Supplementary Information

Long-term exposure to air pollution and liver cancer incidence in six European cohorts

Rina So, Jie Chen, Amar J. Mehta, Shuo Liu, Maciej Strak, Kathrin Wolf, Ulla A. Hvidtfeldt, Sophia Rodopoulou, Massimo Stafoggia, Jochem O. Klompmaker, Evangelia Samoli, Ole Raaschou-Nielsen, Richard Atkinson, Mariska Bauwelinck, Tom Bellander, Marie-Christine Boutron-Ruault, Jørgen Brandt, Bert Brunekreef, Giulia Cesaroni, Hans Concin, Francesco Forastiere, Carla H. van Gils, John Gulliver, Ole Hertel, Barbara Hoffmann, Kees de Hoogh, Nicole Janssen, Youn-hee Lim, Rudi Westendorp, Jeanette T. Jørgensen, Klea Katsouyanni, Matthias Ketzel, Anton Lager, Alois Lang, Petter L. Ljungman, Patrik K.E. Magnusson, Gabriele Nagel, Mette K. Simonsen, Göran Pershagen, Raphael S. Peter, Annette Peters, Matteo Renzi, Debora Rizzuto, Torben Sigsgaard, Danielle Vienneau, Gudrun Weinmayr, Gianluca Severi, Daniela Fecht, Anne Tjønneland, Karin Leander, Gerard Hoek, Zorana J. Andersen

Table of contents

Characteristics of each of the six cohorts and their participants

Supplementary tables

Supplementary figures

Characteristics of each of the six cohorts and their participants

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 Screening Across the Lifespan Twin Study (SALT) sampled 7,043 individuals from the Swedish Twin Register born 1958 and earlier, who lived in Stockholm County[1]. The Stockholm Diabetes Preventive Program (SDPP) is a population-based prospective study of 7,949 subjects aged 35–54 years[2]. The Stockholm cohort of 60-year-olds (Sixty) sub-cohort 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[3]. Lastly, The Swedish National Study on Aging and Care in Kungsholmen (SNAC-K) consists of randomly sampled individuals 60 years old and over from a central area (Kungsholmen) in Stockholm[4].

Characteristics	CEANS Stockholm, Sweden				
	Total	SALT	SDPP	Sixty	SNAC-K
Enrolled, N	22,587	7,043	7,949	4,232	3,363
Pooled, N	21,987	6,724	7,835	4,180	3,248
Exclusions[a], N	1,933	626	432	324	551
Missing on covariates[b], N	423	141	34	151	97
Included, N	19,631	5,957	7,369	3,705	2,600
Baseline period, year	1992–2004	1998–2002	1992–1998	1997–1999	2001–2004
End of follow-up	31-12-2011	31-12-2011	31-12-2011	31-12-2011	31-12-2011
Person-year	234,274.4	58944.6	113079.2	44273.8	17976.9
Follow-up time, year	11.9	9.9	15.3	11.9	6.9
Liver cancer, N	18	3	4	10	1
Baseline age, year (Mean ± SD)	56.3 ± 11.7	57.9 ± 10.8	47 ± 4.9	60 ± 0	73.7 ± 11.1
Age categories, N (%)					
< 65 years old	16,147 (82.3)	4,452 (74.7)	7,369 (100)	3,705 (100)	621 (23.9)
≥ 65 years old	3,484 (17.7)	1,505 (25.3)	0 (0)	0 (0)	1,979 (76.1)
Women, N (%)	11,059 (56.3)	3,181 (53.4)	4,369 (59.3)	1,848 (49.9)	1,661 (63.9)
Smoking status, N (%)					
Never smoker	8,137 (41.4)	2,636 (44.3)	2,773 (37.6)	1,498 (40.4)	1,230 (47.3)
Previous smoker	7,198 (36.7)	2,109 (35.4)	2,681 (36.4)	1,424 (38.4)	984 (37.8)
Current smoker	4,296 (21.9)	1,212 (20.3)	1,915 (26.0)	783 (21.1)	386 (14.8)
Unemployed, N (%)	6,016 (30.6)	2,125 (35.7)	674 (9.1)	1,194 (32.2)	2,023 (77.8)
Intake of alcohol, N (%)					
Low (<4 g/day)	2,352 (21)	255 (5.9)	2,097 (30.5)	-	-
Medium (4-15 g/day)	5,870 (52.5)	2,611 (60.6)	3,259 (47.4)	-	-
High (15> g/day)	2,968 (26.5)	1,442 (33.5)	1,526 (22.2)	-	-
Education levels, N (%)					
Low level	5,971 (30.8)	1,575 (26.6)	2,244 (31.3)	1,459 (39.8)	693 (26.7)
Medium level	7,174 (37.1)	2,185 (36.9)	2,771 (38.6)	1,185 (32.3)	1,033 (39.8)
High level	6,213 (32.1)	2,160 (36.5)	2,162 (30.1)	1,022 (27.9)	869 (33.5)
Mean income at neighborhood level in 2001, (Mean ± SD)	25.3 ± 5.6	25.3 ± 6.6	24.3 ± 4.2	24.7 ± 6.8	28.6 ± 2.2
NO_2, $\mu g/m^3$ (Mean ± SD)	19.8 ± 6.7	21.3 ± 6.2	15.4 ± 4.3	20.7 ± 6.1	27.4 ± 5.1
$PM_{2.5}$, $\mu g/m^3$ (Mean ± SD)	8.1 ± 1	8.4 ± 0.9	7.6 ± 0.9	8.3 ± 0.9	8.6 ± 0.8
BC, 10^{-5}/m (Mean ± SD)	0.8 ± 0.3	0.8 ± 0.3	0.6 ± 0.2	0.8 ± 0.2	1.1 ± 0.1
O_3 , $\mu g/m^3$ (Mean ± SD)	76.8 ± 2.5	76.6 ± 2.7	77.6 ± 1.9	76.7 ± 2.5	75.1 ± 2.7

Abbreviation: N, number; SD, standard deviation; NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm; BC, black carbon; O_3, ozone.

[a] Due to cancer before baseline or missing information on exposure data, the prevalent cancer status, or the date of start or end of follow-up

[b] Covariates which used the main model: age, sex, calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level
[c] EUR per 1,000. Neighborhood defined as municipality

DCH (Diet, Cancer, and Health)
Participants were recruited among persons aged 50 years and older from the areas of greater Copenhagen and Aarhus, Denmark[5].

Characteristics	DCH Copenhagen and Aarhus, Denmark
Enrolled, N	57,053
Pooled, N	56,308
Exclusions[a], N	907
Missing on covariates[b], N	1,259
Included, N	54,142
Baseline period, year	1993–1997
End of follow-up	31-12-2015
Person-year	912,625
Follow-up time, year	16.9
Liver cancer, N	136
Baseline age, year (Mean ± SD)	56.7 ± 4.4
Age categories, N (%)	
< 65 years old	53,447 (98.7)
≥ 65 years old	695 (1.3)
Women, N (%)	28,302 (52.3)
Smoking status, N (%)	
Never smoker	19,034 (35.2)
Previous smoker	15,567 (28.8)
Current smoker	19,541 (36.1)
Unemployed, N (%)	11,819 (21.8)
Intake of alcohol, N (%)	
Low (<4 g/day)	9,897 (18.7)
Medium (4-15 g/day)	18,671 (35.3)
High (15> g/day)	24,264 (45.9)
Education levels, N (%)	
Low level	8,053 (14.9)
Medium level	33,997 (63)
High level	11,924 (22.1)
Mean income at neighborhood level in 2001[c] (Mean ± SD)	20.1 ± 3.4
NO_2, μg/m^3 (Mean ± SD)	28.1 ± 6.9
$PM_{2.5}$, μg/m^3 (Mean ± SD)	13.2 ± 1.4
BC, 10^{-5}/m (Mean ± SD)	1.3 ± 0.3
O_3 , μg/m^3 (Mean ± SD)	77.4 ± 5.1

Abbreviation: N, number; SD, standard deviation; NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm; BC, black carbon; O_3, ozone.
[a] Due to cancer before baseline or missing information on exposure data, the prevalent cancer status, or the date of start or end of follow-up
[b] Covariates which used the main model: age, sex, calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level
[c] EUR per 1,000. Neighborhood defined as municipality

DNC (Danish Nurse Cohort)

The cohort[6] 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).

Characteristics	DNC Denmark		
	Total	DNC-1993	DNC-1999
Enrolled, N	28,731	19,898	8,833
Pooled, N	28,433	19,664	8,769
Exclusions[a], N	2,492	1,742	750
Missing on covariates[b], N	1,661	1,360	301
Included, N	24,280	16,562	7,718
Baseline period, year	1993/1999	1,993	1,999
End of follow-up	31-12-2012	31-12-2012	31-12-2012
Person-year	377,956.8	277,903.5	100,053.3
Follow-up time, year	15.6	16.8	13
Liver cancer, N	15	14	1
Baseline age, year (Mean ± SD)	53.7 ± 8.4	56.4 ± 8.6	47.9 ± 4.3
Age categories, N (%)			
< 65 years old	21,460 (88.4)	13,869 (83.7)	7,591 (98.4)
≥ 65 years old	2,820 (11.6)	2,693 (16.3)	127 (1.6)
Women, N (%)	24,280 (100)	16,562 (100)	7,718 (100)
Smoking status, N (%)			
Never smoker	8,449 (34.8)	5,464 (33)	2,985 (38.7)
Previous smoker	7,414 (30.5)	4,848 (29.3)	2,566 (33.2)
Current smoker	8,417 (34.7)	6,250 (37.7)	2,167 (28.1)
Unemployed, N (%)	5,486 (22.6)	5,085 (30.7)	401 (5.2)
Intake of alcohol, N (%)			
Low (<4 g/day)	2,732 (13.8)	1,925 (14.8)	807 (11.9)
Medium (4-15 g/day)	7,225 (36.5)	4,722 (36.3)	2,503 (36.8)
High (15> g/day)	9,864 (49.8)	6,366 (48.9)	3,498 (51.4)
Education levels, N (%)			
Low level	0 (0)	0 (0)	0 (0)
Medium level	0 (0)	0 (0)	0 (0)
High level	24,280 (100)	16,562 (100)	7,718 (100)
Mean income at neighborhood level in 2001[c] (Mean ± SD)	19.2 ± 2.5	19.2 ± 2.6	19 ± 2.4
NO_2, $\mu g/m^3$ (Mean ± SD)	23.1 ± 8.3	21.8 ± 8	25.8 ± 8.5
$PM_{2.5}$, $\mu g/m^3$ (Mean ± SD)	13.1 ± 1.6	12.7 ± 1.5	13.8 ± 1.5
BC, $10^{-5}/m$ (Mean ± SD)	1.2 ± 0.4	1.1 ± 0.4	1.3 ± 0.4
O_3 , $\mu g/m^3$ (Mean ± SD)	80.5 ± 3.9	80.4 ± 4	80.6 ± 3.8

Abbreviation: N, number; SD, standard deviation; NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm; BC, black carbon; O_3, ozone.

[a] Due to cancer before baseline or missing information on exposure data, the prevalent cancer status, or the date of start or end of follow-up

[b]Covariates which used the main model: age, sex, calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level

[c]EUR per 1,000, Neighborhood defined as municipality

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

The cohort[7] was selected among French women aged 40 to 65 years who were insured through a national health system that primarily covered teachers.

Characteristics	E3N France
Enrolled, N	98,995
Pooled, N	53,521
Exclusions[a], N	3,781
Missing on covariates[b], N	218
Included, N	49,522
Baseline period, year	1989–1991
End of follow-up	2/3/2014
Person-year	785,460.3
Follow-up time, year	15.9
Liver cancer, N	33
Baseline age, year (Mean ± SD)	52.8 ± 6.7
Age categories, N (%)	
< 65 years old	46,601 (94.1)
≥ 65 years old	2,921 (5.9)
Women, N (%)	49,522 (100)
Smoking status, N (%)	
Never smoker	26,952 (54.4)
Previous smoker	15,988 (32.3)
Current smoker	6,582 (13.3)
Unemployed, N (%)	15,255 (30.8)
Intake of alcohol, N (%)	
Low (<4 g/day)	13,118 (30.3)
Medium (4-15 g/day)	16,620 (38.4)
High (15> g/day)	13,554 (31.3)
Education levels, N (%)	
Low level	1,841 (3.9)
Medium level	3,851 (8.1)
High level	41,943 (88.1)
Mean income at neighborhood level in 2001[c] (Mean ± SD)	11.2 ± 3
NO_2, µg/m^3 (Mean ± SD)	26.5 ± 9.8
$PM_{2.5}$, µg/m^3 (Mean ± SD)	17.1 ± 2.9
BC, 10^{-5}/m (Mean ± SD)	1.8 ± 0.5
O_3 , µg/m^3 (Mean ± SD)	87.6 ± 8

Abbreviation: N, number; SD, standard deviation; NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 µm; BC, black carbon; O_3, ozone.

[a] Due to cancer before baseline or missing information on exposure data, the prevalent cancer status, or the date of start or end of follow-up

[b] Covariates which used the main model: age, sex, calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level

[c] EUR per 1,000. Neighborhood defined as IRIS – a small administrative unit of a city

EPIC-NL (European Prospective Investigation into Cancer and Nutrition, the Netherlands)
The EPIC-NL[8] 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.

Characteristics	EPIC-NL Netherland		
	Total	MORGEN	Prospect
Enrolled, N	40,011	22,654	17,357
Pooled, N	36,905	20,711	16,194
Exclusions[a], N	1,755	615	1,140
Missing on covariates[b], N	1,657	827	830
Included, N	33,493	19,269	14,224
Baseline period, year	1993–1997	1993–1997	1993–1997
End of follow-up	31-12-2012	31-12-2012	31-12-2012
Person-year	539,231.9	316,539.7	222,692.3
Follow-up time, year	16.1	16.4	15.7
Liver cancer, N	18	8	10
Baseline age, year (Mean ± SD)	49.0 ± 11.9	42.6 ± 11.2	57.6 ± 6
Age categories, N (%)			
< 65 years old	30,828 (92)	19,153 (99.4)	11,675 (82.1)
≥ 65 years old	2,665 (8)	116 (0.6)	2,549 (17.9)
Women, N (%)	24,631 (73.5)	10,407 (54)	14,224 (100)
Smoking status, N (%)			
Never smoker	12,705 (37.9)	6,638 (34.4)	6,067 (42.7)
Previous smoker	10,506 (31.4)	5,581 (29.0)	4,925 (34.6)
Current smoker	10,282 (30.7)	7,050 (36.6)	3,232 (22.7)
Unemployed, N (%)	12,740 (38.0)	5,842 (30.3)	6,898 (48.5)
Intake of alcohol, N (%)			
Low (<4 g/day)	7,314 (28.2)	3,950 (25.6)	3,364 (32)
Medium (4-15 g/day)	9,661 (37.3)	5,728 (37.2)	3,933 (37.4)
High (15> g/day)	8,943 (34.5)	5,736 (37.2)	3,207 (30.5)
Education levels, N (%)			
Low level	5,283 (15.8)	2,132 (11.1)	3,151 (22.2)
Medium level	25,958 (77.8)	15,427 (80.4)	10,531 (74.1)
High level	2,144 (6.4)	1,620 (8.4)	524 (3.7)
Mean income at neighborhood level in 2001[c] (Mean ± SD)	12.6 ± 1.6	12.2 ± 1.6	13.1 ± 1.4
NO_2, μg/m^3 (Mean ± SD)	35.1 ± 5.8	34.5 ± 6	35.9 ± 5.4
$PM_{2.5}$, μg/m^3 (Mean ± SD)	17.5 ± 1.1	18.0 ± 1	16.9 ± 0.8
BC, 10^{-5}/m (Mean ± SD)	1.7 ± 0.3	1.7 ± 0.3	1.7 ± 0.3
O_3 , μg/m^3 (Mean ± SD)	73.1 ± 6.1	73.5 ± 7.7	72.7 ± 2.7

Abbreviation: N, number; SD, standard deviation; NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm; BC, black carbon; O_3, ozone.
[a] Due to cancer before baseline or missing information on exposure data, the prevalent cancer status, or the date of start or end of follow-up
[b]Covariates which used the main model: age, sex, calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level
[c]EUR per 1,000. Neighborhood defined as a neighborhood of a larger city

VHM&PP (Vorarlberg Health Monitoring and Prevention Programme)
The VHM&PP[9] is a population-based cohort recruited among all adults of the province of Vorarlberg, Austria.

Characteristics	VHM&PP Vorarlberg, Austria
Enrolled, N	181,350
Pooled, N	170,250
Exclusions[a], N	4,892
Missing on covariates[b], N	16,362
Included, N	148,996
Baseline period, year	1985–2005
End of follow-up	31-12-2014
Person-year	3,121,637.2
Follow-up time, year	21
Liver cancer, N	292
Baseline age, year (Mean ± SD)	41.5 ± 14.9
Age categories, N (%)	
< 65 years old	137,261 (92.1)
≥ 65 years old	11,735 (7.9)
Women, N (%)	82,498 (55.4)
Smoking status, N (%)	
Never smoker	105,426 (70.8)
Previous smoker	8,792 (5.9)
Current smoker	34,778 (23.3)
Unemployed, N (%)	42,909 (28.8)
Intake of alcohol, N (%)	
Low (<4 g/day)	-
Medium (4-15 g/day)	-
High (15> g/day)	-
Education levels, N (%)	
Low level	-
Medium level	-
High level	-
Mean income at neighborhood level in 2001[c] (Mean ± SD)	22.9 ± 1.7
NO_2, μg/m^3 (Mean ± SD)	22 ± 5.3
$PM_{2.5}$, μg/m^3 (Mean ± SD)	15.7 ± 2.6
BC, 10^{-5}/m (Mean ± SD)	1.6 ± 0.3
O_3 , μg/m^3 (Mean ± SD)	92.6 ± 3.6

Abbreviation: N, number; SD, standard deviation; NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm; BC, black carbon; O_3, ozone.

[a] Due to cancer before baseline or missing information on exposure data, the prevalent cancer status, or the date of start or end of follow-up

[b] Covariates which used the main model: age, sex, calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level

[c] EUR per 1,000. Neighborhood defined as municipality

Table S1. Pearson correlations per each cohort between annual mean concentration to NO_2, $PM_{2.5}$, BC, and O_3 among participants with full information in the main model (N=330,064).

	NO_2	$PM_{2.5}$	BC		NO_2	$PM_{2.5}$	BC		NO_2	$PM_{2.5}$	BC
All (Average)				CEANS-SNAC-K				E3N			
$PM_{2.5}$	0.62			$PM_{2.5}$	0.76			$PM_{2.5}$	0.82		
BC	0.83	0.57		BC	0.44	0.30		BC	0.92	0.75	
O_{3w}	-0.64	-0.38	-0.58	O_{3w}	-0.66	-0.50	-0.75	O_{3w}	-0.51	-0.49	-0.38
CEANS-SALT				DCH				EPIC- MORGEN			
$PM_{2.5}$	0.67			$PM_{2.5}$	0.73			$PM_{2.5}$	0.21		
BC	0.84	0.56		BC	0.92	0.68		BC	0.84	0.41	
O_{3w}	-0.74	-0.48	-0.76	O_{3w}	-0.62	-0.60	-0.56	O_{3w}	-0.78	0.15	-0.55
CEANS-SDPP				DNC-1993				EPIC- PROSPECT			
$PM_{2.5}$	0.61			$PM_{2.5}$	0.64			$PM_{2.5}$	0.48		
BC	0.67	0.49		BC	0.92	0.70		BC	0.91	0.41	
O_{3w}	-0.69	-0.18	-0.33	O_{3w}	-0.41	-0.32	-0.42	O_{3w}	-0.86	-0.43	-0.84
CEANS-Sixty				DNC-1999				VHM&PP			
$PM_{2.5}$	0.69			$PM_{2.5}$	0.61			$PM_{2.5}$	0.65		
BC	0.84	0.59		BC	0.93	0.64		BC	0.91	0.76	
O_{3w}	-0.72	-0.45	-0.71	O_{3w}	-0.22	-0.16	-0.21	O_{3w}	-0.83	-0.69	-0.88

Abbreviation: The 'Cardiovascular Effects of Air Pollution and Noise in Stockholm' [CEANS] from Stockholm county of Sweden, which is comprised of the four sub-cohorts: Swedish National Study on Aging and Care in Kungsholmen [SNAC-K], Stockholm Screening Across the Lifespan Twin study [SALT], Stockholm cohort of 60-year-olds [Sixty], and Stockholm Diabetes Prevention Program [SDPP]; the 'Diet, Cancer and Health cohort' [DCH] from Copenhagen and Aarhus of Denmark; the 'Danish Nurse Cohort' [DNC] from entire Denmark, which included two sub-cohorts from recruitment rounds in 1993 and 1999; the 'Dutch European Investigation into Cancer and Nutrition' [EPIC-NL] from four cities in the Netherland, consisting of 'EPIC-Monitoring Project on Risk Factors' [EPIC-MORGEN] and 'EPIC-Chronic Diseases in the Netherlands' [EPIC-PROSPECT]; the 'Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale' [E3N] from entire France; and 6) the 'Vorarlberg Health Monitoring and Prevention Programme' [VHM&PP] from Vorarlberg, Austria; NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm; BC, black carbon; O_3, ozone.

Table S2. Effect modifications of associations between long-term exposure to air pollution and liver cancer incidence by age, alcohol intake, and smoking status.

Modifier variable	N	Case	Hazard ratio (95% CI)[a]				P value[b]
			NO_2	$PM_{2.5}$	BC	O_{3w}	
Age, years							NO_2: 0.01; $PM_{2.5}$: 0.29; BC: 0.06; O_{3w}: 0.12
< 65	305,744	456	1.12 (0.96 to 1.30)	1.08 (0.88 to 1.33)	1.11 (0.96 to 1.29)	0.72 (0.59 to 0.87)	
≥ 65	24,320	56	1.85 (1.30 to 2.64)	1.44 (0.87 to 2.38)	1.66 (1.13 to 2.42)	0.51 (0.33 to 0.78)	
Alcohol intake[c]							NO_2: 0.62; $PM_{2.5}$: 0.09; BC: 0.39; O_{3w}: 0.38
Low (<4 g/day)	35,413	33	1.01 (0.66 to 1.53)	0.54 (0.25 to 1.19)	0.93 (0.62 to 1.40)	0.87 (0.54 to 1.41)	
Medium (4-15 g/day)	58,047	51	1.29 (0.92 to 1.81)	1.35 (0.76 to 2.40)	1.30 (0.95 to 1.78)	0.81 (0.53 to 1.25)	
High (15> g/day)	59,493	100	1.24 (0.95 to 1.62)	1.07 (0.64 to 1.79)	1.12 (0.86 to 1.44)	0.62 (0.45 to 0.87)	
Smoking status							NO_2: 0.07; $PM_{2.5}$: 0.52; BC: 0.47; O_{3w}: <.01
Never	180,703	242	1.19 (0.98 to 1.46)	1.08 (0.85 to 1.37)	1.24 (1.02 to 1.51)	0.77 (0.61 to 0.97)	
Ex-smoker	65,465	104	0.93 (0.71 to 1.21)	1.01 (0.70 to 1.46)	1.04 (0.81 to 1.33)	0.88 (0.67 to 1.15)	
Current Smoker	83,896	166	1.34 (1.08 to 1.67)	1.27 (0.94 to 1.72)	1.13 (0.91 to 1.41)	0.57 (0.46 to 0.71)	

Results are presented as hazard ratio and 95% confidence interval [HR (95% CI)] for the following increments: 5 $\mu g/m^3$ for $PM_{2.5}$, 10 $\mu g/m^3$ for NO_2, 0.5 $10^{-5}/m$ for BC and 10 $\mu g/m^3$ for O_3.

[a] In addition to the adjustments in the main model (age (time scale), sex (strata), sub-cohort (strata), calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level in 2001), we included an interaction term of the modifier and the exposure in the model.

[b] From the likelihood ratio test between models with and without the interaction term of the modifier and the exposure.

[c] n=153,043 (The entire participants in the VHM&PP, CEANS-Sixty, and CEANS-SNAC-K cohort dropped out from this analysis because of missing information on alcohol intake)

Table S3. Associations between time-varying estimates of long-term exposure to air pollution and liver cancer incidence in four pooled cohorts with available information based on the main model (DNC and E3N were excluded; N=188,453, Cases=367).

Pollutants	Main model[a]	Time-varying analyses[a] with further adjustment of the below strata term.			
	Reduced dataset (N=188,453)	Strata per year of follow-up time		Strata per 5-year of follow-up time	
		Ratio method	Difference method	Ratio method	Difference method
NO_2	1.14 (0.96 to 1.36)	1.15 (0.98 to 1.35)	1.17 (0.99 to 1.38)	1.16 (1.00 to 1.36)	1.18 (1.00 to 1.40)
$PM_{2.5}$	1.12 (0.88 to 1.43)	1.12 (0.90 to 1.38)	1.14 (0.89 to 1.45)	1.07 (0.90 to 1.28)	1.08 (0.89 to 1.31)
BC	1.12 (0.93 to 1.33)	1.14 (0.95 to 1.36)	1.14 (0.96 to 1.35)	1.15 (0.97 to 1.36)	1.15 (0.97 to 1.37)
O_3	0.68 (0.54 to 0.85)	0.83 (0.74 to 0.93)	0.83 (0.74 to 0.93)	0.88 (0.79 to 0.97)	0.88 (0.79 to 0.97)

Abbreviation: NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm; BC, black carbon; O_3, ozone.

[a] Models were adjusted for age (time scale), sex (strata), sub-cohort (strata), calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level in 2001.

Results are presented as hazard ratio and 95% confidence interval [HR (95% CI)] for the following increments: 5 $\mu g/m^3$ for $PM_{2.5}$, 10 $\mu g/m^3$ for NO_2, 0.5 $10^{-5}/m$ for BC and 10 $\mu g/m^3$ for O_3.

Table S4. Associations of long-term exposure to air pollution estimated from either ELAPSE or ESCAPE with liver cancer incidence in the subset of the pooled cohort with available information from both exposure models (DNC and E3N were excluded; N=203,787, Cases= 370).

Pollutant	Increment	ELAPSE exposure	ESCAPE exposure
		Hazard ratio (95% CI)[a]	Hazard ratio (95% CI)[a]
NO_2	10 μg/m^3	1.08 (0.87 to 1.36)	1.22 (1.03 to 1.45)
$PM_{2.5}$	5 μg/m^3	1.00 (0.76 to 1.31)	1.38 (0.87 to 2.18)
BC	0.5 10^{-5}/m	1.03 (0.83 to 1.29)	1.11 (0.88 to 1.41)

Abbreviation: NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm; BC, black carbon; O_3, ozone.

[a] From models adjusted for age (time scale), sex (strata), sub-cohort (strata), calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level in 2001.

Table S5. Associations of long-term exposure to air pollution estimated from either ELAPSE or MAPLE with liver cancer incidence in the subset of the pooled cohort with available information from both exposure models (N=330,064, Cases=512).

Pollutant	Increment	ELAPSE exposure	MAPLE exposure	
			The year 2001[a]	The year 2010
		Hazard ratio (95% CI) [b]	Hazard ratio (95% CI)[b]	Hazard ratio (95% CI) [b]
$PM_{2.5}$	5 $\mu g/m^3$	1.12 (0.92 to 1.36)	1.26 (1.03 to 1.54)	1.33 (1.11 to 1.60)

[a] The earliest available year from MAPLE

[b] From models adjusted for age (time scale), sex (strata), sub-cohort (strata), calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level in 2001.

Table S6. Associations between long-term exposure to air pollution and liver cancer incidence with including additional confounders (educational level and alcohol intake) in the subset of the pooled cohort with the available information.

Dataset	Model	Cases	N	Hazard ratio (95% CI)			
				NO_2	$PM_{2.5}$	BC	O_3
Full dataset with available information for the main model[a]	Main model[a]	512	330,064	1.17 (1.02 to 1.35)	1.12 (0.92 to 1.36)	1.15 (1.00 to 1.33)	0.70 (0.58 to 0.85)
Dataset with available information for the main model and education level[b]	Main model[a]	219	178,632	1.12 (0.92 to 1.35)	0.94 (0.63 to 1.41)	1.05 (0.87 to 1.27)	0.71 (0.56 to 0.91)
	Main model[a] + Education level			1.12 (0.93 to 1.35)	0.94 (0.63 to 1.41)	1.05 (0.87 to 1.27)	0.71 (0.56 to 0.91)
Dataset with available information for the main model and alcohol intake[c]	Main model[a]	184	153,053	1.21 (0.99 to 1.49)	1.04 (0.68 to 1.60)	1.14 (0.93 to 1.39)	0.71 (0.55 to 0.93)
	Main model[a] + Alcohol intake			1.20 (0.98 to 1.48)	1.03 (0.67 to 1.59)	1.13 (0.92 to 1.39)	0.72 (0.55 to 0.93)

Abbreviation: NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm; BC, black carbon; O_3, ozone.
Results are presented as hazard ratio and 95% confidence interval [HR (95% CI)] for the following increments: 5 $\mu g/m^3$ for $PM_{2.5}$, 10 $\mu g/m^3$ for NO_2, 0.5 10^{-5}/m for BC and 10 $\mu g/m^3$ for O_3.
[a] The model adjusted for age (time scale), sex (strata), sub-cohort (strata), calendar year of baseline, smoking status, employment status, and mean income at the neighborhood level in 2001.
[b] Education level (low/medium/high); The entire participants in the VHM&PP cohort dropped out because of missing information on education level.
[c] Alcohol intake (low-<4g/day; medium-4-15g/day; high->15g/day); The entire participants in the VHM&PP, CEANS-Sixty, and CEANS-SNAC-K cohort dropped out because of missing information on alcohol intake.

Table S7. Associations between long-term exposure to air pollution and the risk of liver cancer incidence after excluding a single cohort at a time from the pooled cohort.

Dataset	Cases, N (%)	Cohort, N	NO_2		
			Model 1[a] Hazard ratio (95% CI)	Model 2[b] Hazard ratio (95% CI)	Model 3[c] Hazard ratio (95% CI)
Pooled cohort	512 (0.16)	330,064	1.14 (1.00 to 1.31)	1.12 (0.98 to 1.29)	1.17 (1.02 to 1.35)
Excluding CEANS	494 (0.16)	310,433	1.16 (1.01 to 1.33)	1.14 (0.99 to 1.31)	1.19 (1.03 to 1.38)
Excluding DCH	376 (0.14)	275,922	1.12 (0.95 to 1.32)	1.10 (0.93 to 1.30)	1.11 (0.93 to 1.33)
Excluding DNC	497 (0.16)	305,784	1.13 (0.99 to 1.31)	1.11 (0.97 to 1.28)	1.17 (1.01 to 1.35)
Excluding E3N	479 (0.17)	280,542	1.15 (0.99 to 1.34)	1.13 (0.97 to 1.32)	1.18 (1.01 to 1.38)
Excluding EPIC-NL	494 (0.17)	296,571	1.15 (1.00 to 1.32)	1.13 (0.98 to 1.30)	1.19 (1.03 to 1.37)
Excluding VHM&PP	220 (0.12)	181,068	1.10 (0.92 to 1.33)	1.08 (0.90 to 1.30)	1.13 (0.93 to 1.36)
Dataset	Case, N (%)	Cohort, N	$PM_{2.5}$		
			Model 1[a] Hazard ratio (95% CI)	Model 2[b] Hazard ratio (95% CI)	Model 3[c] Hazard ratio (95% CI)
Pooled cohort	512 (0.16)	330,064	1.10 (0.91 to 1.33)	1.09 (0.90 to 1.32)	1.12 (0.92 to 1.36)
Excluding CEANS	494 (0.16)	310,433	1.11 (0.92 to 1.35)	1.10 (0.91 to 1.34)	1.13 (0.93 to 1.37)
Excluding DCH	376 (0.14)	275,922	1.09 (0.90 to 1.34)	1.09 (0.89 to 1.34)	1.09 (0.89 to 1.34)
Excluding DNC	497 (0.16)	305,784	1.09 (0.90 to 1.32)	1.08 (0.89 to 1.31)	1.11 (0.92 to 1.35)
Excluding E3N	479 (0.17)	280,542	1.12 (0.92 to 1.38)	1.12 (0.91 to 1.37)	1.14 (0.93 to 1.39)
Excluding EPIC-NL	494 (0.17)	296,571	1.11 (0.92 to 1.34)	1.10 (0.91 to 1.33)	1.13 (0.93 to 1.37)
Excluding VHM&PP	220 (0.12)	181,068	0.96 (0.65 to 1.43)	0.92 (0.62 to 1.37)	0.95 (0.64 to 1.42)
Dataset	Case, N (%)	Cohort, N	BC		
			Model 1[a] Hazard ratio (95% CI)	Model 2[b] Hazard ratio (95% CI)	Model 3[c] Hazard ratio (95% CI)
Pooled cohort	512 (0.16)	330,064	1.13 (0.98 to 1.30)	1.11 (0.97 to 1.28)	1.15 (1.00 to 1.33)
Excluding CEANS	494 (0.16)	310,433	1.15 (1.00 to 1.33)	1.13 (0.98 to 1.30)	1.17 (1.01 to 1.36)
Excluding DCH	376 (0.14)	275,922	1.13 (0.95 to 1.35)	1.12 (0.94 to 1.33)	1.12 (0.94 to 1.35)
Excluding DNC	497 (0.16)	305,784	1.13 (0.97 to 1.30)	1.11 (0.96 to 1.28)	1.15 (0.99 to 1.33)

Excluding E3N	479 (0.17)	280,542	1.14 (0.98 to 1.34)	1.13 (0.96 to 1.31)	1.17 (1.00 to 1.37)
Excluding EPIC-NL	494 (0.17)	296,571	1.13 (0.98 to 1.30)	1.11 (0.96 to 1.28)	1.16 (1.00 to 1.34)
Excluding VHM&PP	220 (0.12)	181,068	1.05 (0.87 to 1.26)	1.02 (0.85 to 1.23)	1.06 (0.87 to 1.28)
Dataset	Case, N (%)	Cohort, N	O3		
			Model 1[a] Hazard ratio (95% CI)	Model 2[b] Hazard ratio (95% CI)	Model 3[c] Hazard ratio (95% CI)
Pooled cohort	512 (0.16)	330,064	0.69 (0.58 to 0.84)	0.71 (0.59 to 0.86)	0.70 (0.58 to 0.85)
Excluding CEANS	494 (0.16)	310,433	0.69 (0.57 to 0.83)	0.70 (0.58 to 0.85)	0.70 (0.58 to 0.84)
Excluding DCH	376 (0.14)	275,922	0.83 (0.65 to 1.05)	0.84 (0.66 to 1.06)	0.83 (0.65 to 1.07)
Excluding DNC	497 (0.16)	305,784	0.68 (0.56 to 0.82)	0.70 (0.58 to 0.84)	0.69 (0.57 to 0.83)
Excluding E3N	479 (0.17)	280,542	0.67 (0.54 to 0.82)	0.69 (0.56 to 0.85)	0.68 (0.56 to 0.84)
Excluding EPIC-NL	494 (0.17)	296,571	0.69 (0.57 to 0.84)	0.71 (0.58 to 0.86)	0.70 (0.57 to 0.84)
Excluding VHM&PP	220 (0.12)	181,068	0.67 (0.53 to 0.86)	0.70 (0.55 to 0.9)	0.70 (0.55 to 0.90)

Abbreviation: The 'Cardiovascular Effects of Air Pollution and Noise in Stockholm' [CEANS] from Stockholm county of Sweden, which is comprised of the four sub-cohorts: Swedish National Study on Aging and Care in Kungsholmen [SNAC-K], Stockholm Screening Across the Lifespan Twin study [SALT], Stockholm cohort of 60-year-olds [Sixty], and Stockholm Diabetes Prevention Program [SDPP]; the 'Diet, Cancer and Health cohort' [DCH] from Copenhagen and Aarhus of Denmark; the 'Danish Nurse Cohort' [DNC] from entire Denmark, which included two sub-cohorts from recruitment rounds in 1993 and 1999; the 'Dutch European Investigation into Cancer and Nutrition' [EPIC-NL] from four cities in the Netherland, consisting of 'EPIC-Monitoring Project on Risk Factors' [EPIC-MORGEN] and 'EPIC-Chronic Diseases in the Netherlands' [EPIC-PROSPECT]; the 'Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale' [E3N] from entire France; and 6) the 'Vorarlberg Health Monitoring and Prevention Programme' [VHM&PP] from Vorarlberg, Austria; NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm; BC, black carbon; O_3, ozone.

Results are presented as hazard ratio and 95% confidence interval [HR (95% CI)] for the following increments: 5 $\mu g/m^3$ for $PM_{2.5}$, 10 $\mu g/m^3$ for NO_2, 0.5 $10^{-5}/m$ for BC and 10 $\mu g/m^3$ for O_3.

[a] Model1 was adjusted for age (time scale), sex (strata), sub-cohort (strata), and calendar year of baseline.

[b] Model2 was adjusted for age (time scale), sex (strata), sub-cohort (strata), and calendar year of baseline, smoking status, and employment status.

[c] Model3 was adjusted for age (time scale), sex (strata), sub-cohort (strata), and calendar year of baseline, smoking status, employment status, and mean income at the neighborhood level in 2001.

Table S8. Associations between $PM_{2.5}$ components and liver cancer incidence among participants with full information in the main model (N=330,064. Cases=512).

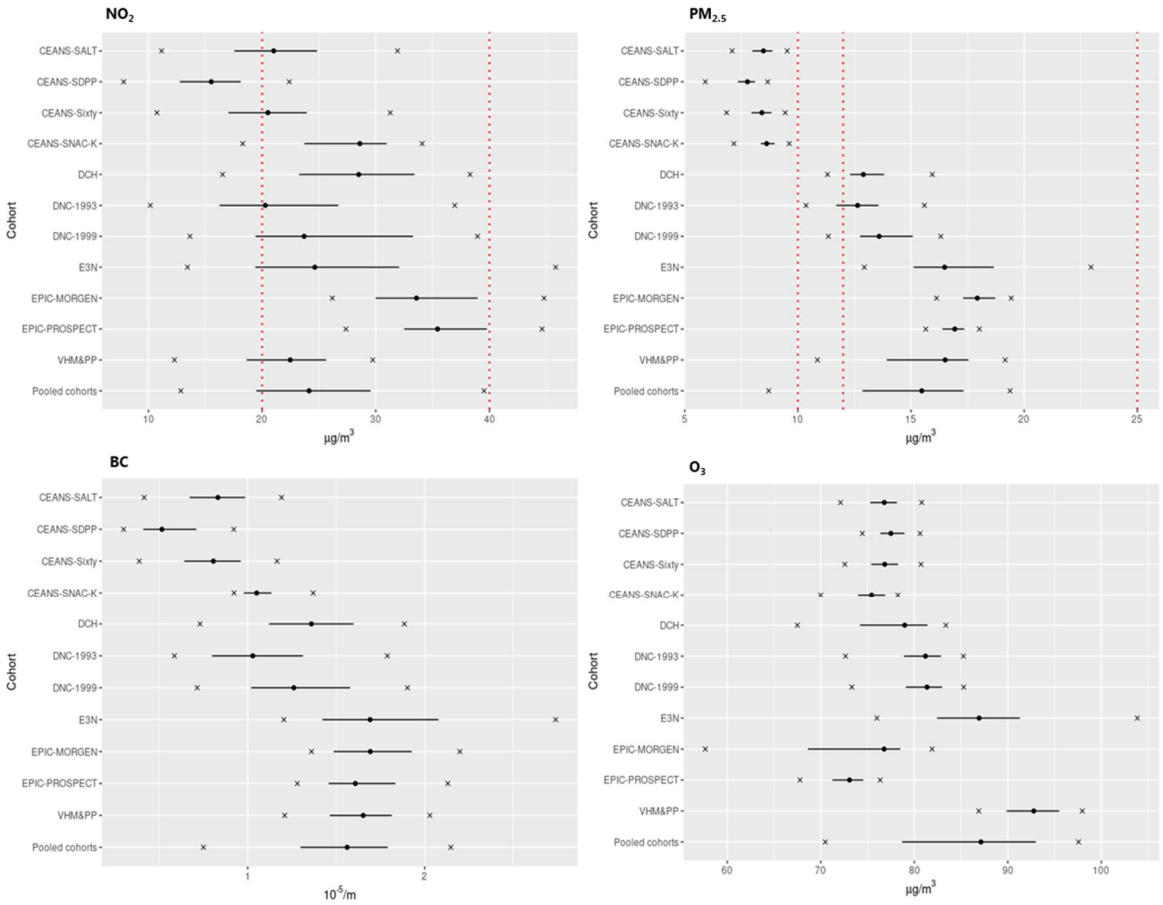
Pollutant	Exposure estimate method							
	Supervised linear regression				Random Forest			
	Unit, ng/m^3 (IQR)	main model [a]	Two-pollutant model [a] (Further adjusted for pollutants below)		Unit, ng/m^3 (IQR)	main model [a]	Two-pollutant model [a] (Further adjusted for pollutants below)	
			$PM_{2.5}$	NO_2			$PM_{2.5}$	NO_2
Cu	3.7	1.24 (1.06 to 1.44)	1.32 (1.07 to 1.63)	1.23 (0.94 to 1.62)	1.9	1.09 (0.95 to 1.24)	1.06 (0.90 to 1.24)	0.90 (0.72 to 1.13)
Fe	55.8	1.19 (1.04 to 1.36)	1.22 (1.04 to 1.45)	1.19 (0.92 to 1.56)	34.1	1.08 (0.95 to 1.22)	1.06 (0.92 to 1.21)	0.92 (0.76 to 1.12)
Zn	10.7	1.19 (1.09 to 1.31)	1.21 (1.09 to 1.34)	1.17 (1.05 to 1.31)	9.6	1.20 (0.96 to 1.49)	1.16 (0.90 to 1.50)	1.08 (0.84 to 1.40)
S	212.2	1.41 (1.09 to 1.81)	1.67 (1.15 to 2.43)	1.31 (0.95 to 1.81)	121.3	1.30 (1.02 to 1.65)	1.28 (0.97 to 1.68)	1.18 (0.90 to 1.56)
Ni	0.8	1.20 (1.06 to 1.35)	1.19 (1.05 to 1.35)	1.16 (1.00 to 1.34)	0.9	1.14 (0.81 to 1.62)	1.13 (0.80 to 1.60)	1.03 (0.72 to 1.49)
V	1.7	1.28 (1.14 to 1.44)	1.28 (1.13 to 1.46)	1.26 (1.10 to 1.44)	1.6	1.34 (1.00 to 1.79)	1.32 (0.98 to 1.77)	1.23 (0.90 to 1.68)
Si	24.1	1.12 (1.00 to 1.26)	1.11 (0.98 to 1.26)	1.04 (0.88 to 1.24)	23.0	0.96 (0.86 to 1.07)	0.96 (0.86 to 1.07)	0.94 (0.84 to 1.05)
K	82.3	1.16 (0.95 to 1.42)	1.14 (0.87 to 1.48)	1.08 (0.87 to 1.34)	201.0	1.05 (0.66 to 1.66)	0.89 (0.52 to 1.51)	0.84 (0.51 to 1.37)

Abbreviation: Cu, copper; Fe, iron; Zn, zinc; S, sulfur; Ni, nickel; V, vanadium; Si, Silicon; K, potassium; NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 μm.

Results are presented as hazard ratio and 95% confidence interval [HR (95% CI)] for interquartile range (IQR) increase for each $PM_{2.5}$ components.

[a] Models adjusted for age (time scale), sex (strata), sub-cohort (strata), calendar year of baseline, smoking status, employment status, and 2001 mean income at the neighborhood level in 2001.

Figure S1. Bar plots of the annual mean concentration of NO_2, $PM_{2.5}$, BC, and O_3 by each cohort study.



Abbreviation: The 'Cardiovascular Effects of Air Pollution and Noise in Stockholm' [CEANS] from Stockholm county of Sweden, which is comprised of the four sub-cohorts: Swedish National Study on Aging and Care in Kungsholmen [SNAC-K], Stockholm Screening Across the Lifespan Twin study [SALT], Stockholm cohort of 60-year-olds [Sixty], and Stockholm Diabetes Prevention Program [SDPP]; the 'Diet, Cancer and Health cohort' [DCH] from Copenhagen and Aarhus of Denmark; the 'Danish Nurse Cohort' [DNC] from entire Denmark, which included two sub-cohorts from recruitment rounds in 1993 and 1999; the 'Dutch European Investigation into Cancer and Nutrition' [EPIC-NL] from four cities in the Netherland, consisting of 'EPIC-Monitoring Project on Risk Factors' [EPIC-MORGEN] and 'EPIC-Chronic Diseases in the Netherlands' [EPIC-PROSPECT]; the 'Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale' [E3N] from entire France; and 6) the 'Vorarlberg Health Monitoring and Prevention Programme' [VHM&PP] from Vorarlberg, Austria; NO_2, nitrogen dioxide; $PM_{2.5}$, particulate matters with aerodynamic diameters of less than 2.5 µm; BC, black carbon; O_3, ozone.
Red dotted lines for NO_2 indicate the 40 µg/m^3 (the WHO guideline), and 20 µg/m^3, all annual averages.
Red dotted lines for $PM_{2.5}$ indicate the 10 µg/m^3 (the WHO guideline),12 µg/m^3 (the US EPA NAAQS), and 25 µg/m^3 (the EU standard), all annual averages.
The solid circle and bars shows the median, 25th, and 75th percentiles of concentrations; the x shows the 5th and 95th percentile values.

Supplementary reference

1. Magnusson PKE, Almqvist C, Rahman I, Ganna A, Viktorin A, Walum H, Halldner L, Lundström S, Ullén F, Långström N, Larsson H, Nyman A, Gumpert CH, Råstam M, Anckarsäter H, Cnattingius S, Johannesson M, Ingelsson E, Klareskog L, de Faire U, Pedersen NL, Lichtenstein P. The Swedish Twin Registry: Establishment of a Biobank and Other Recent Developments. *Twin Res Hum Genet*. 2013;16:317-329.
2. Eriksson AK, Ekbom A, Granath F, Hilding A, Efendic S, Östenson CG. Psychological distress and risk of pre-diabetes and type 2 diabetes in a prospective study of swedish middle-aged men and women. *Diabet Med*. 2008;25:834-842.
3. Wändell PE, Wajngot A, de Faire U, Hellénius ML. Increased prevalence of diabetes among immigrants from non-European countries in 60-year-old men and women in Sweden. *Diabetes Metab*. 2007;33:30-36.
4. Lagergren M, Fratiglioni L, Hallberg IR, Berglund J, Elmståhl S, Hagberg B, Holst G, Rennemark M, Sjölund BM, Thorslund M, Wiberg I, Winblad B, Wimo A. A longitudinal study integrating population, care and social services data. The Swedish National study on Aging and Care (SNAC). *Aging Clin Exp Res*. 2004;16:158-168.
5. Tjønneland A, Olsen A, Boll K, Stripp C, Christensen J, Engholm G, Overvad K. 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;35:432-441.
6. Hundrup YA, Simonsen MK, Jørgensen T, Obel EB. Cohort profile: The danish nurse cohort. *Int J Epidemiol*. 2012;41:1241-1247.
7. Clavel-Chapelon F. Cohort Profile: The French E3N Cohort Study. *Int J Epidemiol*. 2015;44:801-809.
8. Beulens JWJ, Monninkhof EM, Verschuren WMM, van der Schouw YT, Smit J, Ocke MC, Jansen EHJM, van Dieren S, Grobbee DE, Peeters PHM, Bueno-de-Mesquita HB. Cohort profile: the EPIC-NL study. *Int J Epidemiol*. 2010;39:1170-1178.
9. Ulmer H, Kelleher CC, Fitz-Simon N, Diem G, Concin H. Secular trends in cardiovascular risk factors: an age-period cohort analysis of 698,954 health examinations in 181,350 Austrian men and women. *J Intern Med*. 2007;261:566-576.