Research

Examining the Shape of the Association between Low Levels of Fine Particulate Matter and Mortality across Three Cycles of the Canadian Census Health and Environment Cohort

Amanda J. Pappin,[1]* Tanya Christidis,[1] Lauren L. Pinault,[1] Dan L. Crouse,[2,3] Jeffrey R. Brook,[4] Anders Erickson,[5] Perry Hystad,[6] Chi Li,[7] Randall V. Martin,[7,8,9] Jun Meng,[7,9] Scott Weichenthal,[10,11] Aaron van Donkelaar,[7,9] Michael Tjepkema,[1] Michael Brauer,[5] and Richard T. Burnett[12]

[1]Health Analysis Division, Statistics Canada, Ottawa, Ontario, Canada
[2]Department of Sociology, University of New Brunswick, Fredericton, New Brunswick, Canada
[3]New Brunswick Institute for Research, Data, and Training, Fredericton, New Brunswick, Canada
[4]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
[5]School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada
[6]College of Public Health and Human Sciences, Oregon State University, Corvallis, Oregon, USA
[7]Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Nova Scotia, Canada
[8]Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA
[9]Department of Energy, Environmental & Chemical Engineering, Washington University in St. Louis, St. Louis, Missouri, United States
[10]Department of Epidemiology, Biostatistics & Occupational Health, McGill University, Montreal, Quebec, Canada
[11]Air Health Science Division, Health Canada, Ottawa, Ontario, Canada
[12]Population Studies Division, Health Canada, Ottawa, Ontario, Canada

BACKGROUND: Ambient fine particulate air pollution with aerodynamic diameter $\leq 2.5\,\mu m$ ($PM_{2.5}$) is an important contributor to the global burden of disease. Information on the shape of the concentration–response relationship at low concentrations is critical for estimating this burden, setting air quality standards, and in benefits assessments.

OBJECTIVES: We examined the concentration–response relationship between $PM_{2.5}$ and nonaccidental mortality in three Canadian Census Health and Environment Cohorts (CanCHECs) based on the 1991, 1996, and 2001 census cycles linked to mobility and mortality data.

METHODS: Census respondents were linked with death records through 2016, resulting in 8.5 million adults, 150 million years of follow-up, and 1.5 million deaths. Using annual mailing address, we assigned time-varying contextual variables and 3-y moving-average ambient $PM_{2.5}$ at a $1 \times 1\,km$ spatial resolution from 1988 to 2015. We ran Cox proportional hazards models for $PM_{2.5}$ adjusted for eight subject-level indicators of socioeconomic status, seven contextual covariates, ozone, nitrogen dioxide, and combined oxidative potential. We used three statistical methods to examine the shape of the concentration–response relationship between $PM_{2.5}$ and nonaccidental mortality.

RESULTS: The mean 3-y annual average estimate of $PM_{2.5}$ exposure ranged from 6.7 to 8.0 $\mu g/m^3$ over the three cohorts. We estimated a hazard ratio (HR) of 1.053 [95% confidence interval (CI): 1.041, 1.065] per 10-$\mu g/m^3$ change in $PM_{2.5}$ after pooling the three cohort-specific hazard ratios, with some variation between cohorts (1.041 for the 1991 and 1996 cohorts and 1.084 for the 2001 cohort). We observed a supralinear association in all three cohorts. The lower bound of the 95% CIs exceeded unity for all concentrations in the 1991 cohort, for concentrations above $2\,\mu g/m^3$ in the 1996 cohort, and above $5\,\mu g/m^3$ in the 2001 cohort.

DISCUSSION: In a very large population-based cohort with up to 25 y of follow-up, $PM_{2.5}$ was associated with nonaccidental mortality at concentrations as low as $5\,\mu g/m^3$. https://doi.org/10.1289/EHP5204

Introduction

Exposure to ambient fine particulate air pollution with aerodynamic diameter $\leq 2.5\,\mu m$ ($PM_{2.5}$) consistently ranks among the leading risk factors for premature death and disease worldwide (Burnett et al. 2018; GBD 2017 Risk Factors Collaborators 2018; Lim et al. 2012). A number of studies supporting this work have found that the relationship between $PM_{2.5}$ concentrations and mortality risk (for various causes) was supralinear across the

global range (Burnett et al. 2014; Pope et al. 2009, 2011; Yin et al. 2017). In a detailed examination of the shape of the $PM_{2.5}$– mortality association in 15 of the world's largest cohorts (Burnett et al. 2018), 12 displayed a supralinear association. A supralinear concentration–response curve is characterized by a positively sloped curve of decreasing steepness, such that risk initially rises rapidly with a decreasing slope as concentrations increase. Studies that specifically characterize the shape of concentration–response relationships at low-$PM_{2.5}$ mass concentrations offer great value given the steady decline in $PM_{2.5}$ levels over recent decades in North America (ECCC 2017). Further, a substantial proportion of the global $PM_{2.5}$ disease burden is from relatively low level exposures (Apte et al. 2015). Canada is an ideal setting to conduct such analyses, given the availability of large, national cohorts with sufficient sample sizes and detailed exposure information at low $PM_{2.5}$ concentrations.

Canadian cohort studies have shown consistent positive associations between $PM_{2.5}$ and mortality from various causes at low $PM_{2.5}$ concentrations (i.e., annual concentrations generally below $20\,\mu g/m^3$ even in large urban areas) (Crouse et al. 2012, 2015; Nasari et al. 2016; Pinault et al. 2016b, 2017; Weichenthal et al. 2017). Crouse et al. (2012) used the 1991 Canadian Census Health and Environment Cohort (CanCHEC) to conduct the first nationwide cohort analysis and identified a hazard ratio (HR) for nonaccidental mortality of 1.07 [95% confidence interval (CI): 1.06, 1.08] per 10-$\mu g/m^3$ change in $PM_{2.5}$ among nonimmigrant

adults. In a recent analysis of the 2001 CanCHEC, Pinault et al. (2017) reported a larger HR of 1.18 (95% CI: 1.15, 1.21) for $PM_{2.5}$ and nonaccidental mortality. While these studies made important contributions to the evidence base for mortality risks at low $PM_{2.5}$ levels, they also had several important limitations. For example, with the exception of Pinault et al. (2017), past studies used coarser-resolution $PM_{2.5}$ models (i.e., 10×10 km) to assign exposures to census respondents. Furthermore, most of the previous studies excluded immigrants, although this group represents nearly 20% of the Canadian population. Additionally, most of these studies had only 10 y of follow-up.

The present study specifically investigated the shape of the concentration–response function between $PM_{2.5}$ and nonaccidental mortality at low levels of exposure among Canadian adults. We examined data from the 1991, 1996, and 2001 CanCHECs with follow-up until 2016. We address a number of limitations of previous cohort studies in Canada by extending the period of follow-up to 25 y (i.e., for individuals in the 1991 cohort), including all but recent immigrants in the analysis, using annual 1 km^2 $PM_{2.5}$ estimates from 1988–2015, using time-varying contextual covariates over the duration of follow-up, and applying a validated marginalization index to represent four orthogonal dimensions of neighborhood- or community-level socioeconomic status. We examined the shape of associations at low levels of $PM_{2.5}$ exposure by applying restricted cubic splines (RCS) (Harrell 2015), monotonically increasing smoothing splines (MISS) (Pya and Wood 2015), and the Shape Constrained Health Impact Function (SCHIF) (Nasari et al. 2016).

Methods

Analytical Cohort

We created three new, separate analytical cohorts from the 1991, 1996, and 2001 CanCHECs. Briefly, the CanCHECs are population-based, administrative data cohorts that link eligible census respondents (i.e., noninstitutional respondents to the mandatory Statistics Canada long-form census questionnaire that is distributed to 20% of all Canadian households) to their annual mailing address (1981–2016) and follow subjects for mortality. Information on a number of variables capturing the social and economic status of the subjects was available from the long-form census (Table 1).

The linkage was approved by Statistics Canada (linkage requests 037-2016 and 045-2015) and is governed by the Directive on Microdata Linkage (Statistics Canada 2017a). Eligible respondents were first linked probabilistically to tax records using sex, date of birth, postal code (PC), and spousal date of birth (if available).

Table 1. $PM_{2.5}$ Distribution by cohort with lowest (2nd percentile) and highest (98th percentile) knot values for restricted cubic spline.

	2001	1996	1991
100% max	18.50	20.00	20.00
99%	12.30	15.00	17.26
98% (highest knot)	11.70	13.97	17.03
95%	10.70	12.20	14.63
90%	9.80	10.70	12.60
75% Q3	8.23	8.84	9.83
50% median	6.40	6.75	7.40
25% Q1	4.87	5.04	5.38
10%	3.97	4.10	4.26
5%	3.57	3.67	3.80
2% (lowest knot)	3.00	3.29	3.43
1%	3.00	3.05	3.13
0% min	0.37	0.37	0.37
Mean	6.68	7.18	7.95
SD	2.24	2.70	3.28

Note: SD, standard deviation.

This initial linkage was necessary since linkage to the mortality database is based on the social insurance number (SIN), a unique personal identifier. The long form censuses did not capture the SIN, but they are available on tax records. The linkage rate to tax records near the time of cohort inception was approximately 80%, of which 99% were determined to be accurate matches (Christidis et al. 2018; Pinault et al. 2016a; Wilkins et al. 2008).

Mortality and PC history data were attached to the census-tax cohorts using Statistics Canada's Social Data Linkage Environment (SDLE) Derived Record Depository (DRD) (Statistics Canada 2017b), a dynamic relational database. About 99.8% of all deaths that occurred in Canada between 1991 and 2016 were linked to the DRD before being linked to eligible census respondents. From this linkage, we obtained death date and underlying cause of death if it occurred between census day and 31 December 2016. Mortality data were coded by underlying cause of death according to the *International Classification of Diseases, 9th Revision*, prior to 2000 (ICD-9; WHO 1977), and *10th Revision* post-2000 (ICD-10; WHO 2016).

We enhanced the cohort with a number of data elements characterizing the environment in which each subject lived, using PC histories from tax records, of which the primary source was income tax filings (1981 to 2016). We assigned a representative point (latitude and longitude) to each PC (Statistics Canada 2017c). In large cities, PCs often correspond to a single block face, though in rural areas, they can range over much larger areas. Similarly, the point estimates of PCs are accurate within 0.2 km in urban centers and 5.6 km in rural areas (Khan et al. 2018). These point estimates were used to derive estimates of air pollution and location-based contextual risk factors.

We note that these three linked cohorts are newly created using an enhanced linkage environment (SDLE) and thus are not identical to the CanCHEC cohorts used in previous publications (Crouse et al. 2015; Pinault et al. 2017).

Outdoor Air Pollution Concentrations

We used annual ambient $PM_{2.5}$ surfaces as our main exposure of interest at a $0.01° \times 0.01°$ resolution ($\sim 1 \text{ km}^2$) over North America for 1981–2016 (Meng et al. 2019; van Donkelaar et al. 2015). $PM_{2.5}$ estimates for the years 1998–2012 were developed by relating satellite-based retrievals of total column aerosol optical depth to near-surface $PM_{2.5}$ concentrations using the geophysical relationship simulated by a chemical transport model (CTM). These estimates were constrained using ground-based monitoring from the National Air Pollution Surveillance (NAPS) program stations, along with other North America–based measurements, land-use information, and simulated composition in a geographically weighted regression (V4.NA.01; van Donkelaar et al. 2015). For years outside this period, we used $PM_{2.5}$ surfaces developed using a backcasting method (Meng et al. 2019) that applied observed annual trends in ground monitoring data for $PM_{2.5}$ and coarser size fractions to adjust pregridded $PM_{2.5}$ estimates backwards or forwards in time. We estimated a 3-y moving-average exposure window with 1-y lag for assigning $PM_{2.5}$ exposures for consistency with previous studies, as ambient $PM_{2.5}$ is regulated based on a 3-y time window in Canada (CCME 2012).

We assigned estimates of exposures to ambient ozone (O_3; as a May–September daily maximum 8-h average) and nitrogen dioxide (NO_2; annual) for inclusion in multipollutant models. Additionally, we estimated a measure of the combined oxidant capacity of O_3 and NO_2, expressed as $O_x = 2/3 O_3 + 1/3 NO_2$ (Weichenthal et al. 2017). We estimated a 3-y average with 1-y lag for each of O_3, NO_2, and O_x for inclusion in the hazard models. Modeled O_3 surfaces at 21-km spatial resolution were developed by Environment and Climate Change Canada (ECCC) for

2002–2015 using chemical transport modeling informed by surface observations (Robichaud and Ménard 2014; Robichaud et al. 2016). Estimates of ambient NO_2 were based on a national land-use regression model (LUR) developed for 2006 (Hystad et al. 2011) with a spatial resolution of 100 m. The LUR estimates were built using satellite-derived NO_2 (with 10-km resolution), distances to highways and major roads, and roadway kernel density gradients as predictive variables.

We temporally adjusted the O_3 and NO_2 models to obtain exposure estimates over our study period (i.e., 1988–2015). Our adjustment was based on observed trends in ground monitoring data for NO_2 and O_3 from the NAPS in Canada. For each of 24 census divisions (CDs) that had monitoring data available, we estimated yearly adjustment factors from the ratio of observed CD-average concentration in a specific year to the reference year(s) for which the original surfaces were estimated (i.e., 2006 for NO_2 and 2002–2015 average for O_3). We assigned adjustment factors for each PC from the closest CD.

Contextual Covariates

We assigned contextual risk factors describing neighborhood-level characteristics and geographic identifiers using residential PC and data from the closest census (every 5 y from 1991 through 2016). We included in our analysis the Canadian Marginalization Index (CAN-Marg), population size of home community or city, an indicator of urban form, and regional airshed to capture risk factors beyond those captured at the subject level. We assigned these four categories of contextual covariates to residential PCs linked to census geography for each census year.

CAN-Marg is a publicly available index of neighborhood marginalization in Canada that was developed by Matheson et al. (2012) using an analysis of the 2001 and 2006 long-form census cycles. CAN-Marg consists of four dimensions that aim to capture different aspects of marginalization: material deprivation, residential instability, ethnic concentration, and dependency. Following the methodology of Matheson et al. (2012), we developed CAN-Marg using the 1991 and 1996 censuses. We assigned CAN-Marg to PC locations and then created quintiles (based on the cohort distribution) of the continuous values in the four Can-MARG dimensions in order to account for any potentially nonlinear associations with mortality.

We used a variable to describe the population size of a subject's community (Pinault et al. 2017) (Table 2). We categorized geographic locations into the following: census metropolitan areas (CMAs) or census agglomerations (CAs; Statistics Canada 2003) with a population exceeding 1.5 million; 500,000–1.49 million; 100,000–499,999; 30,000–99,999; or 10,000–29,999, as well as non-CMAs/CAs. We note that although non-CMA/CAs are always rural areas, CMAs cover both the urban core of a city and the urban–rural fringe, such that some rural locations fall within a CMA/CA. As such, this variable does not perfectly delineate subjects living in rural vs. urban settings.

To further differentiate between the kinds of built environments and neighborhoods within communities, we created an urban form variable following the methodology developed by Gordon and Janzen (2013). This measure of urban form is informed by population density and the most frequently reported mode of transportation (active or transit) in each census tract as reported on each census cycle. The categories of this variable include an active urban core, transit-reliant suburb, car-reliant suburb, exurban, and non-CMA/CA. We note that mode of commute was not reported on the 1991 census cycle and was derived from the 1996 census.

We included airshed as a geographic covariate in our analysis (Crouse et al. 2016). Airsheds divide Canada into six regions (Western, Prairie, West Central, Southern Atlantic, East Central, and Northern) based on large-scale differences in air masses and meteorology. Airsheds can also be used to represent regional differences in mortality rates across Canada that remain uncaptured by other geographic covariates.

Exclusion of Person-Years of Follow-Up

PC history was not available for each person in every year of follow-up, either because they did not file a tax return or from gaps in administrative data. Any gaps in PCs that had the same PC prior to and after the gap were assigned that PC for all years of the gap. After this imputation, 87.8% of person-years had an available PC. We imputed an additional 2.1% of person-years of missing PCs if the bounding PCs shared the first two characters (Finès et al. 2017; Pinault et al. 2017), totaling 89.9% of person-years with a PC.

After imputation, person-years were excluded if they did not have an assigned PC. Further exclusions of person-years occurred due to: immigrated to Canada less than 10 y before survey date (9,364,400 person-years), age during follow-up period exceeded 89 y (7,357,200), could not be linked to air pollution values (17,814,400), could not be linked to CAN-Marg values (25,973,900), could not be linked to CMA/CA size (25,613,100), could not be linked to airshed (25,545,500), 3-y moving average being informed by only 1 y of exposure (20,056,400), and year after subject death (17,936,100). The above are not mutually exclusive numbers of exclusions. The total available person-years for analyses were 150,996,500 after all exclusions (Figure S1).

Statistical Analysis

Our primary statistical model relating exposure to mortality was the Cox proportional hazards model (Cox 1972). Participants were at least 25 y of age at the beginning of each cohort, and the time axis was the year of follow-up until 2016. Person-years before a census year and after a subject's death year were excluded from analysis. Events were determined by year of death for nonaccidental causes. The Cox model baseline hazard function was stratified by age (5-y groups), sex, and immigrant status (yes or no). This latter strata variable was included since immigrants to Canada live longer, on average, than do Canadian-born citizens (Ng 2011). We excluded immigrants living in Canada for less than 10 y at cohort commencement due to the healthy immigrant effect (Ng 2011) and lack of knowledge of their historical air pollution exposures. Each subject was censored at 89 y of age, either at the start of each cohort or during follow-up, due to evidence suggesting an increased mismatch between home address and the tax return mailing address with increasing age (Bérard-Chagnon 2017). We postulate that relatives of elderly people were completing their tax returns. Each of the three CanCHEC cohorts (1991, 1996, and 2001) was examined separately. Estimates of the cohort-specific hazard ratios were then pooled to form a single summary hazard ratio. We also conducted a test for differences in the hazard ratios between cohorts (Cochran 1950).

We fit two covariate adjustment models for each cohort. The first was based on a directed acyclic graph (DAG; Figure S2) and consisted of all the geographically based predictors: CAN-Marg (four dimensions), airshed, urban form, and CMA/CA size. The second model, denoted as "Full," additionally included the subject-level predictors (income, education, occupational class, Indigenous status, visible minority status, employment status, and marital status), which are not *a priori* causes of outdoor $PM_{2.5}$ concentrations, but which may contribute to confounding owing to a chance imbalance across the $PM_{2.5}$ distribution.

We also conducted analysis by categories of: immigrant status (yes or no), sex (male or female), and age during follow-up (<65, 65–74, or ≥75 y) for each cohort separately, again pooling the

Table 2. Descriptive statistics of 1991, 1996, and 2001 Canadian Census Health and Environment Cohort (CanCHEC) study cohorts.

| | 1991 CanCHEC | | | | 1996 CanCHEC | | | | 2001 CanCHEC | | | |
| | Person-years | | PM$_{2.5}$ concentration (µg/m^3) | | Person-years | | PM$_{2.5}$ concentration (µg/m^3) | | Person-years | | PM$_{2.5}$ concentration (µg/m^3) | |
Covariate	n	%	Mean	SD	n	%	Mean	SD	n	%	Mean	SD
Total	54,042,100	100.0%	8.10	3.44	54,082,700	100.0%	7.18	2.70	42,871,700	100.0%	6.68	2.24
Sex												
Male	27,769,300	51.4	8.14	3.43	28,240,300	52.2	7.23	2.69	22,308,500	52.0	6.72	2.24
Female	26,272,800	48.6	8.06	3.45	25,842,400	47.8	7.13	2.70	20,563,200	48.0	6.64	2.24
Age group												
24–34 y	3,540,300	6.6	10.56	3.96	3,170,900	5.9	8.37	3.18	2,659,100	6.2	7.18	2.52
35–44 y	10,088,100	18.7	8.75	3.58	10,368,100	19.2	7.37	2.82	8,518,300	19.9	6.64	2.29
45–54 y	14,381,600	26.6	7.72	3.25	14,364,600	26.6	6.99	2.61	11,112,700	25.9	6.58	2.21
55–64 y	11,986,800	22.2	7.55	3.19	11,839,400	21.9	6.93	2.56	9,401,200	21.9	6.57	2.17
65–74 y	8,227,800	15.2	7.84	3.32	8,259,400	15.3	7.12	2.64	6,335,700	14.8	6.69	2.21
75–89 y	5,817,700	10.8	7.88	3.13	6,080,300	11.2	7.26	2.55	4,844,600	11.3	6.90	2.20
Immigrant status												
Nonimmigrant	45,568,900	84.3	7.82	3.34	45,280,200	83.7	6.94	2.62	35,465,100	82.7	6.46	2.20
Immigrant, 11–20 y	2,711,900	5.0	9.48	3.48	2,114,600	3.9	8.40	2.60	1,871,300	4.4	7.82	2.00
Immigrant, 21–30 y	2,585,500	4.8	9.57	3.57	3,148,000	5.8	8.45	2.69	2,055,200	4.8	7.69	2.06
Immigrant, >30 y	3,175,800	5.9	9.63	3.74	3,539,800	6.6	8.47	2.84	3,480,100	8.1	7.69	2.23
Visible minority status												
No	51,309,700	94.9	8.02	3.42	51,075,900	94.4	7.10	2.69	37,791,200	88.2	6.69	2.22
Yes	2,732,400	5.1	9.61	3.40	3,006,700	5.6	8.56	2.49	5,080,500	11.9	6.60	2.41
Indigenous status												
No	51,920,400	96.1	8.17	3.43	51,916,000	96.0	7.28	2.68	40,921,000	95.5	6.78	2.22
Yes	2,121,800	3.9	6.28	3.06	2,166,700	4.0	4.90	1.99	1,950,700	4.6	4.61	1.69
Marital status												
Never married/not common-law	6,776,600	12.5	8.52	3.49	6,597,000	12.2	7.57	2.74	5,233,700	12.2	7.05	2.29
Common-law	4,035,500	7.5	7.73	3.24	5,066,100	9.4	6.81	2.50	4,693,200	11.0	6.50	2.15
Married	37,316,200	69.1	7.95	3.40	36,029,200	66.6	7.07	2.68	27,590,800	64.4	6.57	2.22
Separated	1,275,500	2.4	8.46	3.53	1,323,000	2.5	7.49	2.78	1,032,800	2.4	6.89	2.29
Divorced	2,524,900	4.7	8.62	3.46	2,861,000	5.3	7.65	2.68	2,404,100	5.6	7.09	2.21
Widowed	2,113,400	3.9	9.11	3.76	2,206,300	4.1	7.83	2.91	1,917,000	4.5	7.09	2.37
Educational attainment												
<High school graduation	17,025,100	31.5	8.00	3.55	16,190,200	29.9	7.01	2.80	11,564,900	27.0	6.50	2.34
High school, with or without trades certificate	20,516,400	38.0	8.00	3.39	19,575,600	36.2	7.11	2.65	15,491,200	36.1	6.60	2.22
Postsecondary nonuniversity	8,940,200	16.5	8.11	3.35	10,185,400	18.8	7.23	2.63	8,542,100	19.9	6.71	2.17
University degree	7,560,400	14.0	8.55	3.38	8,131,400	15.0	7.64	2.62	7,273,600	17.0	7.08	2.16
Income inadequacy												
Q1 (lowest income)	8,373,700	15.5	8.25	3.61	8,693,400	16.1	7.21	2.81	7,216,300	16.8	6.76	2.36
Q2	9,989,100	18.5	8.22	3.50	9,949,900	18.4	7.28	2.75	8,078,000	18.8	6.74	2.28
Q3	11,417,600	21.1	8.09	3.42	11,248,900	20.8	7.21	2.69	8,772,600	20.5	6.70	2.23
Q4	12,023,900	22.3	8.03	3.37	11,875,400	22.0	7.15	2.65	9,194,600	21.5	6.64	2.19
Q5 (highest income)	12,237,800	22.6	7.97	3.34	12,315,200	22.8	7.08	2.62	9,610,200	22.4	6.58	2.17
Employment status												
Employed	38,679,600	71.6	8.00	3.36	36,133,000	66.8	7.13	2.64	28,781,900	67.1	6.65	2.20
Unemployed	3,380,300	6.3	7.65	3.42	3,018,000	5.6	6.72	2.71	1,739,800	4.1	6.06	2.29
Not in labor force	11,982,200	22.2	8.53	3.63	14,931,700	27.6	7.41	2.82	12,350,000	28.8	6.82	2.32
Occupational class												
Management	4,811,500	8.9	8.17	3.36	4,107,400	7.6	7.25	2.62	3,806,700	8.9	6.75	2.18
Professional	6,718,300	12.4	8.25	3.35	6,598,700	12.2	7.39	2.62	5,593,100	13.1	6.87	2.17
Skilled, technical, and supervisory	14,058,800	26.0	7.77	3.33	12,379,800	22.9	6.89	2.61	10,290,500	24.0	6.45	2.18
Semi-skilled	14,023,100	26.0	7.99	3.40	13,401,200	24.8	7.11	2.67	9,410,200	22.0	6.62	2.22
Unskilled	4,339,400	8.0	7.92	3.46	4,091,000	7.6	6.95	2.72	2,996,100	7.0	6.46	2.29
Not applicable	10,090,900	18.7	8.64	3.66	13,504,600	25.0	7.47	2.82	10,775,000	25.1	6.88	2.33
Residential instability (CAN-Marg)												
Q1 (lowest marginalization)	12,129,000	22.4	7.28	3.19	12,537,400	23.2	6.50	2.53	10,200,700	23.8	6.06	2.09
Q2	13,959,900	25.8	7.46	3.29	14,328,200	26.5	6.63	2.61	11,519,100	26.9	6.20	2.16
Q3	11,234,900	20.8	8.18	3.56	11,059,600	20.5	7.23	2.79	8,645,800	20.2	6.69	2.29
Q4	9,674,400	17.9	8.86	3.41	9,488,700	17.5	7.92	2.62	7,407,900	17.3	7.37	2.14
Q5 (highest marginalization)	7,044,000	13.0	9.58	3.25	6,668,800	12.3	8.53	2.37	5,098,300	11.9	7.97	1.95
Dependence (CAN-Marg)												
Q1 (lowest marginalization)	8,881,200	16.4	8.30	3.48	8,958,200	16.6	7.14	2.68	7,416,500	17.3	6.44	2.12
Q2	9,310,000	17.2	8.44	3.43	8,908,400	16.5	7.38	2.65	6,938,000	16.2	6.73	2.11
Q3	9,079,900	16.8	8.67	3.55	8,702,200	16.1	7.65	2.75	6,663,500	15.5	7.06	2.25
Q4	11,665,500	21.6	8.22	3.43	11,497,400	21.3	7.36	2.71	8,882,800	20.7	6.93	2.31

Note: CA, census agglomeration; CAN-Marg, Canadian Marginalization Index; CMA, census metropolitan area; Pop, population; SD, standard deviation.

Table 2. (*Continued.*)

	1991 CanCHEC				1996 CanCHEC				2001 CanCHEC			
	Person-years		PM$_{2.5}$ concentration (µg/m^3)		Person-years		PM$_{2.5}$ concentration (µg/m^3)		Person-years		PM$_{2.5}$ concentration (µg/m^3)	
Covariate	n	%	Mean	SD	n	%	Mean	SD	n	%	Mean	SD
Q5 (highest marginalization)	15,105,600	28.0	7.32	3.22	16,016,500	29.6	6.72	2.62	12,970,900	30.3	6.41	2.27
Material deprivation (CAN-Marg)												
Q1 (lowest marginalization)	11,497,200	21.3	7.59	3.07	10,947,700	20.2	7.00	2.52	8,651,800	20.2	6.61	2.04
Q2	12,268,900	22.7	8.18	3.24	11,270,800	20.8	7.29	2.49	8,383,500	19.6	6.86	2.06
Q3	10,965,300	20.3	8.46	3.43	10,652,500	19.7	7.44	2.63	8,375,900	19.5	6.85	2.16
Q4	8,826,900	16.3	8.59	3.49	9,190,500	17.0	7.61	2.73	7,335,900	17.1	7.08	2.29
Q5 (highest marginalization)	10,483,800	19.4	7.76	3.88	12,021,200	22.2	6.70	2.98	10,124,600	23.6	6.15	2.47
Ethnic concentration (CAN-Marg)												
Q1 (lowest marginalization)	15,066,600	27.9	6.81	3.20	17,014,800	31.5	6.08	2.38	14,272,200	33.3	5.71	1.96
Q2	12,404,500	23.0	7.79	3.22	13,274,500	24.5	7.02	2.54	10,882,100	25.4	6.57	2.16
Q3	9,435,300	17.5	8.37	3.33	9,457,600	17.5	7.48	2.63	7,569,000	17.7	6.93	2.22
Q4	8,678,400	16.1	9.18	3.34	7,620,700	14.1	8.26	2.66	5,616,300	13.1	7.71	2.16
Q5 (highest marginalization)	8,457,300	15.7	9.43	3.48	6,715,100	12.4	8.65	2.64	4,532,100	10.6	8.28	1.80
CMA/CA size												
Pop: > 1,500,000	15,000,000	27.8	10.07	3.32	14,932,200	27.6	8.85	2.36	12,159,300	28.4	8.13	1.83
Pop: 500,000–1,499,999	8,747,700	16.2	8.16	2.82	8,679,700	16.1	7.40	2.18	6,991,200	16.3	6.95	1.81
Pop: 100,000–499,999	9,759,400	18.1	8.68	3.56	9,751,700	18.0	7.83	2.92	7,826,800	18.3	7.16	2.42
Pop: 30,000–99,999	5,510,600	10.2	7.66	3.27	5,267,500	9.7	6.68	2.42	4,081,700	9.5	6.14	1.98
Pop: 10,000–29,000	2,111700	3.9	6.44	2.51	2,107,400	3.9	5.73	1.88	1,699,900	4.0	5.27	1.42
Non-CMA/CA	12,912,700	23.9	5.78	2.33	13,344,100	24.7	5.13	1.72	10,112,800	23.6	4.83	1.39
Urban form												
Active urban core	4,152,200	7.7	10.02	3.26	4,006,700	7.4	8.95	2.40	3,220,700	7.5	8.32	1.92
Transit-reliant suburb	3,490,900	6.5	10.50	3.26	3,405,600	6.3	9.31	2.30	2,689,000	6.3	8.58	1.69
Car-reliant suburb	21,595,500	40.0	9.16	3.30	21,787,500	40.3	8.18	2.50	17,930,300	41.8	7.53	2.00
Exurban	2,951,100	5.5	6.57	2.59	3,000,100	5.6	5.98	2.06	2,471,500	5.8	5.68	1.72
Non-CMA/CA	21,852,400	40.4	6.50	2.87	21,882,700	40.5	5.70	2.16	16,560,200	38.6	5.27	1.74
Airshed												
Western	6,532,200	12.1	7.92	3.44	6,404,500	11.8	6.58	2.08	5,137,600	12.0	5.95	1.55
Prairie	6,942700	12.9	6.45	2.07	7,016,900	13.0	5.92	1.73	5,675,500	13.2	5.61	1.54
West Central	3,205,600	5.9	5.86	1.73	3,322,900	6.1	5.30	1.41	2,589,400	6.0	5.01	1.25
Southern Atlantic	5,312,600	9.8	5.41	1.87	5,324,000	9.8	4.80	1.30	4,044,400	9.4	4.54	1.05
East Central	31,626,600	58.5	9.23	3.48	31,439,700	58.1	8.25	2.71	24,932,700	58.2	7.65	2.17
Northern	422,300	0.8	4.19	1.37	574,600	1.1	3.80	1.11	492,100	1.2	3.67	1.05

cohort-specific hazard ratio estimates among the three cohorts. In addition, we examined the PM$_{2.5}$ association, adjusting for O$_3$, NO$_2$, or O$_x$ by cohort.

Shape of the Association between PM$_{2.5}$ and Mortality

The main purpose of the current study was to describe the association between PM$_{2.5}$ and mortality in a manner that can be used for risk and benefits assessment. The standard approach is the log-linear (LL) model that relates the logarithm of the hazard ratio to exposure in a linear manner: $logHR(PM_{2.5}) = \beta PM_{2.5}$. Here, β represents a change in relative risk per unit change in concentration estimated using the Cox model. Nasari et al. (2016) developed the SCHIF in order to extend the LL model to nonlinear transformations of exposure, $T(PM_{2.5})$, with the form: $SCHIF(PM_{2.5}) = \theta T(PM_{2.5})$. Nasari et al. (2016) proposed a specific family of transformations based on a sigmodal function that could accommodate a variety of shapes they suggested would be suitable for risk and benefits assessment. Here, θ represents a change in risk per unit change in $T(PM_{2.5})$. The SCHIF can then be used in benefits assessment in a manner similar to the LL model after a suitable transformation of concentration.

The SCHIF approach has two major limitations: The first is in defining an appropriate number of transformations of a sigmodal function that can capture all shapes of interest; the second is that the method requires considerable computational capacity if the selected family is very large. This can be a serious limitation when cohort sizes are very large, such as with the CanCHECs.

Spline methods have also been proposed to characterize the shape. RCS with a few knots have been used (Beelen et al. 2014) in addition to smoothing splines (Di et al. 2017). However, the manner in which splines are presented by graphic representation of the mean predictions and uncertainty bounds over the concentration range limits their use in risk and benefits assessment, as these assessments typically require a differentiable algebraic function in addition to a quantitative estimate of uncertainty by concentration.

We developed and applied a new method that combines the flexibility of splines and the ease of use of the SCHIF in benefits assessment. Our method involves three steps. The first step is a data reduction step in which we fit a RCS with a large number of knots in order to characterize the shape of the concentration–response relationship in sufficient detail. RCS can easily be fit to large cohorts, as they only involve a series of transformations of concentration. Here, we have converted millions of person-years of data into a few hundred observations of RCS predictions over the observed concentration range. In step 2, we smooth the potential erratic predictions due to the large number of knots using a MISS, and in step 3, we fit our SCHIF function to the MISS predictions. In addition, we model the uncertainty in the spline fit as a cubic polynomial in concentration in a manner that assigns all uncertainty to the θ parameter in the SCHIF model, but unlike the LL model, uncertainty can vary with concentration. We now have a differentiable algebraic function of both relative risk and its uncertainty by concentration. This approach also allows for visualization of the SCHIF as well as its representation of the underlying data (as summarized by the RCS).

Specifically, we selected 15 knots defined at the 2nd, 4th, 10th, 14th, 18th, 22nd, 26th, 50th, 74th, 78th, 82nd, 86th, 90th, 94th, and 98th percentiles of the $PM_{2.5}$ person-year distribution. We selected a large number of knots covering both the lower and upper quartiles in order to capture a variety of desired shapes. From this first step, we obtain estimates of the logarithm of the RCS hazard ratio (logRCS) and the associated standard error at several hundred concentrations between the minimum and the 99th percentile of the exposure distribution. We do not include predictions above the 99th percentile, since RCS are linear beyond the highest knot concentration. This linear form can have some influence on the shape of the SCHIF throughout the concentration range, and especially over the higher concentrations, since the SCHIF is a single algebraic function. We also fixed the logRCS to zero at the minimum concentration, and its associated standard error was also set to zero.

In step 2, we smooth the potentially erratic logRCS predictions with a MISS in order to obtain predictions suitable to model with the SCHIF algebraic function, which itself is monotonically increasing (Pya and Wood 2015). The SCHIF hazard ratio function has the form:

$$logSCHIF(z) = \theta f(z) l(z)$$

with $l(z) = \frac{1}{\left[1 + \exp\left(-\frac{z-\mu}{\tau r}\right)\right]}$ a logistic function in concentration. Here, θ, μ, and τ are unknown parameters to be estimated from the data, r is the range in the translated exposure, and $z = PM_{2.5} - \min(PM_{2.5})$ such that $logSCHIF(0) = 0$. The function $f(z)$ can take two forms: $f(z) = z$ (linear) and $f(z) = log(z+1)$ (log). We have constructed the SCHIF to be similar to the LL model, $logLL(z) = \beta z$, by writing: $logSCHIF(z) = \theta T(z)$, where $T(z) = f(z)l(z)$ is a specific transformation of concentration.

The linear form $f(z) = z$ can model both linear and sublinear associations, while the log form $f(z) = log(z+1)$ can model supralinear associations with mortality. Both forms can accommodate S-shaped functions through $l(z)$. Sets of values (μ, τ) are selected that define the shape of $l(z)$. Larger values of μ result in larger ranges of concentration for which a sublinear association is modeled at lower concentrations due to the property of the logistic function. Larger values of τ generate shapes for $l(z)$ with less curvature. By limiting the ranges for (μ, τ), we can limit the amount of curvature in the SCHIF.

A linear regression model was constructed using each transformation as the single predictor and the MISS prediction as the response. Using the MISS predictions, we were then able to select a wide range of values of the parameters to examine a wide variety of shapes that is not possible by modeling the subject-level cohort data. We selected values of μ ranging from 0 to r by integers, and τ ranging from 0.1 to 1 by 0.1 increments. For each set of parameters and the two forms of $f(z)$, we obtained an estimate of θ and its standard error. We then created a single SCHIF curve by a weighted average of all the SCHIF curves examined, with weights determined by the fit of each curve on the MISS values. However, as the model averaged predictions at each concentration are themselves a potentially complicated function, these predictions can be summarized as a single algebraic function. Specifically, we fit a generalization of the SCHIF model

$$logSCHIF(z) = \frac{\theta \log\left(\frac{z}{\alpha} + 1\right)}{1 + \exp\left[-(z-\mu)/v\right]}$$

to the mean SCHIF predicted curve over the concentration range. We added an additional parameter α to model the combination of the linear and log forms of $f(z)$ used in the fitting step. The function $\log\left(\frac{z}{\alpha} + 1\right)$ is nearly linear in z for large values of α. We

collapse the product τr into a single parameter v to simplify the reporting of the parameter estimates.

In the LL model, all uncertainty in the hazard ratio is assigned to the single unknown parameter, β. We aim to make a similar characterization of uncertainty in the SCHIF predictions, where all the uncertainty is ascribed to the parameter θ. We do this by considering a model of the standard error in the RCS predictions. However, unlike the LL model, RCS standard errors can vary in a nonlinear manner with concentration. We therefore consider a model for the standard error as a function of concentration of the form: $se_{RCS}(z) = se_\theta(z) \times T(z)$, with $se_\theta(z)$ denoting our standard error model of θ, dependent on concentration. We select a general model that can accommodate a variety of shapes such as a cubic polynomial with the form: $se_\theta(z) = \sigma_0 + \sigma_1 z + \sigma_2 z^2 + \sigma_3 z^3$.

Finally, we construct pooled SCHIF models among the three cohorts in the following manner: Let $v_c(z)$ be the variance of the logarithm of the SCHIF prediction, $logSCHIF_c(z)$, at concentration z for cohort $c = 1,2,3$. We construct a meta-analytic summary of the SCHIF predictions among the three cohorts as:

$$logSCHIF_{Pooled}(z) = \sum_{c=1}^{3} w_c \, logSCHIF_c(z),$$

where $w_c(z) = [1/v_c(z)] / \sum\limits_{c=1}^{3} 1/v_c(z)$. For the variance of $logSCHIF_{Pooled}(z)$, we include the variation in predictions among the cohorts in addition to the sampling uncertainty for each cohort as:

$$\sum_{c=1}^{3} w_c^2(z) \{v_c(z) + [logSCHIF_c(z) - logSCHIF_{Pooled}(z)]^2\}.$$

In order to obtain an algebraic function for the pooled SCHIF, we used nonlinear regression to estimate the SCHIF parameters, with $logSCHIF_{Pooled}(z)$ defining the data for the regression. We also modeled the standard error of the pooled SCHIF in a manner similar to that for each cohort separately. The variance of the pooled SCHIFs is a function of both the variance of each cohort-specific SCHIF prediction and the squared difference between the cohort-specific SCHIF predictions and the pooled SCHIF prediction. This latter term captures the uncertainty in both the shape and magnitude of the hazard ratio predictions among the three cohorts.

Results

Main Analysis

$PM_{2.5}$ by cohort and covariate categories. Table 1 presents percentiles of the $PM_{2.5}$ distribution based on person-years for each of the three cohorts separately. Concentrations were highest for the 1991 cohort, moderate for the 1996 cohort, and lowest for the 2001 cohort. Concentration differences were well within 1 $\mu g/m^3$ between cohorts for median and lower percentiles, with greater differences for the higher percentiles, suggesting that greater declines in exposure were observed in locations with higher levels. The spatial distribution of $PM_{2.5}$ across Canada is presented for selected 3-y averages (Figure 1). Concentrations declined over time in the heavily populated areas of Southern Ontario and Quebec. Moderate concentrations were observed in the earlier time periods for Northern Canada and the Prairies. These levels declined through the 1990s but then increased during the latter part of our cohort follow-up period.
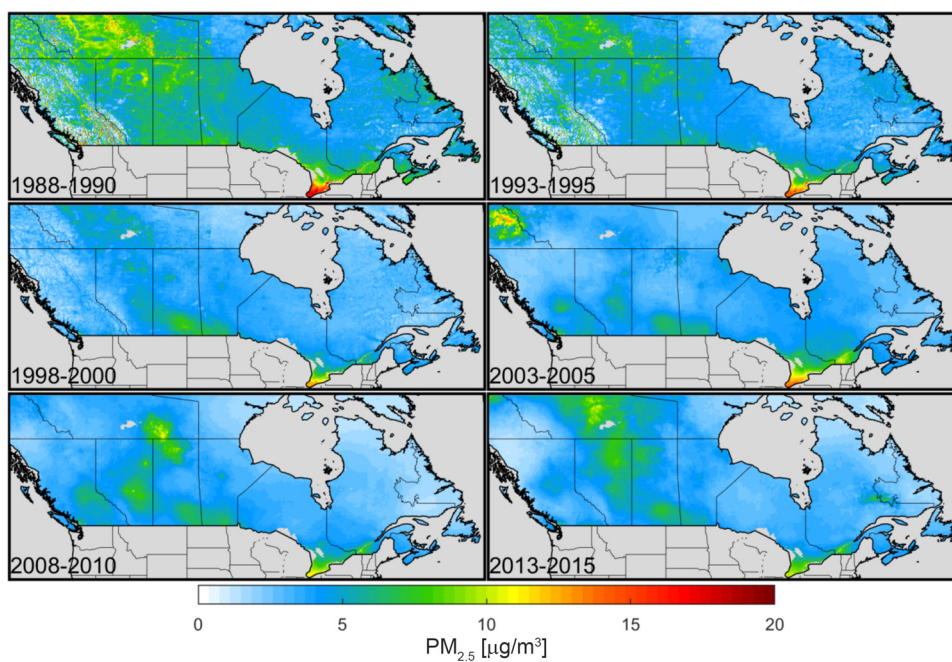
Figure 1. Spatial distribution of particulate matter with aerodynamic diameter ≤2.5 μm (PM$_{2.5}$) across Canada for selected 3-y averages: 1998–2000 (first exposure assigned to the 1991 cohort), 1993–1995 (first exposure assigned to 1996 cohort), 1998–2000 (first exposure assigned to 2001 cohort), 2003–2005, 2008–2010, and 2013–2015 (exposure assigned to last year of follow-up, 2015).

Table 2 reports both the number of person-years and percentages among the categories of mortality predictors for each cohort separately, in addition to the mean and standard deviation of PM$_{2.5}$ assigned to each category. Males tended to be assigned higher concentrations than females in all three cohorts, although the difference was very small ($<1\,\mu g/m^3$). There was a U-shaped pattern with age at cohort commencement for all three cohorts, with concentration declining with age up to the 55- to 64-y-old group and then increasing. Immigrants were consistently assigned higher concentrations than nonimmigrants; however, concentrations were similar over the length an immigrant subject lived in Canada. Subjects who defined themselves as visible minorities had higher assigned concentrations than those subjects who did not in the 1991 and 1996 cohorts. Subjects of Indigenous identity had lower concentrations. Married and common-law subjects had lower assigned exposures compared to other marital categories in all cohorts. Exposure monotonically increased with educational attainment in all cohorts. However, exposure monotonically declined with income. Employed subjects at the time of interview had higher exposures compared to those unemployed subjects. Exposure tended to decline over the occupational class categories moving from management/professional to semi- and unskilled workers. Note that the "not in the labor force" and "not-applicable occupational class" categories had the highest exposures, possibly to due to older subjects who tended to have higher than average exposures. There was a tendency for exposure to increase over the quintiles of three of the CAN-Marg dimensions: residential instability, material deprivation, and ethnic concentration, with no clear trend for the fourth dimension, dependence. Outdoor concentrations increased with CMA/CA size and for the inner-city categories of urban form. Of the six airsheds, the East Central contained 58% of person-years and had the highest concentrations. Based on the associations between several geographic and subject based covariates, there is some potential that adjustment for these variables could influence the magnitude of our estimates of the PM$_{2.5}$–mortality association.

Hazard ratio estimates. Table 3 reports the hazard ratio and 95% confidence limits per 10-μg/m³, for each cohort separately and pooled among the three cohorts by categories of immigrant status, age, and sex, for both the DAG and Full models. There was a tendency for the hazard ratio to be larger under the Full model compared to the DAG for the 1991 and 1996 cohorts, but smaller for the 2001 cohort. Consequently, there was less variation among the hazard ratios between cohorts under the Full compared to the DAG models. The Full model was a better predictor of mortality compared to the DAG model based on its much lower Akaike Information Criterion/Schwarz's Bayesian Criterion values (see Table S1). We therefore focus our interpretation on the results using the Full model.

When all subjects were considered together, hazard ratio estimates were similar for the 1991 and 1996 cohorts (HR = 1.041), with a larger estimate observed for the 2001 cohort (HR = 1.084). The pooled cohort HR estimate was 1.053 (95% CI: 1.041, 1.065). Hazard ratio estimates for nonimmigrants were higher than for immigrants in the 1991 and 1996 cohorts, but lower in the 2001 cohort. Hazard ratio estimates for males were higher than for females in the 1991 and 1996 cohorts but lower in the 2001 cohort. Hazard ratio estimates declined with age in all three cohorts, however.

Hazard ratio estimates based on interquartile range changes in concentrations were larger for O$_x$ compared to O$_3$, and lowest for NO$_2$ (Table 3). The PM$_{2.5}$ HR estimate was moderately sensitive to adjustment for NO$_2$, declining from 1.053 to 1.043 per 10-μg/m³, but very sensitive to adjustment for either O$_3$, declining to 0.982, and O$_x$, declining to 0.955.

Shape of PM$_{2.5}$–mortality association. The shape of the association between PM$_{2.5}$ concentrations and mortality for the Full model is displayed in Figure 2 for each of the three cohorts separately and pooled among cohorts using the SCHIF. MISS predictions (dashed black line) and RCS predictions (dashed red line) over the concentration range are also displayed. A similar shape is observed in each cohort for the MISS, with a steep increase below 5 μg/m³ followed by a much shallower increase for higher concentrations. The SCHIF predictions also display a supralinear association with concentration. Note that the SCHIF predictions display much less curvature than the MISS; a design feature of

Table 3. Hazard ratio (HR) estimates and 95% confidence intervals (CIs) for the association between PM$_{2.5}$ and nonaccidental mortality, as well as for copollutants (NO$_2$, Ozone, oxidative potential), within the Canadian Health and Environment Cohorts (CanCHECs) from 1991, 1996, 2001, and pooled cohorts. Effect modification analyses by immigrant status, sex, and age, and multi-pollutant models are also provided.

Subgroup/model	Model form	1991 Cohort HR	95% CI		1996 Cohort HR	95% CI		2001 Cohort HR	95% CI		Pooled results[a] HR	95% CI		p-Value
All subjects														
—	DAG[b]	0.982	0.959	1.006	1.033	1.016	1.051	1.120	1.096	1.146	1.044	1.031	1.056	<0.01
—	Full[c]	1.041	1.016	1.066	1.041	1.024	1.059	1.084	1.060	1.108	1.053	1.041	1.065	<0.01
Immigrant status[d]														
No	DAG	0.975	0.951	1.000	1.024	1.005	1.043	1.105	1.078	1.133	1.032	1.019	1.045	<0.01
No	Full	1.049	1.022	1.076	1.058	1.039	1.078	1.089	1.062	1.116	1.064	1.050	1.078	0.09
Yes	DAG	1.016	0.945	1.092	1.082	1.040	1.125	1.190	1.131	1.253	1.104	1.073	1.136	<0.01
Yes	Full	1.006	0.935	1.081	1.027	0.987	1.068	1.109	1.053	1.167	1.049	1.019	1.079	0.03
Sex[d]														
Female	DAG	0.956	0.921	0.993	1.001	0.976	1.026	1.121	1.084	1.160	1.022	1.004	1.040	<0.01
Female	Full	1.009	0.972	1.048	1.008	0.983	1.034	1.093	1.056	1.130	1.031	1.013	1.050	<0.01
Male	DAG	0.993	0.963	1.024	1.055	1.032	1.078	1.116	1.083	1.150	1.055	1.039	1.071	<0.01
Male	Full	1.053	1.021	1.086	1.062	1.039	1.086	1.071	1.040	1.104	1.062	1.046	1.079	0.74
Age during follow-up[d]														
<65 y	DAG	1.022	0.971	1.075	1.057	1.019	1.097	1.176	1.119	1.236	1.078	1.051	1.106	<0.01
<65 y	Full	1.079	1.026	1.136	1.095	1.056	1.136	1.165	1.108	1.225	1.109	1.081	1.137	0.07
65–74 y	DAG	0.984	0.939	1.031	1.079	1.044	1.116	1.176	1.122	1.232	1.077	1.052	1.103	<0.01
65–74 y	Full	1.069	1.020	1.120	1.092	1.057	1.130	1.130	1.078	1.184	1.096	1.070	1.122	0.25
>75 y	DAG	0.929	0.899	0.961	0.986	0.964	1.009	1.062	1.031	1.094	0.994	0.978	1.010	<0.01
>75 y	Full	0.972	0.940	1.005	0.985	0.963	1.008	1.031	1.001	1.062	0.995	0.979	1.011	0.02
Single pollutant														
NO$_2$	DAG	1.009	1.004	1.015	0.997	0.993	1.001	1.003	0.998	1.008	1.002	0.999	1.004	<0.01
NO$_2$	Full	1.015	1.009	1.020	1.001	0.997	1.005	1.003	0.998	1.009	1.005	1.002	1.008	<0.01
O$_3$	DAG	1.016	1.006	1.027	1.035	1.029	1.041	1.041	1.034	1.049	1.034	1.030	1.038	<0.01
O$_3$	Full	1.044	1.033	1.055	1.076	1.069	1.082	1.081	1.073	1.088	1.073	1.068	1.077	<0.01
O$_x$	DAG	1.030	1.018	1.043	1.037	1.029	1.044	1.049	1.040	1.058	1.040	1.035	1.045	0.03
O$_x$	Full	1.068	1.056	1.081	1.086	1.078	1.093	1.094	1.085	1.103	1.086	1.080	1.091	<0.01
Two pollutant														
Adjusted for NO$_2$[e]														
PM$_{2.5}$	DAG	0.966	0.942	0.991	1.038	1.02	1.057	1.115	1.089	1.142	1.040	1.028	1.053	<0.01
NO$_2$	DAG	1.010	1.004	1.015	0.997	0.993	1.001	1.003	0.998	1.009	1.002	0.999	1.005	<0.01
PM$_{2.5}$	Full	1.014	0.989	1.041	1.039	1.021	1.058	1.078	1.052	1.104	1.043	1.030	1.056	<0.01
NO$_2$	Full	1.015	1.010	1.021	1.001	0.997	1.006	1.004	0.998	1.009	1.006	1.003	1.009	<0.01
Adjusted for O$_3$[e]														
PM$_{2.5}$	DAG	0.969	0.944	0.994	0.996	0.978	1.014	1.073	1.048	1.098	1.011	0.998	1.024	<0.01
O$_3$	DAG	1.016	1.006	1.026	1.034	1.028	1.04	1.040	1.033	1.047	1.033	1.029	1.037	<0.01
PM$_{2.5}$	Full	1.003	0.978	1.029	0.963	0.946	0.981	0.996	0.973	1.020	0.982	0.970	0.994	0.01
O$_3$	Full	1.043	1.033	1.054	1.074	1.068	1.08	1.079	1.072	1.086	1.071	1.067	1.075	<0.01
Adjusted for O$_x$[e]														
PM$_{2.5}$	DAG	0.950	0.925	0.977	0.988	0.97	1.007	1.056	1.031	1.083	0.998	0.985	1.011	<0.01
O$_x$	DAG	1.028	1.017	1.039	1.034	1.027	1.04	1.045	1.037	1.053	1.037	1.032	1.041	0.03
PM$_{2.5}$	Full	0.967	0.941	0.994	0.941	0.923	0.959	0.970	0.946	0.994	0.955	0.943	0.968	0.10
O$_x$	Full	1.062	1.051	1.074	1.078	1.071	1.085	1.086	1.077	1.094	1.078	1.073	1.083	<0.01

Note: —, no data; NO$_2$, nitrogen dioxide; O$_x$, combined oxidant capacity of O$_3$ and NO$_2$; O$_3$, ambient ozone; PM$_{2.5}$, particulate matter with aerodynamic diameter ≤2.5 μm.
[a]Tests for heterogeneity of hazard ratio among cohorts: *$p < 0.05$, **$p < 0.01$.
[b]Directed acyclic graph (DAG) model is stratified by 5-y age groups by age at baseline, sex, and immigrant status and included the following geographic-based covariates: four Canadian Marginalization Index dimensions, urban form, CMA/CA size and airshed.
[c]Full model is stratified by 5-y age groups by age at baseline, sex, and immigrant status and included the geographic based covariates: four Canadian Marginalization Index dimensions, urban form, CMA/CA size and airshed, and the subject-based covariates: marital status, education, income quintile, Indigenous status, visible minority status, employment status, and occupational class.
[d]Note that the models by immigrant status are not stratified by immigrant status. The models by sex are not stratified by sex.
[e]PM$_{2.5}$ always uses 10 units, copollutants use: O$_3$, 10.20 ppb; NO$_2$, 6.63 ppb; O$_x$, 8.05 ppb.

constraining the shape of the SCHIF. The SCHIF 95% CIs (gray-shaped area) are clearly widest in the 2001 cohort, as it had the shortest follow-up time (15 y) and lowest concentrations (Table 1) compared to the 1991 and 1996 cohorts. The steepness of the increase in the HR below 5 μg/m^3 appears to dampen between the 1991 to 1996 to 2001 cohorts, with the SCHIF predictions of the MISS improving over these lower concentrations as the start date of the cohorts increased. The lower bound of the CIs exceeded unity for all concentrations for the 1991 cohort, above 2 μg/m^3 for the 1996 cohort, and above 5 μg/m^3 for the 2001 cohort.

We display the association between both θ and its standard error and show that association varies with concentration by the ratio $N(z) = \theta/se_\theta(z)$ in Figure 3. Here, $N(z)$ represents the signal-to-noise ratio by concentration. This ratio increases with concentration for all three cohorts, exceeding the 1.96 value (dashed black line) for all concentrations in the 1991 cohort, above 2 μg/m^3 in the 1996 cohort, and above 5 μg/m^3 in the 2001 cohort. For the pooled SCHIF, the ratio increases for concentrations less than 7 μg/m^3 and then is relatively stable for higher concentrations. This pattern is due to the additional variation between the SCHIF values among the three cohorts at higher levels. The parameter estimates for both the SCHIF and its standard error are given in Table 4.

Discussion

The three CanCHEC cohorts included 8.5 million adults with 150 million person-years of follow-up and nearly 1.5 million deaths,
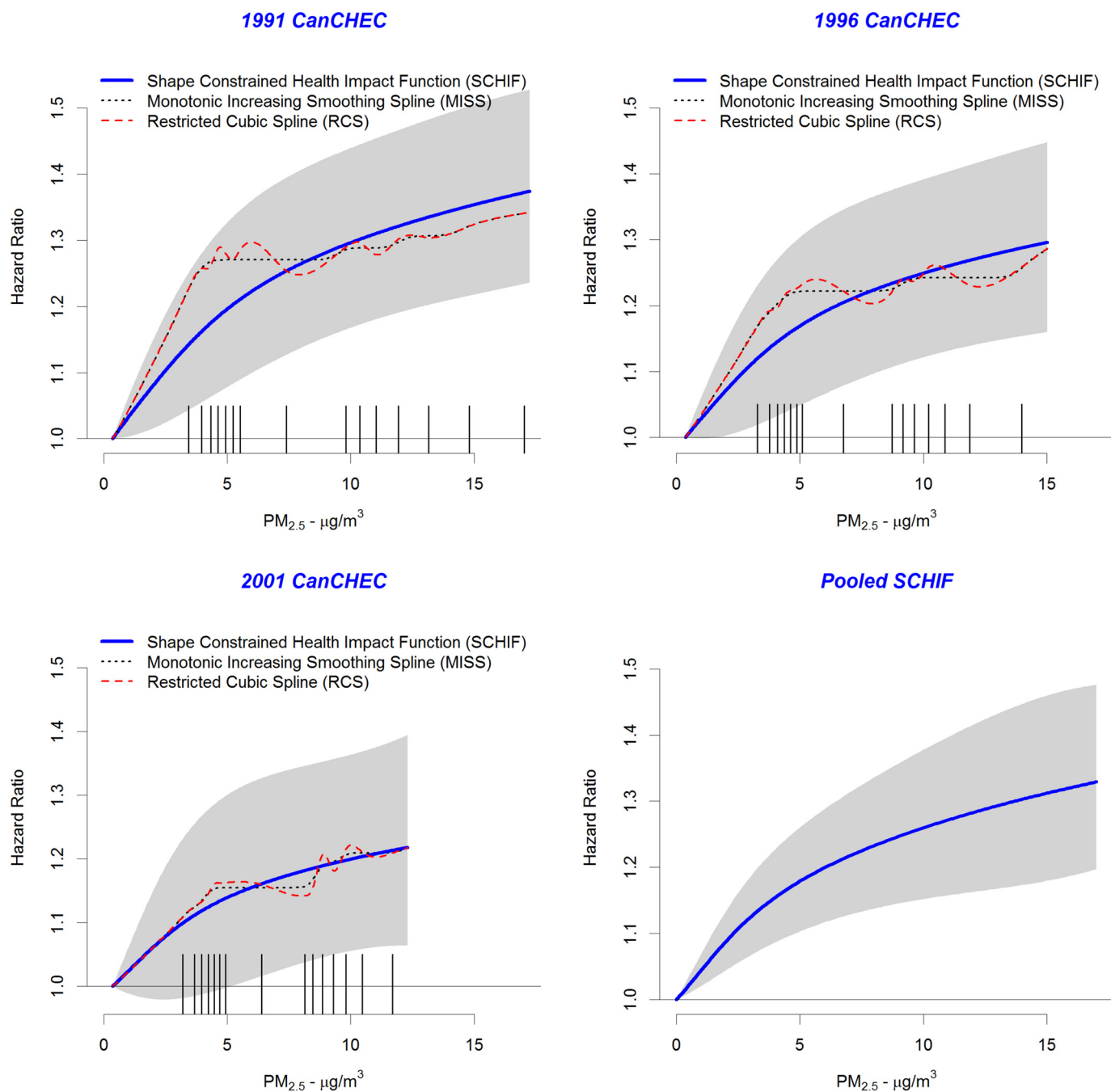
Figure 2. Shape Constrained Health Impact Function (SCHIF) (solid blue line), monotonically increasing smoothing spline (MISS) (dotted black line), and restricted cubic spline (RCS) (dashed red line) predictions by particulate matter with aerodynamic diameter $\leq 2.5\,\mu m$ ($PM_{2.5}$) concentration and Canadian Census Health and Environment Cohort (CanCHEC) (1991, 1996, and 2001). Hazard ratio predictions based on pooling cohort-specific SCHIFs are also presented. Uncertainty bounds are displayed as gray shaded area. Tick marks on $PM_{2.5}$ axis represent the 15-RCS knot locations.

representing one of the largest population-based air pollution cohort analyses conducted to date. We found a HR of 1.053 (95% CI: 1.041, 1.065) per $10\text{-}\mu g/m^3$ change in $PM_{2.5}$ after pooling the three cohort-specific hazard ratios. Hazard ratio estimates based on a LL model do not fully characterize the relationship between $PM_{2.5}$ exposure and mortality, as in each cohort, a supra-linear association was observed (Figure 1).

We found variation in the sensitivity of covariate model specification. The full model yielded larger hazard ratio estimates compared to the DAG model for both the 1991 and 1996 cohorts, with the opposite pattern observed in the 2001 cohort. It is not completely clear why such patterns occur, as the change in

exposure among the categories of the covariates was similar in all three cohorts (Table 1), although the differences in exposure among the categories decreases with more recent cohort start dates due to generally declining concentrations over time.

We observed an increase in the hazard ratio for the immigrant population over time, starting with the weakest association for the 1991 cohort (HR = 1.006; 95% CI: 0.935, 1.081), a slightly stronger association in the 1996 cohort (HR = 1.027; 95% CI: 0.987, 1.068), with the strongest association in the 2001 cohort (HR = 1.109; 95% CI: 1.053, 1.167). We previously also observed no association in the 1991 cohort (Crouse et al. 2015) for immigrants to Canada. There also appeared to be little evidence
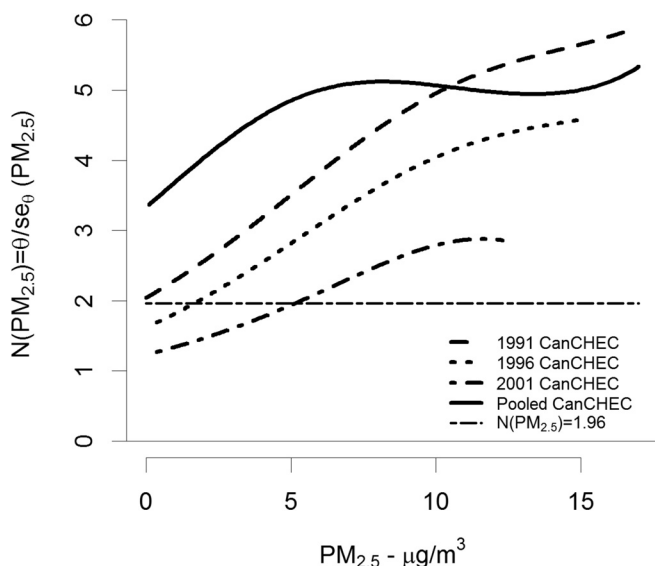
Figure 3. Signal-to-noise ratio, $N(z) = \theta/se_\theta(z)$, by concentration and Canadian Census Health and Environment Cohort (CanCHEC; 1991, 1996, 2001). Dashed-dotted horizontal black line represents a ratio of 1.96.

of an association for females in the 1991 and 1996 cohorts, with much stronger evidence for males. However, the association was similar for males and females in the 2001 cohort (Table 3).

The observation that the $PM_{2.5}$ hazard ratio can be partially explained by NO_2 and fully explained by O_3 also supports previously reported results (Crouse et al. 2015). As these gaseous pollutants may exhibit varying correlations with different components of $PM_{2.5}$, these multipollutant results can provide insight into the source components of $PM_{2.5}$ that are most influential in the relationship between $PM_{2.5}$ mass and mortality. That is, the results of our models with NO_2 and O_3 may, in part, reflect differences in the mixture/composition of $PM_{2.5}$ across space and time. Earlier work has shown that the distribution of components, such as sulfate, organic matter, and black carbon, vary by total mass concentration in Canada (van Donkelaar et al. 2019). Using this spatial variation, Crouse et al. (2016) showed that inclusion of components in addition to total mass improved the prediction of mortality in the 1991 cohort. Mortality from ischemic heart disease has been shown to vary by source of $PM_{2.5}$ in the U.S.-based American Cancer Society (ACS) cohort (Thurston et al. 2016). Therefore, differences over time in the component distribution could explain some of the differences in our hazard ratio estimates between cohorts. However, the empirical evidence for such temporal changes in the component distribution is not supported by our

recent modeling efforts (van Donkelaar et al. 2019), which suggest that the proportion of components is relatively stable over time, even as total mass concentrations have clearly declined.

We find a substantially lower hazard ratio estimate based on the LL model than in previous CanCHEC analyses. There are a number of possible explanations for these differences. First, this new version of the cohort includes an improved death and mobility linkage with updated methodology. Second, the period of follow-up in our analysis ranges from 15 to 25 y (up to 2016). Our analysis further differs from previous CanCHEC studies (e.g., Crouse et al. 2012; Pinault et al. 2017; Weichenthal et al. 2017) in our inclusion of immigrants in the analytical cohort. Immigrants generally have lower mortality rates than the Canadian-born population due to screening criteria for immigration into Canada (Newbold 2005; Ng 2011). Immigrants also tend to live in cities with higher $PM_{2.5}$ concentrations, leading to different risks of air pollution exposure. Our methods of $PM_{2.5}$ averaging at the PC level and imputation differed slightly from previous analyses. Specifically, we required a minimum of a two-digit PC (e.g., K1) to assign a $PM_{2.5}$ value based on a population-weighted average of the geographic area. Some previous analyses assigned the national average $PM_{2.5}$ value in absence of any PC information. We also removed all person-years that resulted in missing environmental ($PM_{2.5}$, O_3, or NO_2) or contextual covariates that may have informed previous analyses, while in previous analyses, we either imputed all missing person-years or included dummy variables for geographically based predictors. Finally, these analyses employ updated $PM_{2.5}$ estimates for the 2012–2016 period based on enhancements to the satellite retrievals (van Donkelaar et al. 2019) and backcasted estimates prior to 2000 that were not available previously. We do note, however, that the relationship between $PM_{2.5}$ and mortality, in both shape and magnitude, is similar to that previously reported for the 1991 (Crouse et al. 2015) and 2001 (Pinault et al. 2017) cohorts. This observation suggests that caution is required in interpreting the magnitude of an association solely based on the LL model without further consideration of the shape of the association.

This work supports observations of a supralinear concentration–response between $PM_{2.5}$ exposure and risk of nonaccidental mortality similar to that reported in Crouse et al. (2012, 2015) for the 1991 cohort and Pinault et al. (2017) for the 2001 cohort in more limited analyses. Most of the previous major cohort studies on $PM_{2.5}$ did not examine the shape of the concentration–mortality association at the low levels that are observed in our cohort, in which the 25th percentile is $5.1 \, \mu g/m^3$, and the median is $6.9 \, \mu g/m^3$ among all person-years of the three cohorts combined. For example, the lower end of the exposure distribution in the Medicare cohort was $6.2 \, \mu g/m^3$ (minimum; Di et al. 2017), the National Health Interview Survey cohort (NHIS) was $7.6 \, \mu g/m^3$ (minimum; Pope et al. 2018), and the ACS Cancer Prevention Study II cohort was $6.7 \, \mu g/m^3$ (first percentile; Turner et al. 2016). Our mortality HR estimate of 1.053 (95% CI: 1.041, 1.065) is slightly below estimates of U.S. cohorts such as the Medicare cohort (HR = 1.073; 95% CI: 1.071, 1.075), NHIS cohort (HR = 1.056; 95% CI: 1.005, 1.110), and the ACS cohort (HR = 1.07; 95% CI: 1.06, 1.09). Our estimate for the 2001 cohort—the only CanCHEC cohort that did not require backcasted exposure data—was slightly higher than the U.S. cohorts at 1.084 (95% CI: 1.060, 1.108).

Strengths and Limitations

Our study has a number of strengths and limitations. These analyses include a large number of deaths (nearly 1.5 million); population representativeness; a number of subject-level mortality predictors, such as occupation, income, and education, that are

Table 4. Shape-constrained health impact function and standard error parameter estimates by Canadian Census Health and Environment Cohort (CanCHEC) cohorts.

Cohort	Shape-constrained health impact function parameters			
	θ	α	μ	ν
1991	0.1102	1	0	1.688
1996	0.0942	1	0	1.465
2001	0.0703	1	0	1.193
Pooled	0.1126	1.477	−0.233	1.165
	Standard error parameters			
	$\sigma_0 \times 10^{-2}$	$\sigma_1 \times 10^{-3}$	$\sigma_2 \times 10^{-4}$	$\sigma_3 \times 10^{-6}$
1991	5.639	−6.608	4.148	−9.179
1996	5.835	−6.942	4.411	−9.747
2001	6.267	−5.330	1.183	6.406
Pooled	3.383	−3.640	3.593	−11.13

often not available in such large administrative-based cohorts (Di et al. 2017); and annual location information through linkage to income tax filings back to 1981. Our temporally resolved air pollution estimates could then be assigned to each year of follow-up. We also developed a method to combine the flexibility of splines to characterize the shape of the relationship between $PM_{2.5}$ and mortality with the usefulness of the SCHIF in risk and benefits assessments. This new approach is feasible with respect to computing resources for very large cohort studies.

Despite these important strengths, the study includes a number of limitations. Specifically, the cohorts lack individual information on behavioral mortality risk factors such as smoking, diet, and obesity. We have previously examined the influence of further adjustment for these missing risk factors using statistical methods of indirect adjustment (Crouse et al. 2015; Erickson et al. 2019; Shin et al. 2014) and direct adjustment in a cohort based on annual health surveys in Canada where such information was available (Pinault et al. 2016b). For both of these approaches, we found that the $PM_{2.5}$–mortality association remained positive. These observations indicate that it is unlikely that the positive associations we observed are entirely due to lack of inclusion of missing risk factors.

We captured information on social and economic status of each subject only at cohort inception. Thus, information on marital status, income, education, occupation, and employment may have changed over time, and we were not able to model any potential influence of such changes on the $PM_{2.5}$–mortality association. The 1991 cohort may have been most influenced by this lack of temporal adjustment given its 25 y of follow-up.

As we examined a 3-y moving-average $PM_{2.5}$ exposure window lagged by 1 y in all analyses, for the 2001 cohort, the initial exposure window assigned to the 2001 follow-up year was a 1998–2000 average and thus did not include backcasted exposures. The 1991 cohort used more years of backcasted exposures (1988–1997) than the 1996 cohort (1993–1997). Backcasted predictions may have induced greater exposure error in the two earlier cohorts.

We employed three very different exposure models for $PM_{2.5}$, NO_2, and O_3. The $PM_{2.5}$ exposure model included remote sensing information coupled with CTM predictions. Land-use information informed any bias in the CTM predictions for the chemical components of total mass, but were not included as direct predictors of ground measurement data. The NO_2 model used both land-use and remote sensing information, while the O_3 model fused ground data with CTM predictions. The spatial resolution of the models was also different, with NO_2 having the finest resolution of 100 m and $PM_{2.5}$ at 1 km, while the O_3 model was at a resolution of 21 km. It was therefore difficult to directly compare the associations with mortality between pollutants and even more difficult with multiple-pollutant models subject to potential differential exposure error. However, the observed stronger associations with O_3 may be due to a lower exposure error in both the original surfaces and the temporal adjustments, since O_3 is known to be more spatially homogenous than $PM_{2.5}$, while NO_2 is known to have the largest spatial variation. Additionally, although the quantitative estimates presented here reflect the large and diverse geographic area of Canada, they may not apply to places around the world with notably different sources and compositions of $PM_{2.5}$, for example, in areas where a much higher proportion of the $PM_{2.5}$ is from dust and sand.

In summary, we found positive associations between $PM_{2.5}$ and mortality in all three cohorts, and a similar supralinear concentration–response relationship. Lower uncertainty bounds were elevated above unity for all concentrations in the 1991 cohort, above $2 \mu g/m^3$ in the 1996 cohort, and above $5 \mu g/m^3$ in the 2001 cohort, suggesting our confidence in identifying concentrations where

there exists a positive response is declining as concentrations decline over time. Therefore, interest exists in examining subsequent CanCHEC cohorts as they become available with the administration of more recent long form censuses.

References

Apte JS, Marshall JD, Cohen A, Brauer M. 2015. Addressing global mortality from ambient $PM_{2.5}$. Environ Sci Technol 49(13):8057–8066, PMID: 26077815, https://doi.org/10.1021/acs.est.5b01236.

Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B, et al. 2014. 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(9919):785–795, PMID: 24332274, https://doi.org/10.1016/S0140-6736(13)62158-3.

Bérard-Chagnon J. 2017. Comparison of Place of Residence between the T1 Family File and the Census: Evaluation using record linkage. Ottawa, Canada: Statistics Canada, Catalogue no. 91F0015M — No. 13.

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 USA 115(38):9592–9597, PMID: 30181279, https://doi.org/10.1073/pnas.1803222115.

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(4):397–403, PMID: 24518036, https://doi.org/10.1289/ehp.1307049.

CCME (Canadian Council of Ministers of the Environment). 2012. *Guidance Document on Achievement Determination Canadian Ambient Air Quality Standards for Fine Particulate Matter and Ozone*. 978-1-896997-91-9 PDF. Winnipeg, Manitoba, Canada: Canadian Council of Ministers of the Environment.

Christidis T, Labrecque-Synnott F, Pinault L, Saidi A, Tjepkema M. 2018. *The 1996 CanCHEC: Canadian Census Health and Environment Cohort Profile*. Canada No. 11-633-X no. 013. Ottawa, Ontario, Canada: Statistics Canada.

Cochran WG. 1950. The comparison of percentages in matched samples. Biometrika 37(3–4):256–266, PMID: 14801052, https://doi.org/10.1093/biomet/37.3-4.256.

Cox DR. 1972. Regression models and life tables. J R Stat Soc B 34(2):187–202, https://doi.org/10.1111/j.2517-6161.1972.tb00899.x.

Crouse DL, Peters PA, Hystad P, Brook JR, van Donkelaar A, Martin RV, et al. 2015. Ambient $PM_{2.5}$, O_3, and NO_2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). Environ Health Perspect 123(11):1180–1186, PMID: 26528712, https://doi.org/10.1289/ehp.1409276.

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(5):708–714, PMID: 22313724, https://doi.org/10.1289/ehp.1104049.

Crouse DL, Philip S, van Donkelaar A, Martin RV, Jessiman B, Peters PA, et al. 2016. A new method to jointly estimate the mortality risk of long-term exposure to fine particulate matter and its components. Sci Rep 6(1):18916, https://doi.org/10.1038/srep18916.

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(26):2513–2522, PMID: 28657878, https://doi.org/10.1056/NEJMoa1702747.

ECCC (Environment and Climate Change Canada). 2017. *Canadian Environmental Sustainability Indicators: Air Pollutant Emissions*. En4-144/22-2017E-PDF. Gatineau, Quebec, Canada: Environment and Climate Change Canada.

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, PMID: 31108354, https://doi.org/10.1016/j.envres.2019.05.010.

Finès P, Pinault L, Tjepkema M. 2017. *Imputing Postal Codes to Analyze Ecological Variables in Longitudinal Cohorts: Exposure To Particulate Matter in the Canadian Census Health and Environment Cohort Database*. Catalogue no. 11-633-X—No. 006. Ottawa, Ontario, Canada: Statistics Canada.

GBD 2017 Risk Factors Collaborators. 2018. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 392(10159):1923–1994, PMID: 30496105, https://doi.org/10.1016/S0140-6736(18)32225-6.

Gordon DLA, Janzen M. 2013. Suburban nation? Estimating the size of Canada's suburban population. J Archit Plann Res 30(3):197–220.

Harrell FE. 2015. *Regression Modeling Strategies with Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis.* 2nd ed. New York, NY: Springer.

Hystad P, Setton E, Cervantes A, Poplawski K, Deschenes S, Brauer M, et al. 2011. Creating national air pollution models for population exposure assessment in Canada. Environ Health Perspect 119(8):1123–1129, PMID: 21454147, https://doi.org/10.1289/ehp.1002976.

Khan S, Pinault L, Tjepkema M, Wilkins R. 2018. Positional accuracy of geocoding from residential postal codes versus full street addresses. Health Rep 29(2):3–9, PMID: 29465738.

Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380(9859):2224–2260, PMID: 23245609, https://doi.org/10.1016/S0140-6736(12)61766-8.

Matheson FI, Dunn JR, Smith KLW, Moineddin R, Glazier RH. 2012. Development of the Canadian Marginalization Index: a new tool for the study of inequality. Can J Public Health 103(8 Suppl 2):S12–S16, PMID: 23618065, https://doi.org/10.1007/BF03403823.

Meng J, Li C, Martin RV, van Donkelaar A, Hystad P, Brauer M, et al. 2019. Estimated long-term (1981–2016) concentrations of ambient fine particulate matter across North America from chemical transport modeling, satellite remote sensing, and ground-based measurements. Environ Sci Technol 53(9):5071–5079, PMID: 30995030, https://doi.org/10.1021/acs.est.8b06875.

Nasari MM, Szyszkowicz M, Chen H, Crouse D, 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(8):961–972, PMID: 27867428, https://doi.org/10.1007/s11869-016-0398-z.

Newbold KB. 2005. Self-rated health within the Canadian immigrant population: risk and the healthy immigrant effect. Soc Sci Med 60(6):1359–1370, PMID: 15626530, https://doi.org/10.1016/j.socscimed.2004.06.048.

Ng E. 2011. The healthy immigrant effect and mortality rates. Health Rep 22(4):25–29, PMID: 22352149.

Pinault LL, Weichenthal S, Crouse DL, Brauer M, Erickson A, Donkelaar AV, et al. 2017. Associations between fine particulate matter and mortality in the 2001 Canadian Census Health and Environment Cohort. Environ Res 159:406–415, PMID: 28850858, https://doi.org/10.1016/j.envres.2017.08.037.

Pinault L, Finès P, Labrecque-Synnott F, Saidi A, Tjepkema M. 2016a. *The 2001 Canadian Census-Tax-Mortality Cohort: A 10-Year Follow-Up*. No. 11-633-X no. 003. Ottawa, Ontario, Canada: Statistics Canada.

Pinault L, Tjepkema M, Crouse DL, Weichenthal S, van Donkelaar A, Martin RV, et al. 2016b. Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the Canadian community health survey cohort. Environ Health 15(1):18–31, PMID: 26864652, https://doi.org/10.1186/s12940-016-0111-6.

Pope CA 3rd, Burnett RT, Krewski D, Jerrett M, Shi Y, Calle EE, et al. 2009. Cardiovascular mortality and exposure to airborne fine particulate matter and cigarette smoke: shape of the exposure-response relationship. Circulation 120(11):941–948, PMID: 19720932, https://doi.org/10.1161/CIRCULATIONAHA.109.857888.

Pope CA 3rd, Burnett RT, Turner MC, Cohen A, Krewski D, Jerrett M, et al. 2011. Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure-response relationships. Environ Health Perspect 119(11):1616–1621, PMID: 21768054, https://doi.org/10.1289/ehp.1103639.

Pope CA 3rd, Ezzati M, Cannon JB, Allen RT, Jerrett M, Burnett RT III. 2018. Mortality risk and PM$_{2.5}$ air pollution in the USA: an analysis of a national prospective cohort. Air Qual Atmos Health 11(3):245–252, https://doi.org/10.1007/s11869-017-0535-3.

Pya N, Wood SN. 2015. Shape constrained additive models. Stat Comput 25(3):543–559, https://doi.org/10.1007/s11222-013-9448-7.

Robichaud A, Ménard R. 2014. Multi-year objective analyses of warm season ground-level ozone and PM$_{2.5}$ over North America using real-time observations and Canadian operational air quality models. Atmos Chem Phys 14(4):1769–1800, https://doi.org/10.5194/acp-14-1769-2014.

Robichaud A, Ménard R, Zaïtseva Y, Anselmo D. 2016. Multi-pollutant surface objective analyses and mapping of air quality health index over North America. Air Qual Atmos Health 9(7):743–759, PMID: 27785157, https://doi.org/10.1007/s11869-015-0385-9.

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, PMID: 24972508, https://doi.org/10.1016/j.envres.2014.05.016.

Statistics Canada. 2003. *2001 Census Dictionary*. 92-378-XIE. Ottawa, Ontario, Canada: Statistics Canada.

Statistics Canada. 2017a. Directive on Microdata Linkage. https://www.statcan.gc.ca/eng/record/policy4-1 [accessed 8 February 2019].

Statistics Canada. 2017b. Social Data Linkage Environment (SDLE). https://www.statcan.gc.ca/eng/sdle/index [accessed 28 January 2019].

Statistics Canada. 2017c. *Postal Code^{OM} Conversion File Plus (PCCF+) Version 7A, Reference Guide: June 2017 Postal Codes^{OM}*. Ottawa, Ontario, Canada: Statistics Canada.

Thurston GD, Burnett RT, Turner MC, Shi Y, Krewski D, Lall R, et al. 2016. Ischemic heart disease mortality and long-term exposure to source-related components of U.S. fine particle air pollution. Environ Health Perspect 124(6):785–794, PMID: 26629599, https://doi.org/10.1289/ehp.1509777.

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. Am J Respir Crit Care Med 193(10):1134–1142, PMID: 26680605, https://doi.org/10.1164/rccm.201508-1633OC.

van Donkelaar A, Martin RV, Spurr RJD, Burnett RT. 2015. High resolution satellite-derived PM$_{2.5}$ for optical estimation and geographically-weighted regression over North America. Environ Sci Technol 49(17):10482–10491, PMID: 26261937, https://doi.org/10.1021/acs.est.5b02076.

van Donkelaar A, Martin RV, Li C, Burnett RT. 2019. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. Environ Sci Technol 53(5):2595–2611, PMID: 30698001, https://doi.org/10.1021/acs.est.8b06392.

Weichenthal S, Pinault LL, Burnett RT. 2017. Impact of oxidant gases on the relationship between outdoor fine particulate air pollution and nonaccidental, cardiovascular, and respiratory mortality. Sci Rep 7(1):16401, PMID: 29180643, https://doi.org/10.1038/s41598-017-16770-y.

WHO (World Health Organization). 1977. *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death: Based on the Recommendations of the Ninth Revision Conference, 1975, and Adopted by the Twenty-Ninth World Health Assembly, 1975 Revision*. http://www.who.int/iris/handle/10665/40492 [accessed 12 February 2019].

WHO. 2016. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*. https://www.who.int/classifications/icd/ICD10Volume2_en_2010.pdf [accessed 12 February 2019].

Wilkins R, Tjepkema M, Mustard C, Choinière R. 2008. The Canadian census mortality follow-up study, 1991 through 2001. Health Rep 19(3):25–43, PMID: 18847143.

Yin P, Brauer M, Cohen A, Burnett RT, Liu J, Liu Y, et al. 2017. Long-term fine particulate matter exposure and nonaccidental and cause-specific mortality in a large national cohort of Chinese men. Environ Health Perspect 125(11):117002, PMID: 29116930, https://doi.org/10.1289/EHP1673.