov/air-trends/particulate-matter-pm25-trends) [accessed 24 April 2019].
- van Donkelaar A, Martin RV, Brauer M, Hsu NC, Kahn RA, Levy RC, et al. 2016. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. *Environ Sci Technol* 50:3762–3772; doi: 10.1021/acs.est.5b05833.
- van Donkelaar A, Martin RV, Spurr RJ, Burnett RT. 2015. High-resolution satellite-derived PM<sub>2.5</sub> from optimal estimation and geographically weighted regression over North America. *Environ Sci Technol* 49:10482–10491; doi.org/10.1021/acs.est.5b02076.
- World Health Organization (WHO). 2006. Air Quality Guidelines: Global Update 2005: Particulate Matter, Ozone, Nitrogen Dioxide, and Sulfur Dioxide. Copenhagen:WHO Regional Office for Europe.



## RELATED HEI PUBLICATIONS

Number	Title	Principal Investigator	Date
<b>Research Reports</b>			
203	Mortality–Air Pollution Associations in Low-Exposure Environments (MAPLE): Phase 1	M. Brauer	2019
200	Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution: Phase 1	F. Dominici	2019
193	Particulate Air Pollutants, Brain Structure, and Neurocognitive Disorders in Older Women	J. Chen	2017
192	Multicenter Ozone Study in older Subjects (MOSES) <i>Part 1. Effects of Exposure to Low Concentrations of Ozone on Respiratory and Cardiovascular Outcomes</i>	M. W. Frampton	2017
	<i>Part 2. Effects of Personal and Ambient Concentrations of Ozone and Other Pollutants on Cardiovascular and Pulmonary Function</i>	D.Q. Rich	2020
190	The Effects of Policy-Driven Air Quality Improvements on Children’s Respiratory Health	F. Gilliland	2017
187	Causal Inference Methods for Estimating Long-Term Health Effects of Air Quality Regulations	C. M. Zigler	2016
183	Development of Statistical Methods for Multipollutant Research <i>Part 1. Statistical Learning Methods for the Effects of Multiple Air Pollution Constituents</i>	B. A. Coull	2015
	<i>Part 2. Development of Enhanced Statistical Methods for Assessing Health Effects Associated with an Unknown Number of Major Sources of Multiple Air Pollutants</i>	E. S. Park	2015
	<i>Part 3. Modeling of Multipollutant Profiles and Spatially Varying Health Effects with Applications to Indicators of Adverse Birth Outcomes</i>	J. Molitor	2016
178	National Particle Component Toxicity (NPACT) Initiative Report on Cardiovascular Effects	S. Vedal	2013
177	National Particle Component Toxicity (NPACT) Initiative: Integrated Epidemiologic and Toxicologic Studies of the Health Effects of Particulate Matter Components	M. Lippmann	2013
142	Air Pollution and Health: A European and North American Approach	K. Katsouyanni	2009
140	Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality	D. Krewski	2009
139	Effects of Long-Term Exposure to Traffic-Related Air Pollution on Respiratory and Cardiovascular Mortality in the Netherlands: The NLCS-AIR Study	B. Brunekreef	2009
127	Personal, Indoor, and Outdoor Exposures to PM <sub>2.5</sub> and Its Components for Groups of Cardiovascular Patients in Amsterdam and Helsinki	B. Brunekreef	2005

(Continued)

Copies of these reports can be obtained from HEI; PDFs are available for free downloading at [www.healtheffects.org/publications](http://www.healtheffects.org/publications).

## RELATED HEI PUBLICATIONS

---

<b>Number</b>	<b>Title</b>	<b>Principal Investigator</b>	<b>Date</b>
<b>Special Reports</b>			
	Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality	Health Effects Institute	2000
<b>Perspectives</b>			
1	Airborne Particles and Health: HEI Epidemiologic Evidence	Health Effects Institute	2001
2	Understanding the Health Effects of Components of the Particulate Matter Mix: Progress and Next Steps	Health Effects Institute	2002
3	Understanding the Health Effects of Ambient Ultrafine Particles	Health Effects Institute	2013

---

*Copies of these reports can be obtained from HEI; PDFs are available for free downloading at [www.healtheffects.org/publications](http://www.healtheffects.org/publications).*

## ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion	MAPLE	Mortality–Air Pollution Associations in Low-Exposure Environments study
AOD	aerosol optical density	NAAQS	National Ambient Air Quality standards (U.S.)
BC	black carbon	NO <sub>2</sub>	nitrogen dioxide
BIC	Bayesian information criterion	NUTS-1	Nomenclature of Territorial Units for Statistics
BMI	body mass index	O <sub>3</sub>	ozone
CI	confidence interval	P	percentile
CRF	concentration–response function	PM <sub>10</sub>	particulate matter ≤10 μm in aerodynamic diameter
COPD	chronic obstructive pulmonary disease	PM <sub>2.5</sub>	particulate matter ≤2.5 μm in aerodynamic diameter
DAG	directed acyclic graph	R <sup>2</sup>	coefficient of determination
DEHM	Danish Eulerian Hemispheric Model	RMSE	root mean squared error
ESCAPE	European Study of Cohorts for Air Pollution Effects	SAT	satellite-derived
ELAPSE	Effects of Low-Level Air Pollution: A Study in Europe	SCHIF	Shape-Constrained Health Impact Function
EU	European Union	SES	socioeconomic status
GEOS-Chem	global three-dimensional chemical transport model	Swiss TPH	Swiss Tropical and Public Health Institute
GIS	geographic information system	U.S. EPA	U.S. Environmental Protection Agency
HR	hazard ratio	WHO	World Health Organization
ICD	International Classification of Diseases		
LUR	land use regression		



# HEI BOARD, COMMITTEES, and STAFF

## Board of Directors

**Richard A. Meserve, Chair** *Senior of Counsel, Covington & Burling LLP; President Emeritus, Carnegie Institution for Science; former Chair, U.S. Nuclear Regulatory Commission*

**Enriqueta Bond** *President Emerita, Burroughs Wellcome Fund*

**Jo Ivey Boufford** *President, International Society for Urban Health*

**Homer Boushey** *Emeritus Professor of Medicine, University of California–San Francisco*

**Michael T. Clegg** *Professor of Biological Sciences, University of California–Irvine*

**Jared L. Cohon** *President Emeritus and Professor, Civil and Environmental Engineering and Engineering and Public Policy, Carnegie Mellon University*

**Stephen Corman** *President, Corman Enterprises*

**Martha J. Crawford** *Dean, Jack Welch College of Business and Technology, Sacred Heart University*

**Michael J. Klag** *Dean Emeritus and Second Century Distinguished Professor, Johns Hopkins Bloomberg School of Public Health*

**Alan I. Leshner** *CEO Emeritus, American Association for the Advancement of Science*

**Karen C. Seto** *Frederick C. Hixon Professor of Geography and Urbanization Science, Yale School of the Environment*

## Research Committee

**David A. Savitz, Chair** *Professor of Epidemiology, School of Public Health, and Professor of Obstetrics and Gynecology, Alpert Medical School, Brown University*

**Jeffrey R. Brook** *Senior Research Scientist, Air Quality Research Division, Environment Canada, and Assistant Professor, University of Toronto, Canada*

**Francesca Dominici** *Professor of Biostatistics and Senior Associate Dean for Research, Harvard T.H. Chan School of Public Health*

**David E. Foster** *Phil and Jean Myers Professor Emeritus, Department of Mechanical Engineering, Engine Research Center, University of Wisconsin–Madison*

**Amy H. Herring** *Sara & Charles Ayres Professor of Statistical Science and Global Health, Duke University*

**Barbara Hoffmann** *Professor of Environmental Epidemiology, Institute of Occupational, Social, and Environmental Medicine, University of Düsseldorf, Germany*

**Neil Pearce** *Professor of Epidemiology and Biostatistics, London School of Hygiene and Tropical Medicine, United Kingdom*

**Allen L. Robinson** *Raymond J. Lane Distinguished Professor and Head, Department of Mechanical Engineering, and Professor, Department of Engineering and Public Policy, Carnegie Mellon University*

**Ivan Rusyn** *Professor, Department of Veterinary Integrative Biosciences, Texas A&M University*

## Review Committee

**Melissa Perry, Chair** *Professor and Chair, Department of Environmental and Occupational Health, George Washington University Milken Institute School of Public Health*

**Kiros Berhane** *Professor and Chair, Department of Biostatistics, Mailman School of Public Health, Columbia University*

**Michael Jerrett** *Professor and Chair, Department of Environmental Health Sciences, Fielding School of Public Health, University of California–Los Angeles*

**Frank Kelly** *Henry Battcock Chair of Environment and Health and Director of the Environmental Research Group, Imperial College London School of Public Health, United Kingdom*

*Continues next page*

# HEI BOARD, COMMITTEES, and STAFF

**Jana B. Milford** *Professor, Department of Mechanical Engineering and Environmental Engineering Program, University of Colorado–Boulder*

**Jennifer L. Peel** *Professor of Epidemiology, Colorado School of Public Health and Department of Environmental and Radiological Health Sciences, Colorado State University*

**Roger D. Peng** *Professor of Biostatistics, Johns Hopkins Bloomberg School of Public Health*

## Officers and Staff

**Daniel S. Greenbaum** *President*

**Robert M. O’Keefe** *Vice President*

**Ellen K. Mantus** *Director of Science*

**Donna J. Vorhees** *Director of Energy Research*

**Anne Moon M. van Erp** *Deputy Director of Science*

**Thomas J. Champoux** *Director of Science Communications*

**Jacqueline C. Rutledge** *Director of Finance and Administration*

**Emily Alden** *Corporate Secretary*

**Lee Ann Adelsheim** *Research Assistant*

**Palak Balyan** *Consulting Staff Scientist*

**Hanna Boogaard** *Consulting Principal Scientist*

**Sofia Chang-DePuy** *Digital Communications Manager*

**Aaron J. Cohen** *Consulting Principal Scientist*

**Dan Crouse** *Senior Scientist*

**Robert M. Davidson** *Staff Accountant*

**Philip J. DeMarco** *Compliance Manager*

**Kristin C. Eckles** *Senior Editorial Manager*

**Hope Green** *Editorial Project Manager*

**Joanna Keel** *Research Assistant*

**Yi Lu** *Staff Scientist*

**Lissa McBurney** *Science Administrative Assistant*

**Janet I. McGovern** *Executive Assistant*

**Martha Ondras** *Research Fellow*

**Pallavi Pant** *Senior Staff Scientist*

**Allison P. Patton** *Staff Scientist*

**Hilary Selby Polk** *Managing Editor*

**Anna S. Rosofsky** *Staff Scientist*

**Robert A. Shavers** *Operations Manager*

**Eva Tanner** *Consulting Staff Scientist*

**Eleanne van Vliet** *Staff Scientist*



HEALTH  
EFFECTS  
INSTITUTE

75 Federal Street, Suite 1400  
Boston, MA 02110, USA  
+1-617-488-2300  
[www.healtheffects.org](http://www.healtheffects.org)

RESEARCH  
REPORT

Number 208  
September 2021

