**Appendix**

Hvidtfeldt et al. Long-term low-level ambient air pollution exposure and risk of lung cancer – a pooled analysis of 7 European cohorts.

**Table of content**:

1. Characteristics of each of the seven cohorts and their participants

2. Supplementary tables

Table S1: Subset analyses

Table S2: Threshold analyses

Table S3: Correlation matrices between air pollutant exposures

Table S4: Exposure estimates back-extrapolated to baseline and time-varying exposure analyses

Table S5: ELAPSE versus ESCAPE exposure estimates

Table S6: Sensitivity analyses additional covariates

Table S7: Sensitivity analyses excluding cohorts

Table S8: Effect modification by smoking

Table S9: Findings of previous studies on air pollutants and Lung cancer

3. Supplementary figures

Figure S1.Temporal trends in PM2.5, NO2, BC and O3 concentrations based on the DEHM for different regions in Europe

Figure S2: Box plots of exposures by (sub-) cohort

Figure S3: Natural spline models of pollutants and lung cancer (3 df)

Figure S4:Two- and three pollutant models

Figure S5: Scatterplot of ELAPSE vs. ESCAPE exposure by cohort

Figure S6: Comparing different options to adjust for differences between cohorts

Figure S7: Hazard ratios for associations between air pollution and lung cancer in the individual cohorts

**Characteristics of the included cohorts**

**CEANS** (Cardiovascular Effects of Air Pollution and Noise in Stockholm)

All participants resided in Stockholm County, Sweden. The cohort is comprised of four sub-cohorts:

The Stockholm Diabetes Preventive Program (SDPP) is a population-based prospective study of 7,949 subjects aged 35–54 years.1 The Stockholm Cohort of 60-year-olds (SIXTY) consists of a random population sample of one-third of all men and women living in Stockholm County turning 60 years between August 1997 and March 1999.2 The Screening Across the Lifespan Twin Study (SALT) is a sub-study of the Swedish Twin Registry.3 All Swedish complete twin-pairs born in Sweden before 1959 were contacted. Included in this study are 7,043 SALT participants who lived in Stockholm County. Lastly, The Swedish National Study of Aging and Care in Kungsholmen (SNAC-K) randomly sampled individuals 60+ years of age from a central area in Stockholm.4

|  |  |  |
| --- | --- | --- |
|  |  | **CEANS, sub-cohorts** |
| Variable | **Total** | SDPP | SIXTY | SALT | SNAC-K |
| Baseline year, range  | 1992–2004 | 1992–1998 | 1997–1999 | 1998–2003 | 2001–2004 |
| Enrolled, N | 22,587 | 7,949 | 4,232 | 7,043 | 3,363 |
| Exclusionsa | 2,519 | 536 | 373 | 944 | 666 |
| Missing on covariatesb | 1,105 | 98 | 196 | 473 | 338 |
| Included, N | 18,963 | 7,315 | 3,663 | 5,626 | 2,359 |
| Age at baseline, mean (SD) | 55.8 (11.1) | 47.0 (4.9) | 60 (0) | 57.3 (10.4) | 72.5 (10.4) |
| Women, N (%) | 10,657 (56) | 4,352 (59) | 1,833 (50) | 3,001 (53) | 1,471 (62) |
| Unemployed, N (%) | 5,521 (29) | 669 (9) | 1,176 (32) | 1,882 (33) | 1,794 (76) |
| Marital status |  |  |  |  |  |
|  | Single | 2,525 (13) | 1,186 (16) | 172 (5) | 785 (14) | 382 (16) |
|  | Married | 13,793 (73) | 6,129 (84) | 2,729 (75) | 3,840 (68) | 1,095 (46) |
|  | Divorced | 1,530 (8) | - | 575 (16) | 629 (11) | 326 (14) |
|  | Widowed | 1,115 (6) | - | 187 (5) | 372 (7) | 556 (24) |
| Smoking status, N (%) |  |  |  |  |  |
|  | Current  | 4,201 (22) | 1,899 (26) | 767 (21) | 1,185 (21) | 350 (15) |
|  | Previous | 6,806 (36) | 2,657 (36) | 1,400 (38) | 1,857 (33) | 892 (38) |
|  | Never | 7,956 (42) | 2,759 (38) | 1,496 (41) | 2,584 (46) | 1,117 (47) |
| Smoking intensity, g/d mean (SD)c | 13.1 (7.8) | 13.5 (7.4) | 13.3 (7.7) | 12.7 (8.1) | 11.7 (8.3) |
| Smoking duration, yrs mean (SD)c | 33.4 (10.9) | 27.8 (8.6) | 36.2 (10.1) | 37.6 (9.1) | 43.2 (13.5) |
| BMI, kg/m2  |  |  |  |  |  |
|  | < 18.5 | 209 (1) | 49 (1) | 23 (1) | 75 (1) | 62 (3) |
|  | 18.5–24.9 | 9,114 (48) | 3,494 (48) | 1,282 (35) | 3,297 (59) | 1,041 (44) |
|  | 25.0–29.9 | 7,333 (39) | 2,861 (39) | 1,635 (45) | 1,887 (34) | 950 (40) |
|  | 30.0+ | 2,307 (12) | 911 (12) | 723 (20) | 367 (7) | 306 (13) |
| Neighborhood incomed, mean (SD) | 25.2 (5.6) | 24.3 (4.2) | 24.7 (6.9) | 25.4 (6.6) | 28.7 (2.2) |
| aDue to failed exposure assignment or any cancer before baselinebMain model 3: age, sex, year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status and 2001 mean income at the neighborhood levelcAmong current smokersdEUR per 1,000, year 2001. Neighborhood defined as municipality |

**Main references:**

1. Eriksson AK, Ekbom A, Granath F, *et al*. Psychological distress and risk of pre-diabetes and Type 2 diabetes in a prospective study of Swedish middle-aged men and women. Diabet Med 2008;25:834–42.

2. Wändell PE, Wajngot A, de Faire U, *et al*. Increased prevalence of diabetes among immigrants from non-European countries in 60-year-old men and women in Sweden. Diabetes Metab 2007;33:30–6.

3. Zagai U, Lichtenstein P, Pedersen NL, Magnusson PKE. The Swedish Twin Registry: Content and Management as a Research Infrastructure. Twin Res Hum Genet. 2019 Dec;22(6):672-680.

4. Lagergren M, Fratiglioni L, Hallberg IR, *et al*. A longitudinal study integrating population, care and social services data. The Swedish National study on Aging and Care (SNAC). Aging Clin Exp Res 2004;16:158–68.

**DCH** (Diet, Cancer and Health)

Participants were recruited among persons aged 50 years and older from the areas of greater Copenhagen and Aarhus, Denmark.

|  |  |
| --- | --- |
| Variable | **Total** |
| Baseline year, range  | 1993–1997 |
| Enrolled, N | 57,053 |
| Exclusionsa | 745 |
| Missing on covariatesb | 2,661 |
| Included, N | 53,647 |
| Age at baseline, mean (SD) | 56.7 (4.4) |
| Women, N (%) | 28,134 (52) |
| Unemployed, N (%) | 11,650 (22) |
| Marital status |  |
|  | Single | 3,241 (6) |
|  | Married | 38,382 (72) |
|  | Divorced | 9,056 (17) |
|  | Widowed | 2,968 (6) |
| Smoking status, N (%) |  |
|  | Current  | 19,459 (36) |
|  | Previous | 14,959 (28) |
|  | Never | 19,229 (36) |
| Smoking intensity, g/d mean (SD)c | 16.4 (9.0) |
| Smoking duration, yrs mean (SD)c | 36.6 (7.7) |
| BMI, kg/m2  |  |
|  | < 18.5 | 421 (1) |
|  | 18.5–24.9 | 23,155 (43) |
|  | 25.0–29.9 | 22,311 (42) |
|  | 30.0+ | 7,760 (14) |
| Neighborhood incomed, mean (SD) | 20.2 (3.4) |
| aDue to failed exposure assignment or any cancer before baselinebMain model 3: age, sex, year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status and 2001 mean income at the neighborhood levelcAmong current smokersdEUR per 1,000, year 2001. Neighborhood defined as municipality |

**Main reference:**

Tjonneland A, Olsen A, Boll K et al. Study design, exposure variables, and socioeconomic determinants of participation in Diet, Cancer and Health: a population-based prospective cohort study of 57,053 men and women in Denmark. Scand J Public Health 2007; 35: 432–41

**DNC** (Danish Nurse Cohort)

The cohort was sampled among members of The Danish Nurse Organization (DNO) including both working and retired nurses. Questionnaires were mailed in 1993 to members aged 45+ years and again in 1999 with the inclusion of new members (45+ years).

|  |  |  |
| --- | --- | --- |
|  |  | **DNC, sub-cohorts** |
| Variable | **Total** | DNC-1993 | DNC-1999 |
| Baseline year | 1993,1997 | 1993 | 1999 |
| Enrolled, N | 28,731 | 19,898 | 8,833 |
| Exclusionsa | 2,752 | 1,945 | 807 |
| Missing on covariatesb | 2,961 | 2,372 | 589 |
| Included, N | 23,018 | 15,581 | 7,437 |
| Age at baseline, mean (SD) | 53.4 (8.2) | 56.0 (8.3) | 47.9 (4.1) |
| Women, N (%) | 23,018 (100) | 15,581 (100) | 7,437 (100) |
| Unemployed, N (%) | 4,946 (21) | 4,578 (29) | 368 (5) |
| Marital status |  |  |  |
|  | Single | 2,300 (10) | 1,619 (10) | 681 (9) |
|  | Married | 16,327 (71) | 10,644 (68) | 5,683 (76) |
|  | Divorced | 2,813 (12) | 1,887 (12) | 926 (12) |
|  | Widowed | 1,578 (7) | 1,431 (9) | 147 (2) |
| Smoking status, N (%) |  |  |  |
|  | Current  | 7,822 (34) | 5,762 (37) | 2,060 (28) |
|  | Previous | 6,849 (30) | 4,427 (28) | 2,422 (33) |
|  | Never | 8,347 (36) | 5,392 (35) | 2,955 (40) |
| Smoking intensity, g/d mean (SD)c | 13.7 (7.9) | 13.8 (8.1) | 13.2 (7.4) |
| Smoking duration, yrs mean (SD)c | 30.3 (9.5) | 31.4 (9.9) | 27.1 (7.1) |
| BMI, kg/m2  |  |  |  |
|  | < 18.5 | 588 (3) | 456 (3) | 132 (2) |
|  | 18.5–24.9 | 15,769 (69) | 10,710 (69) | 5,059 (68) |
|  | 25.0–29.9 | 5,336 (23) | 3,599 (23) | 1,737 (23) |
|  | 30.0+ | 1,325 (6) | 816 (5) | 509 (7) |
| Neighborhood incomed, mean (SD) | 19.2 (2.5) | 19.2 (2.6) | 19.0 (2.4) |
| aDue to failed exposure assignment or any cancer before baselinebMain model 3: age, sex, year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status and 2001 mean income at the neighborhood levelcAmong current smokersdEUR per 1,000, year 2001. Neighborhood defined as municipality |

**Main references:**

Hundrup YA, Simonsen M, Jørgensen T, Obel EB. Cohort profile: The Danish Nurse Cohort. International Journal of Epidemiology, 2012;41:1241–47.

**EPIC-NL** (European Prospective Investigation into Cancer and Nutrition, the Netherlands)

The EPIC-NL combines two Dutch EPIC-cohorts: The Monitoring Project on Risk Factors and chronic diseases in the Netherlands (MORGEN) cohort which consists of a general population sample aged 20–59 years from three Dutch towns (Amsterdam, Doetinchem and Maastricht). The Prospect is a prospective cohort study among women aged 49–70, residing in the city of Utrecht or its vicinity, who participated in the nationwide Dutch breast cancer screening programme between 1993 and 1997.

|  |  |  |
| --- | --- | --- |
|  |  | **EPIC-NL, sub-cohorts** |
| Variable | **Total** | MORGEN | PROSPECT |
| Baseline year | 1993–1997 | 1993–1997 | 1993–1997 |
| Enrolled, N | 40,011 | 22,654 | 17,357 |
| Exclusionsa | 4,850 | 2,547 | 2,303 |
| Missing on covariatesb | 3,719 | 2,305 | 1,414 |
| Included, N | 31,442 | 17,802 | 13,640 |
| Age at baseline, mean (SD) | 49.2 (11.9) | 42.7 (11.2) | 57.6 (6.0) |
| Women, N (%) | 23,330 (74) | 9,690 (54) | 13,640 (100) |
| Unemployed, N (%) |  12,093 (38) | 5,470 (31) | 6,623 (49) |
| Marital status |  |  |  |
|  | Single | 5,342 (17) | 4,573 (26) | 769 (6) |
|  | Married | 22,071 (70) | 11,561 (65) | 10,510 (77) |
|  | Divorced | 2,396 (8) | 1,320 (7) | 1,076 (8) |
|  | Widowed | 1,633 (5) | 348 (2) | 1,285 (9) |
| Smoking status, N (%) |  |  |  |
|  | Current  | 9,286 (30) | 6,173 (35) | 3,113 (23) |
|  | Previous | 9,469 (30) | 5,006 (28) | 4,463 (33) |
|  | Never | 12,687 (40) | 6,623 (37) | 6,064 (44) |
| Smoking intensity, g/d mean (SD)c | 15.0 (8.7) | 15.7 (8.6) | 13.6 (8.7) |
| Smoking duration, yrs mean (SD)c | 28.6 (11.3) | 24.5 (10.6) | 36.7 (7.7) |
| BMI, kg/m2  |  |  |  |
|  | < 18.5 | 257 (1) | 178 (1) | 79 (1) |
|  | 18.5–24.9 | 15,023 (48) | 8,923 (50) | 6,100 (45) |
|  | 25.0–29.9 | 12,057 (38) | 6,651 (37) | 5,406 (40) |
|  | 30.0+ | 4,105 (13) | 2,050 (12) | 2,055 (15) |
| Neighborhood incomed, mean (SD) | 12.6 (1.6) | 12.2 (1.6) | 13.6 (1.4) |
| aDue to failed exposure assignment or any cancer before baselinebMain model 3: age, sex, year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status and 2001 mean income at the neighborhood levelcAmong current smokersdEUR per 1,000, year 2001. Neighborhood defined as a neighborhood of a larger city |

**Main references:**

Beulens JWJ, Monninkhof EM, Verschuren WMM et al. Cohort Profile: The EPIC-NL study. International Journal of Epidemiology 2010; 39: 1170–78.

**HNR** (Heinz Nixdorf Recall study)

The cohort consists of randomly sampled persons aged 45 to 75 years from the Ruhr area (Bochum, Essen, and Mülheim), Germany.

|  |  |
| --- | --- |
| Variable | **HNR** |
| Baseline year, range  | 2000–2003 |
| Enrolled, N | 4,814 |
| Exclusionsa | 1,142 |
| Missing on covariatesb | 61 |
| Included, N | 3,611 |
| Age at baseline, mean (SD) | 59.1 (7.7) |
| Women, N (%) | 1,790 |
| Unemployed, N (%) | 2,061 |
| Marital status |  |
|  | Single | 215 |
|  | Married | 2,714 |
|  | Divorced | 367 |
|  | Widowed | 315 |
| Smoking status, N (%) |  |
|  | Current  | 886 |
|  | Previous | 1,219 |
|  | Never | 1,506 |
| Smoking intensity, g/d mean (SD)c | 19.1 (12.5) |
| Smoking duration, yrs mean (SD)c | 33.9 (9.2) |
| BMI, kg/m2  |  |
|  | < 18.5 | 11 |
|  | 18.5–24.9 | 972 |
|  | 25.0–29.9 | 1,652 |
|  | 30.0+ | 976 |
| Neighborhood incomed, mean (SD) | 25.1 (8.1) |
| aDue to failed exposure assignment or any cancer before baselinebMain model 3: age, sex, year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status and 2001 mean income at the neighborhood levelcAmong current smokersdEUR per 1,000, year 2001. Neighborhood defined as city district |

**Main reference:**

Schmermund A, Möhlenkamp S, Stang A et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: Rationale and design of the Heinz Nixdorf RECALL Study. American Heart Journal, 2002; 144: 212–2018.

**E3N** (Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale)

The cohort was selected among French women aged 40 to 65 years who were insured through a national health system that primarily covered teachers.

|  |  |
| --- | --- |
| Variable | **E3N** |
| Baseline year, range  | 1989–1991 |
| Enrolled, N | 98,995 |
| Exclusionsa | 48,790 |
| Missing on covariatesb | 12,193 |
| Included, N | 36,597 |
| Age at baseline, mean (SD) | 52.8 (6.7) |
| Women, N (%) | 36,597 (100) |
| Unemployed, N (%) | 11,542 (32) |
| Marital status |  |
|  | Single | 6,051 (17) |
|  | Married | 30,546 (83) |
|  | Divorced | - |
|  | Widowed | - |
| Smoking status, N (%) |  |
|  | Current  | 4,777 (13) |
|  | Previous | 7,042 (19) |
|  | Never | 24,778 (68) |
| Smoking intensity, g/d mean (SD)c | 11.3 (9.2) |
| Smoking duration, yrs mean (SD)c | 28.5 (7.6) |
| BMI, kg/m2  |  |
|  | < 18.5 | 1,328 (4) |
|  | 18.5–24.9 | 27,783 (76) |
|  | 25.0–29.9 | 6,195 (17) |
|  | 30.0+ | 1,291 (4) |
| Neighborhood incomed, mean (SD) | 11.2 (3.0) |
| aDue to failed exposure assignment or any cancer before baselinebMain model 3: age, sex, year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status and 2001 mean income at the neighborhood levelcAmong current smokersdEUR per 1,000, year 2001. Neighborhood defined as IRIS – a small administrative unit of a city |

**Main reference:**

Francoise Clavel-Chapelon for the E3N Study Group. Cohort Profile: The French E3N Cohort Study. International Journal of Epidemiology 2015; 44: 801–809.

**VHM&PP** (Vorarlberg Health Monitoring and Prevention Programme)

The VHM&PP is a population-based cohort recruited among all adults of the province of Vorarlberg, Austria.

|  |  |
| --- | --- |
| Variable | **VHM&PP** |
| Baseline year, range  | 1985–2005 |
| Enrolled, N | 181,350 |
| Exclusionsa | 15,597 |
| Missing on covariatesb | 25,481 |
| Included, N | 140,272 |
| Age at baseline, mean (SD) | 41.7 (14.9) |
| Women, N (%) | 78,895 (56) |
| Unemployed, N (%) | 41,008 (29) |
| Marital status |  |
|  | Single | 24,569 (18) |
|  | Married | 96,597 (69) |
|  | Divorced | 9,518 (7) |
|  | Widowed | 9,588 (7) |
| Smoking status, N (%) |  |
|  | Current  | 28,370 (20) |
|  | Previous | 8,674 (6) |
|  | Never | 103,228 (74) |
| Smoking intensity, g/d mean (SD)c | 15.6 (8.9) |
| Smoking duration, yrs mean (SD)c | 13.4 (8.2) |
| BMI, kg/m2  |  |
|  | < 18.5 | 4,382 (3) |
|  | 18.5–24.9 | 76,814 (55) |
|  | 25.0–29.9 | 43,963 (31) |
|  | 30.0+ | 15,113 (11) |
| Neighborhood incomed, mean (SD) | 22.9 (1.7) |
| aDue to failed exposure assignment or any cancer before baselinebMain model 3: age, sex, year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status and 2001 mean income at the neighborhood levelcAmong current smokersdEUR per 1,000, year 2001. EUR per 1,000, year 2001. Neighborhood defined as municipality |

**Main reference:**

Ulmer H, Kelleher CC, Fitz-Simon N et al. Secular trends in cardiovascular risk factors: an age-period cohort analysis of 698,954 health examinations in 181,350 Austrian men and women. Journal of Internal Medicine, 2007; 261: 566–576.

|  |
| --- |
| **Table S1. Subset analyses of NO2, BC, and O3 (warm season) and lung cancer** |
|  |  | **All lung cancersa****(N=3,956)** |
| **Subsetb** | **No Obs** | **HR** | **95% CI** |
| NO2 (µg/m3) |  |  |  |  |
|  | Full dataset | 307,550 | 1.02 | 0.97 | 1.07 |
|  | < 40  | 233,303 | 1.04 | 0.98 | 1.10 |
|  | < 30 | 247,039 | 1.05 | 0.96 | 1.13 |
|  | < 20 | 83,229 | 1.06 | 0.84 | 1.33 |
| BC (10-5m-1) |  |  |  |  |
|  | Full dataset | 307,550 | 1.02 | 0.97 | 1.07 |
|  | < 3 | 306,984 | 1.02 | 0.97 | 1.07 |
|  | < 2.5 | 303,132 | 1.02 | 0.97 | 1.08 |
|  | < 2 | 280,462 | 1.03 | 0.97 | 1.09 |
|  | < 1.5 | 134,046 | 1.03 | 0.92 | 1.15 |
|  | < 1 | 33,435 | 1.32 | 0.93 | 1.86 |
|  | < 0.5 | 4,730 | 4.02 | 0.12 | 129.09 |
| O3 |  |  |  |  |
|  | Full dataset | 307,550 | 0.95 | 0.89 | 1.02 |
|  | < 120 | 307,550 | 0.95 | 0.89 | 1.02 |
|  | < 100 | 302,968 | 0.95 | 0.89 | 1.02 |
|  | < 80 | 93,882 | 1.08 | 0.96 | 1.21 |
|  | < 60 | 1,664 | 0.38 | 0.06 | 2.34 |
| HR, hazard ratio; CI, confidence intervalaAdjusted for study (strata), age, sex, year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status and 2001 mean income at the neighborhood levelbConcentrations above a certain value were excluded and compared to analyses of the full exposure range |

|  |
| --- |
| **Table S2. Threshold analyses of PM2.5 and lung cancer** |
|  |  | **All lung cancersa****(N=3,956)** |
| **Thresholdb** |  | **AICc** |
| No threshold |  | 69478.2 |
| 5 µg/m3 |  | 69478.3 |
| 7.5 µg/m3 |  | 69478.6 |
| 10 µg/m3 |  | 69479.2 |
| No. obs = 307,550HR, hazard ratio; CI, confidence intervalaAdjusted for study (strata), age, sex, year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status and 2001 mean income at the neighborhood levelbThe pollutant variable was set to 0 for concentrations below a certain valuecLower AIC represent improved prediction of lung cancer |

|  |
| --- |
| **Table S3.** Spearman correlations per (sub) cohort between NO2, PM2.5, BC, and O3 (warm season) among participants with full information in the main model (N=307,550) |
|  | **PM2.5** | **BC** | **O3** |   |   | **PM2.5** | **BC** | **O3** |   |   | **PM2.5** | **BC** | **O3** |
| **CEANS-SALT** |  | **DCH** |  | **EPIC-Prospect** |
| **NO2** | 0.63 | 0.81 | -0.69 |   | **NO2** | 0.74 | 0.91 | -0.59 |   | **NO2** | 0.42 | 0.89 | -0.85 |
| **PM2.5** | - | 0.52 | -0.46 |  | **PM2.5** | - | 0.70 | -0.52 |  | **PM2.5** | **-** | 0.39 | -0.43 |
| **BC** | - | - | -0.69 |  | **BC** | - | - | -0.58 |  | **BC** | - | - | -0.81 |
| **CEANS-SDPP** |  | **DNC-1993** |   | **HNR** |
| **NO2** | 0.60 | 0.64 | -0.67 |  | **NO2** | 0.58 | 0.90 | -0.35 |  | **NO2** | 0.63 | 0.83 | -0.77 |
| **PM2.5** | - | 0.58 | -0.14 |  | **PM2.5** | - | 0.69 | -0.27 |  | **PM2.5** | **-** | 0.60 | -0.73 |
| **BC** | - | - | -0.27 |  | **BC** | - | - | -0.38 |   | **BC** | - | - | -0.73 |
| **CEANS-SIXTY** |   | **DNC-1999** |  | **E3N** |
| **NO2** | 0.62 | 0.81 | -0.67 |  | **NO2** | 0.57 | 0.93 | -0.16 |  | **NO2** | 0.77 | 0.90 | -0.52 |
| **PM2.5** | - | 0.53 | -0.42 |  | **PM2.5** | - | 0.63 | -0.10 |  | **PM2.5** | **-** | 0.68 | -0.45 |
| **BC** | - | - | -0.66 |  | **BC** | - | - | -0.16 |  | **BC** | - | - | -0.37 |
| **CEANS-SNACK** |   | **EPIC-Morgen** |   | **VHM PP** |  |  |  |
| **NO2** | 0.76 | 0.38 | -0.71 |  | **NO2** | 0.20 | 0.82 | -0.77 |  | **NO2** | 0.62 | 0.90 | -0.82 |
| **PM2.5** | - | 0.28 | -0.63 |  | **PM2.5** | - | 0.42 |  0.18 |  | **PM2.5** | **-** | 0.74 | -0.71 |
| **BC** | - | - | -0.60 |   | **BC** | - | - | -0.50 |   | **BC** | - | - | -0.87 |

|  |
| --- |
| **Table S4.** Air pollutant exposure extrapolated back to the time of enrolment and time-varying exposure analysis using residential history between enrolment and end-of follow-up based on the DEHM. Back-extrapolation was performed using the absolute difference and the ratio between the baseline and 2010 periods. |
|  | **Model 3a**Back-extrapolated to enrolmentN=307,526 | **Model 3a**Main model 2010 exposureN=307,526 |
| **All lung cancers**  | **Mean (SD)** | **HR** | **95% CI** | **Mean (SD)** | **HR** | **95% CI** |
|  | NO2 *difference* | 28.7 (8.9) | 1.02 | 0.97 | 1.07 | 25.0 (8) | 1.02 | 0.97 | 1.07 |
|  | NO2 *ratio* | 35.9 (11.1) | 1.02 | 0.98 | 1.05 |  |  |  |  |
|  | PM2.5 *difference* | 28.8 (7.7) | 1.06 | 1.01 | 1.12 | 15.0 (3.2) | 1.13 | 1.05 | 1.23 |
|  | PM2.5 *ratio* | 28.3 (8.3) | 1.06 | 1.02 | 1.10 |  |  |  |  |
|  | BC *difference* | 1.6 (0.5) | 1.01 | 0.96 | 1.06 | 1.5 (0.4) | 1.02 | 0.97 | 1.09 |
|  | BC *ratio* | 1.9 (0.7) | 1.01 | 0.97 | 1.05 |  |  |  |  |
|  | O3 *difference* | 89.6 (12.0) | 0.95 | 0.89 | 1.01 | 85.4 (9.0) | 0.95 | 0.89 | 1.02 |
|  | O3 *ratio* | 89.7 (12.1) | 0.95 | 0.89 | 1.01 |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  | **Model 3b**Time-varying exposure analysisN=180,921 | **Model 3a**Main model 2010 exposureN=180,921 |
|  | NO2 *difference* | - | 1.04 | 0.99 | 1.11 | - | 1.04 | 0.97 | 1.10 |
|  | NO2 *ratio* | - | 1.04 | 0.99 | 1.10 |  |  |  |  |
|  | PM2.5 *difference* | - | 1.15 | 1.04 | 1.28 | - | 1.18 | 1.06 | 1.31 |
|  | PM2.5 *ratio* | - | 1.13 | 1.03 | 1.24 |  |  |  |  |
|  | BC *difference* | - | 1.03 | 0.97 | 1.10 | - | 1.02 | 0.96 | 1.09 |
|  | BC *ratio* | - | 1.03 | 0.97 | 1.10 |  |  |  |  |
|  | O3 *difference* | - | 0.97 | 0.93 | 1.01 | - | 0.94 | 0.87 | 1.02 |
|  | O3 *ratio* | - | 0.97 | 0.93 | 1.01 |  |  |  |  |
| HR, hazard ratio; CI, confidence interval; O3, ozone warm seasonaAdjusted for study (strata), age, sex (strata), year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status, and 2001 mean income at the neighborhood levelbAdjusted for study (strata), age, sex (strata), year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status, 2001 mean income at the neighborhood level, and calendar time (1-year strata) |

|  |
| --- |
| **Table S5.** Air pollutants exposure from either ELAPSE or ESCAPE and risk of lung cancer in the subset of the pooled cohort with available information from both exposure models. |
|  |  |  |  | **Model 3a**ELAPSE exposureN=197,695b |  | **Model 3a**ESCAPE exposureN=197,695b |
| **All lung cancers** (N=2,701) | *Increment* |  | **HR** | **95% CI** |  | **HR** | **95% CI** |
|  | NO2 | 10 µg/m3 |  | 1.04 | 0.96 | 1.13 |  | 1.00 | 0.94 | 1.06 |
|  | PM2.5 | 5 µg/m3  |  | 1.20 | 1.07 | 1.35 |  | 1.34 | 1.11 | 1.61 |
|  | BC | 0.5 10-5m-1 |  | 1.03 | 0.95 | 1.11 |  | 1.02 | 0.94 | 1.11 |
| HR, hazard ratio; CI, confidence intervalaAdjusted for study (strata), age, sex (strata), year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status, and 2001 mean income at the neighborhood levelbExcluding the DNC, E3N and parts of the DCH and VHM&PP cohorts |

|  |
| --- |
| **Table S6.** Sensitivity analyses (PM2.5) including additional confounders in the subset of the pooled cohort with the available information. |
|  | **Model 3a** |  | **Model 3a****+ additional covariate adjustment** |
| **All lung cancers**  | ***N*** | **HR** | **95% CI** |  | **HR** | **95% CI** |
| Educational levelb | 166,138 | 1.01 | 0.89 | 1.15 |  | 1.02 | 0.89 | 1.16 |
| Neighborhood unemployment ratec | 288,586 | 1.13 | 1.04 | 1.22 |  | 1.13 | 1.04 | 1.22 |
| Neighborhood ethnicityd | 284,975 | 1.13 | 1.04 | 1.22 |  | 1.15 | 1.06 | 1.24 |
| Neighborhood low educational level ratee | 272,496 | 1.12 | 1.03 | 1.21 |  | 1.12 | 1.03 | 1.21 |
| Neighborhood high educational level ratee | 272,496 | 1.12 | 1.03 | 1.21 |  | 1.12 | 1.04 | 1.22 |
| Smoking intensity in ex-smokersb | 167,278 | 1.01 | 0.88 | 1.14 |  | 0.99 | 0.87 | 1.13 |
| Smoking intensity in ex-smokersf | 130,681 | 1.08 | 0.92 | 1.26 |  | 1.07 | 0.92 | 1.25 |
| Occupational classg | 169,003 | 1.19 | 1.07 | 1.33 |  | 1.19 | 1.07 | 1.32 |
| HR, hazard ratio; CI, confidence intervalaAdjusted for study (strata), age, sex (strata), year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status, and 2001 mean income at the neighborhood levelbExcluding the VHM&PP cohortcExcluding the CEANS cohortdExcluding the CEANS and the HNR cohortseExcluding the CEANS, the HNR, and the EPIC-NL cohortsfExcluding the VHM&PP and E3N with suboptimal smoking informationgExcluding DCH, E3N, HNR, EPIC-NL |

|  |
| --- |
| **Table S7.** Sensitivity analyses (PM2.5) excluding specific cohorts with suboptimal case ascertainment. |
|  | **Model 3a** |
| **All lung cancers**  | ***N*** | **HR** | **95% CI** |
| All cohorts | 307,550 | 1.13 | 1.05 | 1.23 |
| Excluding HNR | 303,939 | 1.13 | 1.04 | 1.22 |
| Excluding E3N | 270,953 | 1.17 | 1.08 | 1.28 |
|  |  |  |  |  |
| HR, hazard ratio; CI, confidence intervalaAdjusted for study (strata), age, sex (strata), year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status, and 2001 mean income at the neighborhood level. |

|  |
| --- |
| **Table S8.** Effect modification by smoking on the relation between PM2.5 and lung cancer incidence (N=307,550). |
|  | **Model 3a** |
| **Smoking status**  |  | **HR** | **95% CI** |
| PM2.5 (Never smokers) |  | 1.15 | 1.01 | 1.31 |
| PM2.5 (Previous smokers) |  | 1.02 | 0.88 | 1.19 |
| PM2.5 (Current smokers) |  | 1.15 | 1.08 | 1.26 |
|  |  |  |  |  |
| HR, hazard ratio; CI, confidence intervalaAdjusted for study (strata), age, sex (strata), year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status, and 2001 mean income at the neighborhood level. |

|  |
| --- |
| **Table S9.** Previous studies on PM2.5 and lung cancer |
| **Reference** | **Study** | **Area** | **Lung cancer measure** | **Unit** | **Linear effect estimate****HR (95% CI)** | **Shape of the CRF** |
| Bai et al. (2019) | ONPHEC | Canada | Incidence | per 5.3 μg/m³ | 1.02 (1.01–1.05) | Sub-linear, threshold at 10 µg/m3 |
| Raaschou-Nielsen et al. (2013) | ESCAPE Cohort | Europe | Incidence | per 5 μg/m³ | 1.18 (0.96–1.46) | Linear or Supra-linear |
| Burnett et al. (2018) | Meta-analysis | Worldwide | Mortality |  - | - | Linear |
| Bowe et al. (2019) | US veterans Cohort | US | Mortality | per 10 μg/m³ | 1.11 (1.09–1.14) | Supra-linear |
| Pinault et al. (2017) | CANCHEC | Canada | Mortality | per 10 μg/m³ | 1.16 (1.07–1.25) | Sub-linear, threshold at 5 µg/m3 |
| Yin et al. (2017) | Chinese cohort of men | China | Mortality | per 10 μg/m³ | 1.12 (1.07–1.14) | Sub-linear |
| Hamra et al. (2014) | Meta-analysis | North America + Europe | Incidence + mortality | per 10 μg/m³ | 1.09 (1.04–1.14) | - |
| Hamra et al. (2014) | Meta-analysis | Europe only | Incidence + mortality | per 10 μg/m³ | 1.03 (0.89–1.20) | - |
| Abbreviations: CRF, concentration response function, ONPHEC, the Ontario Population Health Environment Cohort; ESCAPE, the European Study of Cohorts for Air Pollution Effects; CANCHEC, the Canadian Census Health and Environment Cohort |

**Figure S1.** Temporal trends in PM2.5, NO2, BC and O3 concentrations based on the DEHM for different regions in Europe



**Figure S2**. Box plots of exposures by individual (sub-) cohort study. Red dotted lines represent the applied subset values. The boundary of the box closest to zero indicates P25; furthest from zero – P75; bold line in the middle of the box – P50; whiskers indicate P5 and P95. O3, ozone in warm season.









|  |
| --- |
| **Figure S3.** Natural spline models of air pollutants and lung cancer (3 df). |
|  |  |
|  |  |

|  |
| --- |
| **Figure S4.** Two- and three-pollutant models of PM2.5 and co-pollutants and lung cancer |
|  |  |
| Hazard ratios with confidence intervals adjusted for study (strata), age, sex, year of baseline visit, smoking status, duration, intensity, intensity², BMI, marital status, employment status and 2001 mean income at the neighborhood levelO3w, ozone in the warm seasonNote: the correlation with NO2 and BC was high in some (sub) cohorts |

**Figure S5. Scatterplot and Spearman correlations (*r*) of ELAPSE vs. ESCAPE exposure by cohort**

****



****

**Figure S6. Comparing different approaches to account for differences between the (sub) cohorts**

****

**Figure S7. Hazard ratios for associations between air pollution and lung cancer in the individual cohorts**

****