Title

Chronic Exposure to Fine Particulate Matter Increases Mortality Through Pathways of Metabolic and Cardiovascular Disease: Insights From a Large Mediation Analysis

Journal

Journal of the American Heart Association, 11(22)

ISSN

2047-9980

Authors

Bai, Li

Benmarhnia, Tarik

Chen, Chen

et al.

Publication Date

2022-11-15

DOI

10.1161/jaha.122.026660

Journal of the American Heart Association

ORIGINAL RESEARCH

Chronic Exposure to Fine Particulate Matter Increases Mortality Through Pathways of Metabolic and Cardiovascular Disease: Insights From a Large Mediation Analysis

Li Bai, PhD; Tarik Benmarhnia, PhD*; Chen Chen, PhD; Jeffrey C. Kwong, MD; Richard T. Burnett, PhD; Aaron van Donkelaar, PhD; Randall V. Martin, PhD; JinHee Kim, MD; Jay S. Kaufman, PhD; Hong Chen, PhD*

BACKGROUND: Long-term exposure to outdoor fine particulate matter ($PM_{2.5}$) is the leading environmental risk factor for premature mortality worldwide. Characterizing important pathways through which $PM_{2.5}$ increases individuals' mortality risk can clarify the $PM_{2.5}$–mortality relationship and identify possible points of interventions. Recent evidence has linked $PM_{2.5}$ to the onset of diabetes and cardiovascular disease, but to what extent these associations contribute to the effect of $PM_{2.5}$ on mortality remains poorly understood.

METHODS AND RESULTS: We conducted a population-based cohort study to investigate how the effect of $PM_{2.5}$ on nonaccidental mortality is mediated by its impacts on incident diabetes, acute myocardial infarction, and stroke. Our study population comprised ≈200 000 individuals aged 20 to 90 years who participated in population-based health surveys in Ontario, Canada, from 1996 to 2014. Follow-up extended until December 2017. Using causal mediation analyses with Aalen additive hazards models, we decomposed the total effect of $PM_{2.5}$ on mortality into a direct effect and several path-specific indirect effects mediated by diabetes, each cardiovascular event, or both combined. A series of sensitivity analyses were also conducted. After adjusting for various individual- and neighborhood-level covariates, we estimated that for every 1000 adults, each 10 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with ≈2 incident cases of diabetes, ≈1 major cardiovascular event (acute myocardial infarction and stroke combined), and ≈2 deaths annually. Among $PM_{2.5}$-related deaths, 31.7% (95% CI, 17.2%–53.2%) were attributable to diabetes and major cardiovascular events in relation to $PM_{2.5}$. Specifically, 4.5% were explained by $PM_{2.5}$-induced diabetes, 22.8% by $PM_{2.5}$-induced major cardiovascular events, and 4.5% through their interaction.

CONCLUSIONS: This study suggests that a significant portion of the estimated effect of long-term exposure to $PM_{2.5}$ on deaths can be attributed to its effect on diabetes and cardiovascular diseases, highlighting the significance of $PM_{2.5}$ on deteriorating cardiovascular health. Our findings should raise awareness among professionals that improving metabolic and cardiovascular health may reduce mortality burden in areas with higher exposure to air pollution.

Key Words: air pollutants ■ cardiovascular diseases ■ causality ■ cohort studies ■ mediation analysis ■ mortality, premature ■ population

Outdoor fine particulate matter ($PM_{2.5}$) is a leading global health concern.[1] A recent international study estimated that 8.9 million deaths globally may be attributable to outdoor $PM_{2.5}$ in 2015.[2] During the past decade, $PM_{2.5}$ exposure has also been increasingly linked to the incidence of diabetes[3–5] and major cardiovascular

Correspondence to: Li Bai, PhD, G177, 2075 Bayview Avenue, Toronto, Ontario M4N3M5, Canada. Email: li.bai@ices.on.ca

*T. Benmarhnia and H. Chen are co-senior authors.

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.122.026660

For Sources of Funding and Disclosures, see page 12.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- This is the first causal mediation study to date that empirically assessed pathways from long-term exposure to outdoor fine particulate matter to nonaccidental deaths through the incidence of diabetes and cardiovascular diseases and thereby quantified the relative contributions of these potential mediators.
- A third of the estimated adverse effects of exposure to fine particulate matter on mortality operates through its influence on the development of diabetes and cardiovascular events.

What Are the Clinical Implications?

- Beyond interventions to reduce exposure to fine particulate matter, there is an opportunity in clinical practice to help mitigate the adverse effects of fine particulate matter before they become overwhelming and irreversible. This can be achieved by improving metabolic and cardiovascular health in the polluted areas.

Nonstandard Abbreviations and Acronyms

NDE	natural direct effect
NIE	natural indirect effect
$PM_{2.5}$	fine particulate matter

events including acute myocardial infarction (AMI)[6,7] and stroke.[8–10] For example, a recent multicountry cohort study in Europe has shown that even lower levels of air pollution (ie, concentrations lower than the existing World Health Organization guideline limits) were associated with increased incidences of stroke and coronary heart disease.[10] The biological mechanisms underlying the $PM_{2.5}$ effects on these incident outcomes and mortality were hypothesized to include aggravating oxidative stress and inflammation, altering endothelial function, inducing insulin resistance, increasing blood coagulability, accelerating atherosclerosis progression, impairing autonomic balance, and increasing sympathetic tone, all of which can precipitate the development of metabolic diseases such as diabetes and cardiovascular events, ultimately increasing the risk of premature death.[11,12] Despite the hypothesized mechanisms, there is a dearth of empirical evidence about the extent to which the effect of $PM_{2.5}$ on deaths is mediated by its effect on the development of these conditions (ie, the proportion mediated). In the face of continuing mortality burden attributed to $PM_{2.5}$, it is crucial to elucidate the mechanistic structure that may provide valuable insight on the reduction of $PM_{2.5}$-related mortality (eg, intervention on mediators when it is not possible to intervene on the exposure in a timely manner).[13]

Causal mediation analysis is a valuable tool for epidemiologic research to assess how an exposure can affect an outcome of interest through ≥1 intermediate variables (ie, mediators) on complex pathways.[14–17] It reveals new insight toward understanding disease pathogenesis using observational data when large clinical trials to deduce mechanistic pathways are not feasible. Such information can greatly improve our knowledge of the health effects of an exposure by allowing us to suggest and test hypotheses about underlying mechanisms, which will help prioritize intervention targets. Furthermore, recent developments in statistical methods for the mediation analysis allow to study more complex causal pathways, such as when multiple mediators affect each other or when interactions are present.[15,18] However, mediation analysis has not been widely applied in air-health studies, especially for such complex causal structures. So far, relatively few of these studies have evaluated mediation effects in relation to air pollution. None of them has considered mediation by different diseases and disease states nor their interactions,[19–25] therefore it remains elusive to prioritize each potential pathway. Given the pervasive air pollution exposure and the dynamic nature of cardiometabolic health during the life course, it is crucial to determine how air pollution shapes individuals' health trajectories from the onset of disease through its progression and eventually to death. This, however, has not been addressed by previous mediation analyses.[19–25]

We conducted a large, population-based mediation analysis to investigate the onset of diabetes and cardiovascular events as potential mediating pathways between long-term exposure to $PM_{2.5}$ and mortality. Specifically, we estimated the extent to which the effect of $PM_{2.5}$ on nonaccidental deaths was mediated by the following 3 sets of sequential time-varying mediators: (1) incidence of diabetes and AMI, (2) incidence of diabetes and stroke, and (3) incidence of diabetes and major cardiovascular events (AMI and stroke combined). Our analysis improved on existing work by using causal mediation approaches that allow for causally ordered mediators to be a time-varying feature in the context of a time-to-event outcome. In addition, the use of population-based health surveys with rich data on key risk factors offers unique opportunities to elucidate these knowledge gaps.

METHODS

The data set from this study is held securely in coded form at ICES. Although legal data-sharing agreements

between ICES and data providers (eg, health care organizations and government) prohibit ICES from making the data set publicly available, access may be granted to those who meet prespecified criteria for confidential access, available at https://www.ices.on.ca/DAS (email: das@ices.on.ca).

Study Participants

In this retrospective, population-based cohort study, our population was derived from the 1996 to 1997 cycle of the National Population Health Survey[26] and the 2000 to 2001, 2003, 2005, 2007 to 2008, 2009 to 2010, 2011 to 2012, and 2013 to 2014 cycles of the Canadian Community Health Survey (Figure 1; n=254 965 individuals).[27] Details about the methodology of these surveys have been described elsewhere (see Table S1 for time periods of data collection).[26,27] Briefly, the National Population Health Survey and Canadian Community Health Surveys are nationally representative surveys that collect self-reported data related to health status and determinants of health (eg, sociodemographic factors) from a representative sample of people aged ≥12 years who are living in private dwellings in Canada.

We included the respondents who, at the time of survey, resided in the province of Ontario, were aged 20 to 90 years, were not pregnant, were eligible for Ontario's provincial health insurance plan, and provided informed consent to share and link their responses to provincial health administrative data (n=218 158). Eligible respondents were followed up from the time of survey and censored when reaching the end of follow-up (December 31, 2017), becoming ineligible for provincial health insurance, or death.

The use of the data in this study is authorized under section 45 of Ontario's Personal Health Information Protection Act and does not require review by a research ethics board.

Outcome

Consistent with previous studies,[2] we considered nonaccidental death as our outcome. Using unique encoded identifiers (encrypted health card numbers), we anonymously linked the National Population Health Survey and Canadian Community Health Survey respondents to Ontario health administrative databases. We ascertained nonaccidental deaths from the provincial death registry, Ontario's Registrar General Death File (see Data S1). The nonaccidental death was determined based on the underlying cause of death listed on the death certificates.

$PM_{2.5}$ Exposure Assessment

We estimated ground-level $PM_{2.5}$ concentrations by relating satellite retrievals of aerosol optical depth to $PM_{2.5}$

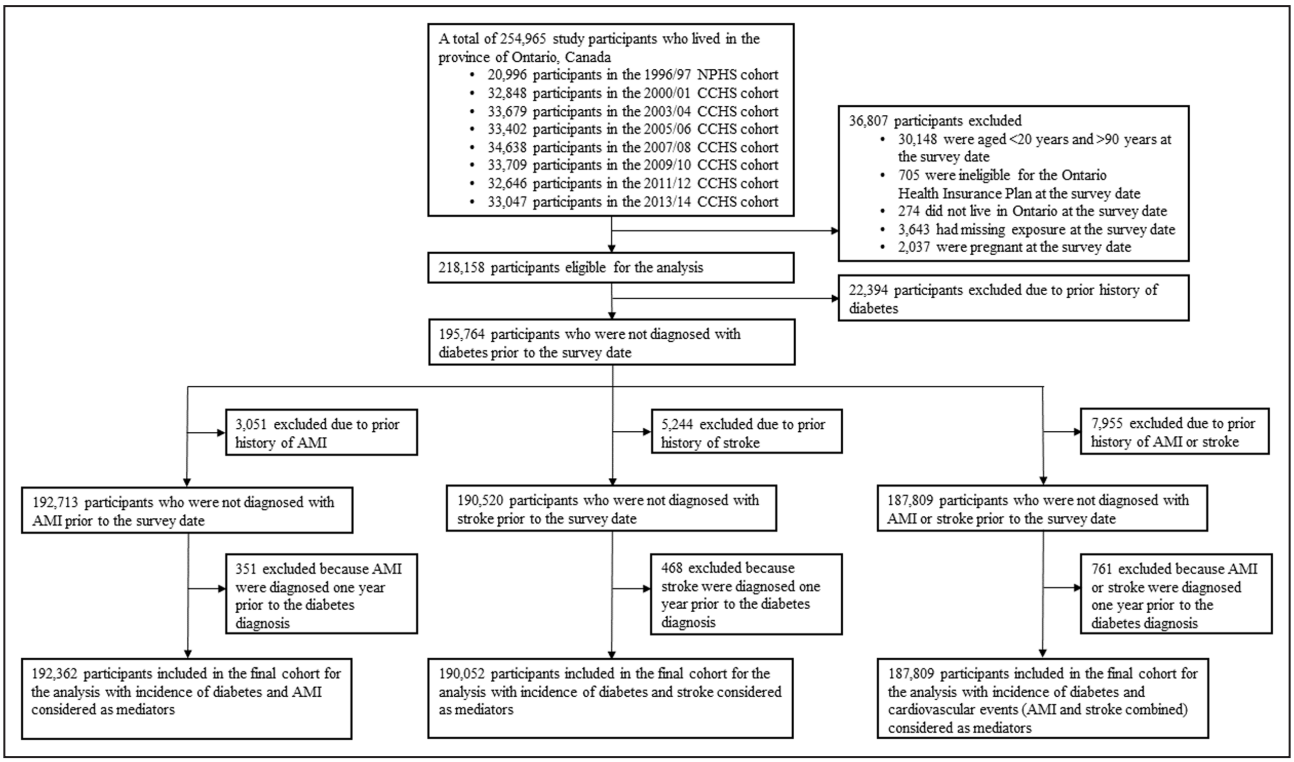

Figure 1. Flowchart of participant selection.
AMI indicates acute myocardial infarction; CCHS, Canadian Community Health Surveys; and NPHS, National Population Health Survey.

using a global atmospheric chemistry transport model combined with a geographically weighted regression model.[28] The $PM_{2.5}$ estimates were available on a grid with a spatial resolution of ≈1 km×1 km for each year between 2000 and 2016. Briefly, satellite retrievals of aerosol optical depth were related to near-surface $PM_{2.5}$ concentrations using the geophysical relationship predicted by a chemical transport model and subsequently calibrated via geographically weighted regression. These results were validated against ground $PM_{2.5}$ concentrations measured at fixed monitoring stations (n=2312) across North America and showed good cross-validated performance (R^2 =0.70). Recent methodological modifications were made (eg, treating topographical changes and urban land cover as separate predictors), which further improved the performance (R^2 =0.73).

Because the $PM_{2.5}$ data are only available for 2000 to 2016, we conducted annual calibration of the surfaces to relevant time periods during the study, similar to previous studies.[29,30] Briefly, we rescaled the annual estimate of $PM_{2.5}$ in 2000 to years from 1991 to 1999 by taking the ratio of the 2000 surface to the average concentration from 1991 to 1999 of $PM_{2.5}$ at all fixed-site monitors across Ontario. We also estimated the concentrations of $PM_{2.5}$ in 2017 by scaling the data in 2016. We were thus able to assign annual estimates of $PM_{2.5}$ exposure from 1991 to 2017 to the centroid of each respondent's residential area determined by the annual 6-character postal code in that year, therefore accounting for residential mobility and long-term trends in exposure. In Canada, a 6-character postal code represents a block face or a large building in urban areas, but much larger areas in rural areas. For each year of follow-up of a participant, we estimated a 5-year moving window of past exposures to $PM_{2.5}$ with a 1-year lag, as was done in previous studies.[31] For example, a respondent's moving window of exposures for 2001 would be computed as the mean concentrations from 1996 to 2000.

Mediators

The mediators of interest in this study are the incidence of diabetes, AMI, and stroke and a composite of the incidence of AMI and stroke (referred to as major cardiovascular events). Using algorithms previously validated against patient charts, we ascertained the incidence of these conditions during follow-up based on health administrative databases in Ontario including hospital discharge abstracts from the Canadian Institute for Health Information (see Data S2), physician service claims from the Ontario Health Insurance Plan database, or claims for prescription drugs from the Ontario Drug Benefit database. An incident case of diabetes was defined as an individual with ≥2 physician claims with a diabetes diagnostic code (250), ≥1 drug claim for diabetes, or ≥1 hospitalization for diabetes within 1 year (*International Classification of Diseases, Ninth Revision* [*ICD-9*] code 250 and *International Classification of Diseases, Tenth Revision* [*ICD-10*] codes E10–E14). This validated algorithm has a high sensitivity (90.0%) and specificity (97.7%).[32] Incident AMI was defined as having ≥1 hospitalization with AMI and having had no hospitalization for AMI in the previous 1 year (*ICD-9* code 410 and *ICD-10* code I21). A previous study has shown high accuracy of coding of AMI, with a sensitivity of 89% and a specificity of 93%.[33] To ascertain incident stroke, we used an algorithm of ≥2 physician claims with diagnostic codes for stroke (436, 432) or transient ischemic attack (435) or ≥1 hospitalization for stroke or transient ischemic attack within 1 year (*ICD-9* codes 362.3, 430, 431, 434.x, 436, and 435.x and *ICD-10* codes I60.x, I61.x, I63.x [excluding I63.6 cerebral infarction attributed to central venous thrombosis], I64, H34.1, G45.x [excluding G45.4 transient global amnesia], and H34.0).[34] This algorithm for the stroke definition has been found to have a sensitivity of 85% and a specificity of 97%.[34] The databases have been also used to identify people who were diagnosed with diabetes, AMI, or stroke before the survey date.

Covariates

Covariates were identified as potential confounders on the basis of a priori assumptions of their relationships with exposure ($PM_{2.5}$) and the mediators (diabetes, AMI, or stroke) and outcome (nonaccidental death) under investigation.[35–37] We considered the following 3 types of confounders: exposure–outcome, exposure–mediators and mediators–outcome.[15] The covariates were considered at both individual- and neighborhood area–levels. Individual-level covariates were identified at baseline only. Sociodemographic characteristics included age, sex, marital status (married or common law, never married, separated, widowed, or divorced), ethnicity (White/non-White race [including Black, Korean, Filipino, Japanese, Chinese, South Asian, South East Asian, Arab, West Asian, Latin American, and others]), nativity (immigrants/nonimmigrants), education (less than high school, high school, some postsecondary, and postsecondary graduation), and household income level (≤$29 999, $30 000–$79 999, and ≥$80 000). Lifestyle factors included smoking status (never smoker, daily smoker, occasional smoker, always occasional smoker, former daily smoker, and former occasional smoker), smoking pack years (available in the Canadian Community Health Survey only), body mass index, and physical activity (based on energy expenditure).

We also created a series of time-varying contextual variables at census division and dissemination levels,

respectively, by linking the survey data to the 1996, 2001, 2006, and 2016 Canadian Census data based on the nearest census year: percentage of the population aged ≥15 years with less than a high school education, percentage of the population aged ≥15 years who are recent immigrants, income quintile (a measure of relative household income accounting for household size and community), and unemployment rate.[38] In addition, we created 2 time-varying geographic indicators of residence (ie, northern/southern Ontario and urban/rural areas) based on the participants' yearly postal code information to control for possible regional differences in the outcome and identification of mediators. Furthermore, we included survey cycles as a covariate.

We obtained information on several additional covariates at the individual-level such as alcohol consumption (regular, occasional, former, and never drinkers), working status (having a job last week, no job last week, and unable/permanent disability), and sense of belonging to local community (very strong, somewhat strong, somewhat weak, and very weak). A summary of all covariates is provided in Table S2.

All data sets were linked using unique encoded identifiers and analyzed at ICES. ICES is an independent, nonprofit research institute whose legal status under Ontario's health information privacy law allows it to collect and analyze health care and demographic data without consent for health system evaluation and improvement.

Statistical Analysis

We conducted a causal mediation analysis to quantify to what extent the estimated effect of PM$_{2.5}$ on nonaccidental deaths might be mediated through its effect on 2 causally related intermediates (and their interaction) during the follow-up (Figure 2). We excluded participants with missing exposure data at baseline (Figure 1; n=3643) (see Data S3). To capture the first-ever incidences of mediators during follow-up, which would facilitate temporality, we excluded respondents who had previous physician diagnoses of diabetes, AMI, and/or stroke at the time of entry (see Data S4). For example, for the analysis with incidence of diabetes and AMI as mediators, respondents with prior diagnoses of diabetes (n=22 394) and AMI (n=3051) were excluded by definition (Figure 1). In addition, we a priori considered the incidence of diabetes as the first mediator and the incidence of AMI, stroke, or major cardiovascular events (ie, AMI and stroke combined) as the second mediator, respectively, because diabetes is known to increase the risk for cardiovascular diseases.[39] Thus, we excluded a small number of respondents who had the second mediator (ie, AMI and/or stroke) diagnosed before the diagnosis of the first mediator (ie, diabetes). This led to an exclusion of 351 (≈0.2%) individuals in the analysis where diabetes and AMI were 2 mediators, for example.

Although the first mediator would causally affect the second mediator, it is also plausible that there could be interactions between the 2 mediators on the mediated effect of PM$_{2.5}$ on mortality. Therefore, for each pair of causally ordered mediators, we created a time-varying joint mediator (a categorical variable) (eg, 1, free of both diabetes and AMI diagnoses during follow-up; 2, developed diabetes only; 3, developed AMI only; and 4, developed both diabetes and AMI). This joint-mediator approach allows us to assess the individual mediation effects as well as their joint effect, which in turn helps elucidate the relative contributions of the mediators (eg, developed either diabetes or AMI) and their interaction (eg, developed both diabetes and AMI).[18] In addition, this approach ensured that none of the confounders for the second mediator–mortality association was affected by the exposure (eg, diabetes).

For the mediation analysis, we applied the method for estimating causal effects described by VanderWeele et al.[15,18] With the identifiability assumptions (eg, no confounding for the exposure–outcome, exposure–mediator, and mediator–outcome relationships), this approach allowed us to decompose the total effect of PM$_{2.5}$ on deaths into the natural direct effect (NDE)

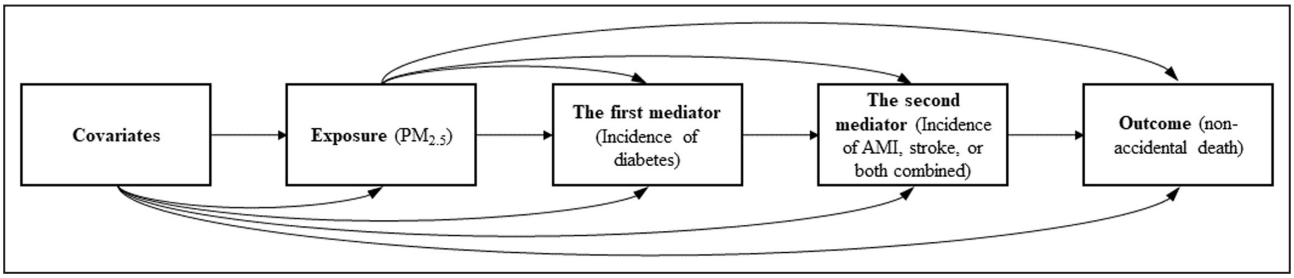

Figure 2. A diagram illustrating the relationships between exposure (PM$_{2.5}$), the first mediator, the second mediator, covariates, and outcome (nonaccidental deaths).
The first mediator was incidence of diabetes. The second mediator was incidence of AMI, stroke, or cardiovascular events (AMI and stroke combined). The confounders included individual-level risk factors at baseline and time-varying area-level factors. AMI indicates acute myocardial infarction; and PM$_{2.5}$, fine particulate matter.

(ie, the estimated effect of $PM_{2.5}$ on deaths via pathways that do not involve mediators) and natural indirect effects (NIEs) (ie, the estimated effect of $PM_{2.5}$ on deaths attributed to the effect of $PM_{2.5}$ on the mediators) (Figure S1). The NIEs were measured through the effect mediated through the first mediator alone, the second mediator alone, and the interactive effect between the 2 mediators. The proportion mediated was calculated as the ratio between the NIE and the total effect.

To implement this approach, we first evaluated the associations between the exposure, the outcome, and the mediators using Aalen additive hazards models (see Data S5 and Data S6), which has been increasingly used in recent mediation studies.[16,40] An advantage of using Aalen additive hazards models is that they can be used with relatively common mediators and outcomes, unlike Cox proportional hazards models.[40] In addition, the Aalen additive hazards models give rise to effect measures on the additive scale rate (ie, rate difference), which is more interpretable from the public health perspective. In all models, we included exposure to $PM_{2.5}$ as a time-varying variable (continuous) and adjusted for the selected individual-level covariates measured at baseline (ie, age, sex, marital status, education, immigration status, household income adequacy, smoking status, smoking pack years, physical activity, and body mass index) and time-varying area-level variables (ie, education, income, percentage of unemployment, percentage of immigrants, indicators for rural/urban and north/south). Then, using the coefficients derived from these models, we calculated NDE, NIEs, and the proportions mediated based on the product coefficient method.[15] We performed bootstrapping based on 500 replications to derive CIs for NDE, NIEs, and the proportions mediated. We reported percentile-based CIs as demonstrated previously.[41]

Sensitivity analyses were performed to test the robustness of these estimates by (1) further incrementally adjusting for additional individual risk factors, including alcohol consumption, working status, and sense of belonging to local community; (2) using the restricted cubic splines of age in the models to allow for the nonlinear effect of age,[42] and (3) further adjusting for a linear term for time to account for potential changes in the diagnosis and disease incidence over time. Because of computational constraints, we only calculated the point estimates of the proportion mediated in these sensitivity analyses.

We also conducted an additional sensitivity analysis to estimate the proportion mediated using an alternative approach developed by Aalen et al.[43] With the framework of dynamic path analysis, Aalen et al combines the g-formula (ie, a causal inference technique)[44] with the additive hazards model and a sequential linear model for the mediator process. It also allows for multiple mediators to be a time-varying feature in the context of survival analysis. Using this method, we derived the proportion mediated by calculating the ratio of the mean cumulative indirect and total effects. Because this method is computationally demanding, we were unable to perform bootstrapping to derive CIs for the point estimates (each bootstrapping replication takes ≈3 days with our large cohort size).

Lastly, to make a comparison with previous studies, we estimated the association between the long-term exposure to $PM_{2.5}$ and nonaccidental deaths using the Cox proportional hazards model with adjustments for all covariates and mediators. To investigate the potential impact of the exclusion of respondents with prior diagnoses of diabetes and cardiovascular events, we also estimated the association between $PM_{2.5}$ and nonaccidental deaths in the cohort with these respondents added back using the Aalen additive model and Cox proportional hazards model without mediators, respectively.

The study cohorts were created using SAS version 9.4 (SAS Institute, Cary, NC); all analyses were done using RStudio and the timereg package.

RESULTS

Participant Characteristics

A total of 192 362 participants were included in the analytical cohort with incidence of diabetes and AMI considered as mediators (referred to as the diabetes–AMI cohort); 190 052 participants were included the analytical cohort with incidence of diabetes and stroke considered as mediators (referred to as the diabetes–stroke cohort), and 187 809 participants were included in the analytical cohort with incidence of diabetes and cardiovascular events (AMI and stroke combined) considered as mediators (referred to as the diabetes–cardiovascular event cohort) (Figure 1). Table 1 summarizes characteristics of the participants in the diabetes–cardiovascular event cohort (characteristics of the participants in the diabetes–AMI and diabetes–stroke cohorts are included in Table S3). For the diabetes–cardiovascular event cohort, the total follow-up time is 2016152.6 person-years. The annual average mortality was ≈10 per 1000 people. For all 3 cohorts, the 5-year moving average exposure to $PM_{2.5}$ at respondents' baseline residences was 8.5 $\mu g/m^3$ (SD, 2.47 $\mu g/m^3$). There had been a decreasing trend in $PM_{2.5}$ during the study period.

At baseline, participant characteristics were broadly similar across the 3 cohorts: the mean ages were ≈50 years, 45% were men, 20% were immigrants, 89% were White race, 86% were urban residents, and 16% had less than a high school diploma. Of the participants, ≈23% to 28% were smokers depending on

Table 1. Demographic Characteristics of the Study Cohort With Incidence of Diabetes and Cardiovascular Events (AMI and Stroke Combined) Analyzed as Mediators

Characteristics	Diabetes–cardiovascular events cohort			
	All participants	Participants who were diagnosed with diabetes during follow-up	Participants who were diagnosed with AMI during follow-up	Participants who died as a result of nonaccidental causes during follow-up
No.	187 048	17 630	11 473	19 267
PM$_{2.5}$ exposure at entry	8.45±2.47	8.70±2.67	8.67±2.76	8.71±2.79
Individual-level risk factors				
Age at entry, y	49.66±17.53	54.80±14.42	64.79±14.30	69.46±12.9
Male sex	44.7	49.0	47.9	54.2
Immigrant	20.2	25.1	23.2	22.6
Non-White race*	11.4	13.1	5.8	4.3
Living in an urban area	86.3	74.0	84.2	72.3
Living in southern region	74.1	84.9	72.7	86.4
Marital status				
Married or common law	59.6	61.7	55.7	48.5
Single (never married)	19.4	13.1	8.8	8.7
Separate, widowed, or divorced	21.0	25.2	35.6	42.9
Education				
Less than high school	16.2	24.8	31.7	37.7
High school graduation	19.5	19.2	18.2	18.3
Some form of postsecondary education	8.0	8.4	7.6	7.5
Postsecondary graduation	56.3	47.6	42.5	36.6
Family income				
≤$29 999	19.9	26.5	32.7	41.0
$30 000 to $79 999	43.2	44.6	41.4	37.4
≥$80 000	29.8	19.7	15.0	9.1
Missing	7.1	9.1	10.9	12.6
Smoking (the NPHS)				
Never smoker	40.6	37.4	39.3	35.2
Daily smoker	25.0	24.9	23.2	25.1
Occasional smoker	2.0	1.9	1.5	1.7
Always occasional smoker	1.5	1.6	0.9	1.0
Former daily smoker	24.0	28.7	29.8	31.4
Former occasional smoker	6.8	5.6	5.3	5.6
Smoking (the CCHS)				
Never smoker	34.0	30.9	29.0	26.9
Current smoker of <10 pack-year	8.5	4.9	2.8	2.1
Current smoker of 10 to 20 pack-year	5.2	4.6	4.6	3.5
Current smoker of ≥20 pack-year	9.1	12.8	14.8	17.5
Current smoker with missing pack-year	0.9	1.0	0.0	1.2
Former smoker who quit within 5 y	6.8	7.5	5.1	6.5
Former smoker who quit >5 y ago	22.2	27.3	32.2	33.2
Missing	13.4	11.1	10.5	9.0
BMI, kg/m^2				
18.5 to 25.0	43.1	18.6	35.8	37.0
25.0 to 30.0	33.8	36.5	34.5	29.1

(*Continued*)

Table 1. (Continued)

Characteristics	Diabetes–cardiovascular events cohort			
	All participants	Participants who were diagnosed with diabetes during follow-up	Participants who were diagnosed with AMI during follow-up	Participants who died as a result of nonaccidental causes during follow-up
30.0 to 35.0	11.9	24.0	12.1	10.1
≥35.0	4.3	12.9	3.7	3.8
<18.5	2.1	0.6	1.8	3.2
Missing	4.8	7.4	12.1	16.7
Physical activity				
Active	24.5	18.2	19.4	16.1
Moderate active	25.6	23.0	23.7	20.4
Inactive	49.9	58.7	56.8	63.5
Area-level risk factor†				
Percentage of recent immigrants	2.56±2.92	2.56±2.99	2.18±2.69	2.17±2.68
Percentage of the population aged ≥15 y without employment	6.98±1.44	7.00±1.42	7.05±1.44	7.04±1.42
Percentage of the population aged ≥15 y with less than a high school education	26.82±4.43	26.77±4.36	27.26±4.10	27.25±4.06
Income quintiles				
Lowest	18.3	21.8	22.4	24.6
Lower	19.6	21.1	20.9	21.4
Middle	19.8	20.6	19.6	19.2
Upper	22.2	19.9	19.4	18.5
Uppermost	20.0	16.6	17.7	16.3

Data are provided as number, mean±SD, or percentage. AMI indicates acute myocardial infarction; BMI, body mass index; CCHS, Canadian Community Health Surveys; NPHS, National Population Health Survey; and $PM_{2.5}$, fine particulate matter.

†Non-white includes Black, Korean, Filipino, Japanese, Chinese, South Asian, South East Asian, Arab, West Asian, Latin American, and others.

‡From Canadian Census 2001, at the dissemination-area level.

the surveys. The neighborhood-level socioeconomic status of the participants was also similar across the cohorts (Table 1). During follow-up, roughly 10% of participants were diagnosed with diabetes, ≈2% developed AMI, ≈5% had a stroke, and ≈11% died from nonaccidental causes. Overall, compared with the entire cohorts, participants who were diagnosed with any of the 3 conditions or who died during follow-up were older and were more likely to be immigrants, have low education, have a history of smoking, and live in the lowest income neighborhoods.

Main Analysis

The associations between $PM_{2.5}$, mediators, and nonaccidental deaths estimated by rate differences using Aalen additive hazard models are presented in Table 2. After adjusting for all covariates, every 10 $\mu g/m^3$ increase in exposure to $PM_{2.5}$ was associated with 1.51 to 1.76 excess deaths for every 1000 person-years, depending on the analytical cohort investigated.

As expected, having diabetes or cardiovascular events (AMI or stroke) was associated with an increased risk for death (Table 2). For every 1000 person-years, the impact on mortality estimated by the rate difference is 4.33 (95% CI, 3.44–5.22) and 47.23 (95% CI, 44.90–49.64) for incident diabetes and cardiovascular events, respectively. Patients who were diagnosed with both of diabetes and cardiovascular events during follow-up were at the highest risk (53.40 [95% CI, 45.26–61.69] excess deaths per 1000 person-years). These estimates were in dependent of the impact of $PM_{2.5}$ on death.

In the models for the associations between $PM_{2.5}$ and mediators, we found that every 10 $\mu g/m^3$ increase in exposure to $PM_{2.5}$ was associated with ≈2 excess cases of diabetes, 0.5 excess cases of AMI, 0.5 excess cases of stroke, and 1 excess case of combined cardiovascular events for 1000 person-years, adjusting for all covariates (Table 2).

Table 3 shows the NDE, NIEs through mediators, total effect of long-term exposure to $PM_{2.5}$ on nonaccidental deaths, and the proportion mediated. In the analysis that decomposed the $PM_{2.5}$–mortality association via a pathway involving the incidence of diabetes and AMI, we estimated that the proportion mediated through incident diabetes and AMI jointly was 18.6% (95% CI, 8.8%–32.1%): 4.2% (95% CI, 1.9%–7.3%)

Table 2. Rate Differences and 95% CIs for the Associations Between $PM_{2.5}$, Nonaccidental Deaths, and Selected Mediators

Model*	Rate difference (95% CI)†
Mediators: incidence of diabetes and AMI	
Outcome model	
Association between $PM_{2.5}$ and deaths	1.76 (0.88 to 2.64)
Association between incident diabetes and deaths	4.32 (3.39 to 5.25)
Association between incident AMI and deaths	40.04 (35.93 to 44.15)
Association of the interaction between diabetes and AMI with deaths	59.53 (45.67 to 73.40)
Mediator model	
Association between $PM_{2.5}$ and incident diabetes	1.95 (1.06 to 2.84)
Association between $PM_{2.5}$ and incident AMI	0.55 (0.18 to 0.93)
Association between $PM_{2.5}$ and the interaction between diabetes and AMI	0.10 (−0.05 to 0.24)
Mediators: incidence of diabetes and stroke	
Outcome model	
Association between $PM_{2.5}$ and deaths	1.61 (0.75 to 2.47)
Association between incident diabetes and deaths	4.67 (3.76 to 5.58)
Association between incident stroke and deaths	54.02 (51.10 to 57.31)
Association of the interaction between diabetes and stroke with deaths	59.50 (49.28 to 69.35)
Mediator model	
Association between $PM_{2.5}$ and incident diabetes	2.12 (1.23 to 3.02)
Association between $PM_{2.5}$ and incident stroke	0.51 (−0.07 to 1.10)
Association between $PM_{2.5}$ and the interaction between diabetes and stroke	0.16 (−0.03 to 0.34)
Mediators: incidence of diabetes and cardiovascular events‡	
Outcome model	
Association between $PM_{2.5}$ and deaths	1.51 (0.65 to 2.37)
Association between incident diabetes and deaths	4.33 (3.44 to 5.22)
Association between incident cardiovascular events and deaths	47.23 (44.90 to 49.64)
Association of the interaction between diabetes and cardiovascular events with deaths	53.40 (45.26 to 61.69)
Mediator model	
Association between $PM_{2.5}$ and incident diabetes	1.98 (1.09 to 2.87)
Association between $PM_{2.5}$ and incident cardiovascular events	0.97 (0.29 to 1.65)
Association between $PM_{2.5}$ and the interaction between diabetes and cardiovascular events	0.16 (−0.06 to 0.38)

AMI indicates acute myocardial infarction; and $PM_{2.5}$, fine particulate matter.

*Aalen additive hazards models adjusted for the selected individual-level covariates measured at baseline (ie, age, sex, marital status, education, immigration status, household income adequacy, smoking status, smoking pack years, type of drinker, daily consumption of total fruits and vegetables, physical activity, and body mass index) and time-varying area-level variables (ie, education, income, percentage of unemployment, percentage of immigrants, indicators for rural/urban and north/south).

†The effects are presented as number of additional cases per 1000 person-years for each 10 $\mu g/m^3$ increase in $PM_{2.5}$.

‡A composite indicator of incident acute myocardial infarction and stroke (whichever occurred first).

through diabetes alone, 11.5% (95% CI, 4.0%–22.1%) through AMI alone, and 2.9% (95% CI, –1.7% to 9.1%) through the interaction of diabetes and AMI. When considering the incidence of diabetes and stroke as mediators, 24.0% (95% CI, 6.6%–43.7%) of the estimated effect of $PM_{2.5}$ could be attributed to the pathways through these 2 mediators jointly: 5.3% (95% CI, 2.5%–10.4%) through diabetes alone, 13.9% (95% CI, –2.1% to 30.7%) through stroke alone, and 4.8% (95% CI, –1.0% to 11.2%) through the interaction of diabetes and stroke. Lastly, we found that the pathways through the incidence of diabetes and cardiovascular events (AMI and stroke combined) explained 31.7% (95% CI, 17.2%–53.2%) of the total effect of $PM_{2.5}$: 4.5% (95% CI, 1.9%–8.4%) through diabetes alone, 22.8% (95% CI, 9.4%–40.9%) through cardiovascular events alone, and 4.5% (95% CI, –1.5% to 11.7%) through the interaction of diabetes and cardiovascular events.

Sensitivity Analysis

In the sensitivity analyses, the proportions mediated overall remained consistent with those observed in the main analysis (Table S4). For example, the sensitivity analyses with incidence of diabetes and cardiovascular events evaluated as mediators showed that the proportions mediated by these 2 conditions ranged from 30% to 34%. However, the proportion mediated by diabetes was slightly attenuated with the further adjustment for a linear term for time.

Using the approach of dynamic path analysis,[43] the total effect, NDE, and NIEs through mediators (measured by mean cumulative effect estimates over follow-up) were overall consistent with the effect estimates derived from the primary analysis (Table S4). Despite some differences in indirect effects via individual mediators, the proportions of total indirect effects through mediators in all 3 cohorts were similar to and slightly higher than those estimated using the primary approach. For example, the total proportion mediated by diabetes and cardiovascular events was 38% using this approach compared with 32% using the primary approach.

Using the Cox proportional hazards model, we estimated that the hazard ratio for nonaccidental deaths corresponding to each 10 $\mu g/m^3$ increase in $PM_{2.5}$ exposure was 1.13 (95% CI, 1.05–1.22). The association between $PM_{2.5}$ and deaths derived based on the cohort with people with a prior history of diabetes and cardiovascular events added back were slightly strengthened (Table S5).

DISCUSSION

In this population-based cohort study, we observed a substantial increased risk of nonaccidental deaths

Table 3. Estimated Total Effects, Natural Direct Effects, and Natural Indirect Effects and Proportion Mediated Through Selected Mediators for the Association Between Long-Term Exposure to $PM_{2.5}$ and Nonaccidental Deaths

Path-specific effect*	Main analysis	
	Estimate† (95% CI)	Percentage of effect‡ (95% CI)
Mediators: incidence of diabetes and AMI		
Direct effect§	18.6 (9.25 to 27.0)	81.4 (67.9 to 91.2)
Total indirect effect‖	4.0 (2.2 to 6.2)	18.6 (8.8 to 32.1)
Indirect effect via diabetes¶	0.9 (0.5 to 1.4)	4.2 (1.9 to 7.3)
Indirect effect via AMI#	2.5 (0.9 to 4.2)	11.5 (4.0 to 22.1)
Indirect effect via the interaction between diabetes and AMI**	0.6 (−0.4 to 1.7)	2.9 (−1.7 to 9.1)
Total effect	22.6 (13.4 to 31.8)	…
Mediators: incidence of diabetes and stroke		
Direct effect	16.7 (8.3 to 25.7)	76.0 (56.3 to 93.4)
Total indirect effect	5.1 (1.2 to 9.1)	24.0 (6.6 to 43.7)
Indirect effect via diabetes	1.1 (0.6 to 1.6)	5.3 (2.5 to 10.4)
Indirect effect via stroke	3.0 (−0.4 to 6.4)	13.9 (−2.1 to 30.7)
Indirect effect via the interaction between diabetes and stroke	1.0 (−0.2 to 2.1)	4.8 (−1.0 to 11.2)
Total effect	21.8 (12.5 to 31.6)	…
Mediators: incidence of diabetes and cardiovascular events††		
Direct effect	15.7 (6.2 to 24.9)	68.3 (46.8 to 82.8)
Total indirect effect	7.0 (3.3 to 10.8)	31.7 (17.2 to 53.2)
Indirect effect via diabetes	1.0 (0.5 to 1.4)	4.5 (1.9 to 8.4)
Indirect effect via cardiovascular events	5.1 (1.7 to 8.6)	22.8 (9.4 to 40.9)
Indirect effect via the interaction between diabetes and cardiovascular events	1.0 (−0.3 to 2.3)	4.5 (−1.5 to 11.7)
Total effect	22.7 (11.9 to 33.2)	…

AMI indicates acute myocardial infarction; and $PM_{2.5}$, fine particulate matter.

*Adjusted for the selected individual-level covariates measured at baseline (ie, age, sex, marital status, education, immigration status, household income adequacy, smoking status, smoking pack years, type of drinker, daily consumption of total fruits and vegetables, physical activity, and body mass index) and time-varying area-level variables (ie, education, income, percentage of unemployment, percentage of immigrants, indicators for rural/urban and north/south). These effects were estimated using Aalen additive hazards models and scaled by the mean number of follow-up days (3796 days).

†These effects are presented as number of additional cases per 1000 person-years for every 10 $\mu g/m^3$ increase in $PM_{2.5}$ exposure.

‡The proportions of direct effects and indirect effects through selected mediators.

§The effect of $PM_{2.5}$ on nonaccidental deaths that was not mediated by selected mediators.

‖The total effect of $PM_{2.5}$ on nonaccidental deaths mediated by selected mediators.

¶The effect mediated through the first mediator (ie, incidence of diabetes) alone.

#The effect mediated through the second mediator (ie, incidence of AMI, stroke, or combined) alone.

**The effect mediated through the interaction between the 2 mediators.

††A composite indicator of incident AMI and stroke (whichever occurred first).

in association with long-term exposure to $PM_{2.5}$. In addition, $PM_{2.5}$ was positively associated with the incidence of diabetes, AMI, and stroke, 3 major risk factors for premature mortality. Importantly, we found that incidence of diabetes, AMI, and stroke mediated the relationship between $PM_{2.5}$ and nonaccidental deaths, collectively accounting for roughly a third of the total effect of $PM_{2.5}$. Outdoor $PM_{2.5}$ has been recognized as the leading environmental risk factor for premature deaths worldwide.[1,2] Our findings demonstrate that a significant fraction of the estimated adverse effect of $PM_{2.5}$ on mortality operates through its influence on the development of diabetes and cardiovascular events.

There is consistent evidence that people with higher exposure to $PM_{2.5}$ were more likely to die prematurely.[1,2,45,46] We estimated that each 10 $\mu g/m^3$ increase in $PM_{2.5}$ may give rise to an estimated ≈1.5 excess deaths among every 1000 Canadians in each year. This estimate can translate to ≈16 000 excess deaths in Canada each year attributed to $PM_{2.5}$ exposure given that the population-weighted national average concentration of $PM_{2.5}$ is ≈6.1 $\mu g/m^3$, the background concentration of $PM_{2.5}$ is ≈1.8 $\mu g/m^3$,[43] and the total population of Canada is 27.3 million aged >20 years according to the 2016 Census.[47] This is broadly consistent with a recent report of 14 600 excess deaths in Canada attributed to $PM_{2.5}$ exposure.[46] As well, we observed positive associations of $PM_{2.5}$ with the onset of diabetes, AMI, and stroke. Similar findings were also reported in recent population-based studies.[3–10]

The estimated adverse effects of diabetes and cardiovascular diseases on increasing mortality are well known. In this study, we also estimated that for every 1000 person-years, the impacts of these conditions on mortality estimated by the rate difference are ≈5 for diabetes, 40 for AMI, and 54 for stroke (Table 2). These estimates are in line with those reported elsewhere.[48–50] For example, in a recent cohort study comprising ≈1 million adults receiving care in the US Veterans Affairs Healthcare System, it was estimated that for every 1000 individuals with diabetes, there were 7 excess deaths in each year compared with those without diabetes.[48] In addition, a recent Danish national cohort study reported an estimated 53 excess deaths per 1000 people per year in those with nonfatal AMI compared with the general population.[49] Another Danish study estimated that the excess annual mortality attributed to stroke was 57 per 1000 people among participants aged ≥25 years.[50]

A minimal amount of epidemiological evidence exists about how air pollution may affect the trajectory of human health. As previously mentioned, causal mediation analysis (or effect decomposition) is a valuable epidemiological method to assess the role of intermediates that lie along the paths from the exposure to the outcome. To our knowledge, this is the first causal mediation study that empirically assessed pathways from long-term exposure to $PM_{2.5}$ to nonaccidental deaths through the incidence of diabetes and cardiovascular diseases and therefore quantified the relative contributions of these potential mediators. In this study, people with higher exposure to $PM_{2.5}$ were more likely to die in part because they had developed $PM_{2.5}$-induced diabetes and cardiovascular diseases (31.7% in total including 4.5% mediated through diabetes, 22.8% through cardiovascular events, and 4.5% through the interaction of diabetes and cardiovascular events). The higher proportion mediated through cardiovascular events than that through diabetes may indicate that cardiovascular events are the main mediating pathway between $PM_{2.5}$ and mortality. On the other hand, the increased proportion of the indirect effect mediated by AMI and stroke combined (22.8%) than individually (11.5% and 13.9%, respectively) could be explained by a stronger impact of $PM_{2.5}$ on these 2 conditions combined than each of them alone (Table 2). Furthermore, the proportion mediated through the interaction between diabetes and cardiovascular events was not found statistically significant. This was likely attributed to smaller numbers of joint mediators because the follow-up period might not be long enough for us to detect more participants developing both conditions.

Collectively, our findings provide mechanistic insight from the epidemiological perspective and corroborate past experimental evidence that links exposure to $PM_{2.5}$ to changes in cardiometabolic biomarkers and measures.[51,52] Human and animal studies have shown that repetitive exposure to $PM_{2.5}$ results in alterations in endothelial function, which may precede changes in insulin resistance by reducing insulin-dependent glucose uptake and then increase the risk of diabetes.[51] In addition, PM exposure may impair β cell function, resulting in reduced insulin secretion and metabolic dysfunction. As well, several mechanistic pathways between air pollution and cardiovascular disease have been described previously.[52] One pathway involves the initiation of pulmonary and systemic oxidative stress and inflammation, which could contribute to thrombosis, cardiac dysrhythmias, acute vascular dysfunction, plaque instability, the development or progression of atherosclerosis, and cardiovascular events.[52] Another pathway may involve disturbances of the cardiac autonomic nervous system through activation of pulmonary neural reflexes, which might also contribute to instability of vascular plaques and trigger cardiovascular events.[52] According to the reported causes of death, ≈15% of all deaths were attributable to AMI, stroke, and diabetes in Canada each year.[53] By contrast, the present study estimated that ≈32% of $PM_{2.5}$-related mortality was attributed to the 3 conditions.

Interventions on reducing $PM_{2.5}$ exposure and on mitigating intermediate conditions are both critical to reduce the harmful effect of $PM_{2.5}$ on mortality.[54,55] Population-level strategies such as policies to reduce the amount of fuel burned for electricity generation, industrial production, and transportation are essential to reduce the total emissions of air pollutants.[54,55] Individual behavior change (such as staying indoors during periods of poor air quality and using personal filtration systems) can be also helpful in avoiding exposure. Clinicians and health professionals may also play an important role to communicate and advocate health-protective actions and behaviors at the individual level.[54] Beyond interventions to reduce exposure, by enhancing the mechanistic understanding of air pollution effects, our study further suggests that there is an opportunity in clinical practice to help mitigate the adverse effects of $PM_{2.5}$ before they become overwhelming and irreversible. This can be achieved by improving metabolic and cardiovascular health, even in the polluted areas. For example, for individuals at elevated risk of diabetes and cardiovascular disease, clinicians could offer a variety of tailored recommendations and interventions to reduce the overall risk of these conditions (eg, eating healthy foods, exercising more but being mindful of high pollution days or locations, and controlling blood pressure and cholesterol levels).[54,55] In addition, effective community-level interventions may include targeted screening for diabetes and cardiovascular diseases and improved access to health care for the prevention and treatment of these conditions in the most polluted areas.[54]

Strengths of this study include the large cohort sizes and a long follow-up period (up to 17 years), which allowed us to obtain valid mediation effect estimates with good statistical power. Further strengths are our ability to ascertain mediators (incident cases) using province-wide registries and algorithms with high sensitivity and specificity, which greatly reduced the possibility of misclassification; the use of satellite-based PM$_{2.5}$ exposure estimates at finer spatial scales (1 km×1 km); the ability to account for participants' residential mobility by assigning exposures to their residential postal code for each year of follow-up; and the application of analytic methods that allow multiple, correlated, time-varying mediators in the context of censored survival outcomes. Our study also benefits from the adjustment for extensive information on important risk factors of cardiometabolic diseases and deaths at both the individual and neighborhood levels to ensure better control of potential confounders.

Nevertheless, we also acknowledge several limitations. First, information on socioeconomic level and lifestyle was self-reported and was only measured at baseline, thus measurement errors are inevitable. Second, the Ontario Diabetes Database does not distinguish between type 1 and type 2 diabetes. However, given that 90% to 95% of the entire population with diabetes in Ontario has type 2 diabetes and our study participants were aged ≥20 years at cohort entry, the vast majority of incident cases of diabetes in this analysis should be type 2 diabetes. Third, the incidence rate of diabetes in our cohort may be underestimated, although the proportion of diabetes that remained undiagnosed in Ontario has dropped significantly as a result of increasing screening.[56] In addition, our cohorts could be subject to selection effects, biased toward including healthier respondents who had no prior history of diabetes and selected cardiovascular diseases, although the proportion of excluded participants was small (≈9%). As an attempt to investigate the extent of the potential bias, we estimated the association between PM$_{2.5}$ and deaths based on the cohort with people with prior history added back and found similar estimates with those derived from the cohort with these people excluded (Table S5). Furthermore, unmeasured confounding between exposure, outcomes, and mediators may have affected our findings. As an attempt to assess the extent of possible residual confounding, we conducted a series of sensitivity analyses (eg, adjusting for additional risk factors) and found no appreciable change in our observed indirect effect sizes. Our analysis may also be subject to exposure misclassification. The PM$_{2.5}$ estimates were available at the postal-code level, which may not fully reflect each participant's complete personal exposure as it may be influenced by indoor exposures, occupations, and lifestyles.

CONCLUSIONS

In this large, population-based cohort study in Ontario, Canada, we found that the association between chronic exposure to PM$_{2.5}$ and nonaccidental death is considerably (≈one-third) mediated through the incidence of PM$_{2.5}$-associated diabetes and cardiovascular events (AMI and stroke individually or combined). Our findings signify the influence of PM$_{2.5}$ on deteriorating cardiovascular health, highlighting the importance of improving metabolic and cardiovascular health in reducing the enormous burden of PM$_{2.5}$ on mortality, even in countries with low air pollution levels.

ARTICLE INFORMATION

Received July 13, 2022; accepted August 29, 2022.

Affiliations

ICES, Toronto, Ontario, Canada (L.B., J.C.K., H.C.); Scripps Institution of Oceanography (T.B., C.C.); and Department of Family Medicine and Public Health (T.B.), University of California, San Diego, La Jolla, CA; Public Health Ontario, Toronto, Ontario, Canada (J.C.K., J.K., H.C.); Dalla Lana School of Public Health (J.C.K., J.K., H.C.); and Department of Family and Community Medicine (J.C.K.), University of Toronto, Ontario, Canada; Environmental Health Science and Research Bureau, Health Canada, Ottawa, Ontario, Canada (R.T.B., H.C.); Department of Energy, Environment and Chemical Engineering, Washington University, St Louis, MO, USA (A.v.D., R.V.M.); Department of Epidemiology and Biostatistics (J.S.K.) and Institute for Health and Social Policy, McGill University, Montreal, Quebec, Canada (J.S.K.).

Sources of Funding

This study was funded by the Addressing Air Pollution Horizontal Initiative of the Government of Canada. This study was also supported by Public Health Ontario and ICES, which is funded by an annual grant from the Ontario Ministry of Health and the Ministry of Long-Term Care. Parts of this material are based on data and information compiled and provided by Canadian Institute for Health Information (CIHI). Parts of this study are based on Ontario Registrar General information on deaths, the original source of which is Service Ontario. The opinions, results, and conclusions reported in this article do not necessarily represent the views of the ICES, Ontario Registrar General, Ontario Ministry of Health, Ministry of Long-Term Care, or Canadian Institute for Health Information. No endorsement by ICES, Ontario Ministry of Health, or Ministry of Long-Term Care is intended or should be inferred. Martin acknowledges support from NASA Health Air Quality Applied Sciences Team HAQAST (80NSSC21K0508).

Disclosures

None.

Supplemental Material

Data S1–S6
Tables S1–S5
Figure S1

REFERENCES

1. GBD 2017 Risk Factor Collaborators. 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*. 2018;392:1923–1994. doi: 10.1016/S0140-6736(18)32225-6. PMID: 30496105; PMCID: PMC6227755.
2. Burnett R, Chen H, Szyszkowicz M, Fann N, Hubbell B, Pope CA III, Apte JS, Brauer M, Cohen A, Weichenthal S, et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc Natl Acad Sci U S A*. 2018;115:9592–9597. doi: 10.1073/pnas.1803222115. PMID: 30181279; PMCID: PMC6156628.

3. Chen H, Burnett RT, Kwong JC, Villeneuve PJ, Goldberg MS, Brook RD, van Donkelaar A, Jerrett M, Martin RV, Brook JR, Copes R Risk of incident diabetes in relation to long-term exposure to fine particulate matter in Ontario, Canada. *Environ Health Perspect.* 2013;121:804–810. doi: 10.1289/ehp.1205958. PMID: 23632126; PMCID: PMC3701997.
4. Paul LA, Burnett RT, Kwong JC, Hystad P, van Donkelaar A, Bai L, Goldberg MS, Lavigne E, Copes R, Martin RV, et al. The impact of air pollution on the incidence of diabetes and survival among prevalent diabetes cases. *Environ Int.* 2020;134:105333. doi: 10.1016/j.envint.2019.105333 PMID: 31775094
5. Hansen AB, Ravnskjær L, Loft S, Andersen KK, Bräuner EV, Baastrup R, Yao C, Ketzel M, Becker T, Brandt J, et al. Long-term exposure to fine particulate matter and incidence of diabetes in the Danish Nurse Cohort. *Environ Int.* 2016;91:243–250. doi: 10.1016/j.envint.2016.02.036 PMID: 26989812
6. Bai L, Shin S, Burnett RT, Kwong JC, Hystad P, van Donkelaar A, Goldberg MS, Lavigne E, Copes R, Martin RV, et al. Exposure to ambient air pollution and the incidence of congestive heart failure and acute myocardial infarction: a population-based study of 5.1 million Canadian adults living in Ontario. *Environ Int.* 2019;132:105004. doi: 10.1016/j.envint.2019.105004 PMID: 31387019
7. Madrigano J, Kloog I, Goldberg R, Coull BA, Mittleman MA, Schwartz J. Long-term exposure to PM2.5 and incidence of acute myocardial infarction. *Environ Health Perspect*. 2013;121:192–196. doi: 10.1289/ehp.1205284. PMID: 23204289; PMCID: PMC3569684.
8. Shin S, Burnett RT, Kwong JC, Hystad P, van Donkelaar A, Brook JR, Goldberg MS, Tu K, Copes R, Martin RV, Liu Y, Kopp A, Chen H Ambient air pollution and the risk of atrial fibrillation and stroke: a population-based cohort study. *Environ Health Perspect.* 2019;127:87009. doi: 10.1289/EHP4883. PMID: 31449466; PMCID: PMC6792368.
9. Scheers H, Jacobs L, Casas L, Nemery B, Nawrot TS. Long-term exposure to particulate matter air pollution is a risk factor for stroke: meta-analytical evidence. *Stroke.* 2015;46:3058–3066, PMID: 26463695, 10.1161/STROKEAHA.115.009913.
10. Wolf K, Hoffmann B, Andersen ZJ, Atkinson RW, Bauwelinck M, Bellander T, Brandt J, Brunekreef B, Cesaroni G, Chen J, et al. Long-term exposure to low-level ambient air pollution and incidence of stroke and coronary heart disease: a pooled analysis of six European cohorts within the ELAPSE project. *Lancet Planet Health.* 2021;5:e620–e632. doi: 10.1016/S2542-5196(21)00195-9 PMID: 34508683
11. Hamanaka RB, Mutlu GM. Particulate matter air pollution: effects on the cardiovascular system. *Front Endocrinol (Lausanne).* 2018;9:680. doi: 10.3389/fendo.2018.00680. PMID: 30505291; PMCID: PMC6250783.
12. Brook RD, Rajagopalan S, Pope CA III, Brook JR, Bhatnagar A, Diez-Roux AV, Holguin F, Hong Y, Luepker RV, Mittleman MA, et al. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation.* 2010;121:2331–2378. doi: 10.1161/CIR.0b013e3181dbece1 PMID: 20458016
13. VanderWeele TJ. Mediation analysis: a practitioner's guide. *Annu Rev Public Health.* 2016;37:17–32. doi: 10.1146/annurev-publhealth-032315-021402
14. VanderWeele TJ. *Explanation in causal inference: methods for mediation and interaction*. New York: Oxford University Press; 2015.
15. VanderWeele TJ. Causal mediation analysis with survival data. *Epidemiology.* 2011;22:582–585. doi: 10.1097/EDE.0b013e31821db37e
16. Lange T, Hansen JV. Direct and indirect effects in a survival context. *Epidemiology.* 2011;22:575–581. doi: 10.1097/EDE.0b013e31821c680c
17. Tanz LJ, Stuart JJ, Williams PL, Rimm EB, Missmer SA, Rexrode KM, Mukamal KJ, Rich-Edwards JW. Preterm delivery and maternal cardiovascular disease in young and middle-aged adult women. *Circulation.* 2017;135:578–589. doi: 10.1161/CIRCULATIONAHA.116.025954. PMID: 28153993; PMCID: PMC5308562.
18. VanderWeele TJ, Vansteelandt S. Mediation analysis with multiple mediators. *Epidemiol Methods.* 2014;2:95–115. doi: 10.1515/em-2012-0010
19. Pan WC, Wu CD, Chen MJ, Huang YT, Chen CJ, Su HJ, Yang HI. Fine particle pollution, alanine transaminase, and liver cancer: a Taiwanese Prospective Cohort Study (REVEAL-HBV). *J Natl Cancer Inst.* 2016;108:djv341.
20. Mazidi M, Speakman JR. Impact of obesity and ozone on the association between particulate air pollution and cardiovascular disease and stroke mortality among US adults. *J Am Heart Assoc.* 2018;7:e008006. doi: 10.1161/JAHA.117.008006. PMID: 29848499; PMCID: PMC6015356.
21. Inoue K, Yan Q, Arah OA, Paul K, Walker DI, Jones DP, Ritz B. Air pollution and adverse pregnancy and birth outcomes: mediation analysis using metabolomic profiles. *Curr Environ Health Rep.* 2020;7:231–242. doi: 10.1007/s40572-020-00284-3
22. Havet A, Li Z, Zerimech F, Sanchez M, Siroux V, Le Moual N, Brunekreef B, Künzli N, Jacquemin B, Varraso R, et al. Does the oxidative stress play a role in the associations between outdoor air pollution and persistent asthma in adults? Findings from the EGEA study. *Environ Health.* 2019;18:90. doi: 10.1186/s12940-019-0532-0. PMID: 31665023; PMCID: PMC6819357.
23. Ilango SD, Chen H, Hystad P, van Donkelaar A, Kwong JC, Tu K, Martin RV, Benmarhnia T. The role of cardiovascular disease in the relationship between air pollution and incident dementia: a population-based cohort study. *Int J Epidemiol.* 2020;49:36–44. doi: 10.1093/ije/dyz154. PMID: 31347651; PMCID: PMC7124495.
24. Paul KC, Haan M, Yu Y, Inoue K, Mayeda ER, Dang K, Wu J, Jerrett M, Ritz B. Traffic-related air pollution and incident dementia: direct and indirect pathways through metabolic dysfunction. *J Alzheimers Dis.* 2020;76:1477–1491. doi: 10.3233/JAD-200320. PMID: 32651321; PMCID: PMC7591265.
25. Lucht S, Hennig F, Moebus S, Ohlwein S, Herder C, Kowall B, Jöckel KH, Hoffmann B. All-source and source-specific air pollution and 10-year diabetes Incidence: total effect and mediation analyses in the Heinz Nixdorf recall study. *Environ Int.* 2020;136:105493. doi: 10.1016/j.envint.2020.105493 PMID: 31991234
26. Statistics Canada. 2020. National Population Health Survey – Household Component – Longitudinal (NPHS). Available: https://www.statcan.gc.ca/eng/survey/household/3225. Accessed 31 August 2022.
27. Statistics Canada. 2021. Canadian Community Health Survey – Annual Component (CCHS). Available at: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3226. Accessed 31 August 2022.
28. van Donkelaar A, Martin RV, Li C, Burnett RT. 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.* 2019;53:2595–2611. doi: 10.1021/acs.est.8b06392 PMID: 30698001
29. Bai L, Shin S, Burnett RT, Kwong JC, Hystad P, van Donkelaar A, Goldberg MS, Lavigne E, Weichenthal S, Martin RV, et al. Exposure to ambient air pollution and the incidence of lung cancer and breast cancer in the Ontario population health and environment cohort. *Int J Cancer.* 2020;146:2450–2459. doi: 10.1002/ijc.32575 PMID: 31304979
30. Chen H, Burnett RT, Bai L, Kwong JC, Crouse DL, Lavigne E, Goldberg MS, Copes R, Benmarhnia T, Ilango SD, van Donkelaar A, Martin RV, Hystad P Residential greenness and cardiovascular disease incidence, readmission, and mortality. *Environ Health Perspect.* 2020;128:87005. doi: 10.1289/EHP6161. PMID: 32840393; PMCID: PMC7446772.
31. Crouse DL, Peters PA, van Donkelaar A, Goldberg MS, Villeneuve PJ, Brion O, Khan S, Atari DO, Jerrett M, Pope CA, Brauer M, Brook JR, Martin RV, Stieb D, Burnett RT 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.* 2012;120:708–714. doi: 10.1289/ehp.1104049. PMID: 22313724; PMCID: PMC3346774.
32. Lipscombe LL, Hwee J, Webster L, Shah BR, Booth GL, Tu K. Identifying diabetes cases from administrative data: a population-based validation study. *BMC Health Serv Res.* 2018;18:316. doi: 10.1186/s12913-018-3148-0. PMID: 29720153; PMCID: PMC5932874.
33. Austin PC, Daly PA, Tu JV. A multicenter study of the coding accuracy of hospital discharge administrative data for patients admitted to cardiac care units in Ontario. *Am Heart J.* 2002;144:290–296. doi: 10.1067/mhj.2002.123839 PMID: 12177647
34. Tu K, Wang M, Young J, Green D, Ivers NM, Butt D, Jaakkimainen L, Kapral MK. Validity of administrative data for identifying patients who have had a stroke or transient ischemic attack using EMRALD as a reference standard. *Can J Cardiol.* 2013;29:1388–1394. doi: 10.1016/j.cjca.2013.07.676 PMID: 24075778
35. Havranek EP, Mujahid MS, Barr DA, Blair IV, Cohen MS, Cruz-Flores S, Davey-Smith G, Dennison-Himmelfarb CR, Lauer MS, Lockwood DW, et al. Social determinants of risk and outcomes for cardiovascular disease: a scientific statement from the American Heart Association. *Circulation.* 2015;132:873–898. doi: 10.1161/CIR.0000000000000228 PMID: 26240271
36. Mozaffarian D, Wilson PW, Kannel WB. Beyond established and novel risk factors: lifestyle risk factors for cardiovascular disease. *Circulation.*

2008;117:3031–3038. doi: 10.1161/CIRCULATIONAHA.107.738732 PMID: 18541753
37. Strak M, Janssen N, Beelen R, Schmitz O, Karssenberg D, Houthuijs D, van den Brink C, Dijst M, Brunekreef B, Hoek G. Associations between lifestyle and air pollution exposure: potential for confounding in large administrative data cohorts. *Environ Res.* 2017;156:364–373. doi: 10.1016/j.envres.2017.03.050 PMID: 28395240
38. Statistics Canada. 2021. Census Program. Available: https://www12.statcan.gc.ca/census-recensement/index-eng.cfm. Accessed 31 August 2022.
39. Low Wang CC, Hess CN, Hiatt WR, Goldfine AB. Clinical update: cardiovascular disease in diabetes mellitus: atherosclerotic cardiovascular disease and heart failure in type 2 diabetes mellitus - mechanisms, management, and clinical considerations. *Circulation.* 2016;133:2459–2502. doi: 10.1161/CIRCULATIONAHA.116.022194
40. Huang YT, Yang HI. Causal mediation analysis of survival outcome with multiple mediators. *Epidemiology.* 2017;28:370–378. doi: 10.1097/EDE.0000000000000651. PMID: 28296661; PMCID: PMC5408128.
41. Wang W, Nelson S, Albert JM. Estimation of causal mediation effects for a dichotomous outcome in multiple-mediator models using the mediation formula. *Stat Med.* 2013;32:4211–4228. doi: 10.1002/sim.5830. PMID: 23650048; PMCID: PMC3789850.
42. Chen H, Burnett RT, Copes R, Kwong JC, Villeneuve PJ, Goldberg MS, Brook RD, van Donkelaar A, Jerrett M, Martin RV, Brook JR, Kopp A, Tu JV Ambient fine particulate matter and mortality among survivors of myocardial infarction: population-based cohort study. *Environ Health Perspect.* 2016;124:1421–1428. doi: 10.1289/EHP185. PMID: 27152932; PMCID: PMC5010396.
43. Aalen OO, Stensrud MJ, Didelez V, Daniel R, Røysland K, Strohmaier S. Time-dependent mediators in survival analysis: modeling direct and indirect effects with the additive hazards model. *Biom J.* 2020;62:532–549. doi: 10.1002/bimj.201800263 PMID: 30779372
44. Naimi AI, Cole SR, Kennedy EH. An introduction to g methods. *Int J Epidemiol.* 2017;46:756–762. doi: 10.1093/ije/dyw323. PMID: 28039382; PMCID: PMC6074945.
45. Government of Canada. 2021. Health Impacts of Air Pollution in Canada: Estimates of morbidity and premature mortality outcomes – 2021 Report. Available at: https://www.canada.ca/en/health-canada/services/publications/healthy-living/2021-health-effects-indoor-air-pollution.html. Accessed 15 March 2022.
46. Environment and Climate Change Canada. 2020. Health impacts from air pollution. Available at: https://www.canada.ca/en/environment-climate-change/campaigns/canadian-environment-week/clean-air-day/health-impacts-air-pollution.html. Accessed 15 March 2022.
47. Statistics Canada. 2021. Data products, 2016 Census. Available at: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/index-eng.cfm. Accessed 15 March 2022.
48. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, Lin JK, Farzadfar F, Khang YH, Stevens GA, et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *Lancet.* 2011;378:31–40. doi: 10.1016/S0140-6736(11)60679-X
49. Schmidt M, Szépligeti S, Horváth-Puhó E, Pedersen L, Bøtker HE, Sørensen HT. Long-term survival among patients with myocardial infarction before age 50 compared with the general population: a Danish Nationwide Cohort Study. *Circ Cardiovasc Qual Outcomes.* 2016;9:523–531. doi: 10.1161/CIRCOUTCOMES.115.002661 PMID: 27576336
50. Brønnum-Hansen H, Davidsen M, Thorvaldsen P, Danish MONICA Study Group. Long-term survival and causes of death after stroke. *Stroke.* 2001;32:2131–2136. doi: 10.1161/hs0901.094253 PMID: 11546907
51. Rajagopalan S, Brook RD. Air pollution and type 2 diabetes: mechanistic insights. *Diabetes.* 2012;61:3037–3045. doi: 10.2337/db12-0190. PMID: 23172950; PMCID: PMC3501850.
52. Brook RD, Franklin B, Cascio W, Hong Y, Howard G, Lipsett M, Luepker R, Mittleman M, Samet J, Smith SC Jr, et al; Expert Panel on Population and Prevention Science of the American Heart Association. Air pollution and cardiovascular disease: a statement for healthcare professionals from the expert panel on population and prevention science of the American Heart Association. *Circulation.* 2004;109:2655–2671. doi: 10.1161/01.CIR.0000128587.30041.C8. PMID: 15173049.
53. Statistics Canada. 2022. Mortality, Summary List of Causes, 2009 – ARCHIVED. Available at: https://www150.statcan.gc.ca/n1/pub/84f0209x/84f0209x2009000-eng.pdf. Accessed 15 March 2022.
54. Keswani A, Akselrod H, Anenberg S. Health and clinical impacts of air pollution and linkages with climate change. *NEJM Evidence.* 2022;1:EVIDra2200068. doi: 10.1056/EVIDra2200068
55. Perera F, Nadeau K. Climate change, fossil-fuel pollution, and children's health. *N Engl J Med.* 2022;386:2303–2314. doi: 10.1056/NEJMra2117706
56. Creatore MI, Booth GL, Manuel DG, Moineddin R, Glazier RH. Diabetes screening among immigrants: a population-based urban cohort study. *Diabetes Care.* 2012;35:754–761. doi: 10.2337/dc11-1393

SUPPLEMENTAL MATERIAL

Data S1. ICD codes to ascertain non-accidental deaths

Ontario's Registrar General Death File is an annual dataset containing information on all deaths registered in Ontario starting on January 1, 1990. Ontario switched from International Classification of Diseases (ICD)-9 classifications to ICD-I0 starting in 2000 to code deaths since January 1, 2000. Therefore, we used ICD-9 codes to define non-accidental deaths (001-799) from 1996 to 1999; from 2000 onward, ICD-10 (A00-R99) codes were used to ascertain non-accidental deaths.

Data S2. Hospital discharge abstracts from the Canadian Institute for Health Information

The Canadian Institute for Health Information Discharge Abstract Database [CIHI-DAD] collects data abstracted from patients' medical charts at discharge by professional, certified medical coders. This database includes up to 25 diagnostic codes related to the hospitalization. For hospitalizations due to diabetes, we considered any diagnosis code. To ascertain incident acute myocardial infarction and stroke, we considered the most responsible diagnosis code

Data S3. Missingness of exposure and covariates

We excluded participants with missing exposure data at baseline (n=3,643). To avoid losing substantial statistical power, we included participants with missing covariates. We used different approaches to handle the missingness depending on the amount of missing data. For example, we replaced missing values with the most frequent value (the category with the most participants) for marital status, income quintile, physical activity, working status, alcohol consumption. For family income levels, BMI, and smoking status, we created a separate category of missing values.

Data S4. Ascertainment of prevalent cases of diabetes, acute myocardial infarction (AMI), and stroke

To capture that first-ever incidences of mediators during follow-up, we excluded respondents who had previous physician diagnoses of diabetes, AMI, and/or stroke at the time of entry. We determined prior history of these conditions using the same ICD codes, databases, and algorithms that were used to ascertain incident cases (more details can be found in the main text under the "Mediators" section). The presence of the diagnosis of diabetes, AMI, or stroke between 1991 and the survey date was defined as the prevalent cases.

Data S5. Fitting Aalen additive hazards models

We evaluated the associations between the exposure, the outcome, and the mediators using Aalen additive hazards models. Specifically, for the analysis with incidence of diabetes and AMI considered as mediators, we constructed four Aalen additive hazards models: an outcome model associating $PM_{2.5}$ with deaths by adjusting the joint mediator and three mediator models associating $PM_{2.5}$ with incidence of diabetes, incidence of AMI, and the interaction between diabetes and AMI, respectively. The same approach was applied to the analyses with the other two pairs of sequential mediators considered.

Data S6. R code examples of fitting Aalen additive hazards models and calculating direct, indirect effects, and proportion mediated

```
## ------------------------------------------------------------------------------------------------------------------------
##© 2022 Institute for Clinical Evaluative Sciences. All rights reserved.
##TERMS OF USE:
##Not for distribution: This code and data is provided to the user solely for its own non-commercial
##use by individuals and/or not-for-profit corporations. User shall not distribute without express written
##permission from the Institute for Clinical Evaluative Sciences.
##Not-for-profit: This code and data may not be used in connection with profit generating activities.
##No liability: The Institute for Clinical Evaluative Sciences makes no warranty or representation
##regarding the fitness, quality or reliability of this code and data.
##No Support: The Institute for Clinical Evaluative Sciences will not provide any technological,
##educational or informational support in connection with the use of this code and data.
##Warning: By receiving this code and data, user accepts these terms, and uses the code and data,
##solely at its own risk.
## ------------------------------------------------------------------------------------------------------------------------
## Paper Title: Chronic PM2.5 Exposure Increases Mortality through Pathways of Metabolic and
## Cardiovascular Disease: Insights from a Large Population-based Mediation Analysis
## Authors:  Li Bai, Tarik Benmarhnia, Chen Chen, Jeffrey C. Kwong, Richard T. Burnett, Aaron van
## Donkelaar, Randall V. Martin, JinHee Kim, Jay S. Kaufman, Hong Chen
## Contact (code): Li Bai (li.bai@ices.on.ca) and Hong Chen (hong.chen@hc-sc.gc.ca)
## ------------------------------------------------------------------------------------------------------------------------
## Description: R code example of fitting Aalen additive hazards models and calculating direct, indirect
## effects, and proportion mediated. This example is for investigating the extent to which the effect of
## PM2.5 on deaths is ## mediated by its effect on the development of diabetes and acute myocardial ##
infarction (AMI).
## ------------------------------------------------------------------------------------------------------------------------
## Required Packages: timereg
## ------------------------------------------------------------------------------------------------------------------------
## Requires data in counting process format.
##
## Variable description:
##       event: non-accidental deaths (binary)
##       PM25: long-term exposure to PM2.5 (continuous)
##       m_td:  joint mediator indicator (categorical: 1: free of both diabetes and AMI diagnoses during
##       follow-up, 2: developed diabetes only, 3: developed AMI only, and 4: developed both diabetes
##       and AMI)
##       M_td1_2: incident diabetes (binary)
##       M_td1_3: incident AMI (binary)
##       M_td1_4: incident diabetes and AMI (whichever came first) (binary)
##       start: start time of follow up in days
##       stop: end time of follow up in days (ended when reaching Dec 31, 2017, becoming ineligible for
##       provincial health insurance, or death)
##       stop_m1: end time of follow up for diabetes in days (ended when reaching Dec 31, 2017,
##       becoming ineligible for provincial health insurance, death, or receiving the diagnosis of diabetes)
##       stop_m2: end time of follow up for AMI in days (ended when reaching Dec 31, 2017,
##       becoming ineligible for provincial health insurance, death, or receiving the diagnosis of AMI)
##       cov1:  selected categorial covariates such as BMI and smoking
##       cov2:  selected continuous covariates such as area-level risk factors (e.g., % of low education)
## ------------------------------------------------------------------------------------------------------------------------
```

```
#### load library
library(timereg)

#### load data
load('... /data/dataset')

#### fit the outcome model
m.outcome<-aalen(Surv(start, stop, event)~ const(PM25) + const(factor(m_td)) + const(factor(cov1)) +
const(cov2), data=dataset, robust=0, n.sim = 500)

# extract the effect of PM2.5 on non-accidental deaths measured as
theta1<-m.outcome$gamma[1]
theta1.se<-sqrt(m.outcome$var.gamma[1,1])

# extract the effect of incident diabetes on non-accidental deaths
theta2_1<-m.outcome$gamma[2]
theta2_1.se=sqrt(m.outcome$var.gamma[2,2])

# extract the effect of incident AMI on non-accidental deaths
theta2_2<-m.outcome$gamma[3]
theta2_2.se<-sqrt(m.outcome$var.gamma[3,3])

# extract the effect of incident diabetes and AMI on non-accidental deaths
theta2_3<-m.outcome$gamma[4]
theta2_3.se<-sqrt(m.outcome$var.gamma[4,4])

### fit the mediator model for incidence of diabetes
m.mediator1<-aalen(Surv(start, stop_m1, M_td1_2)~const(PM25) + const(factor(cov1)) + const(cov2),
data=data, robust=0, n.sim = 500)

# extract the effect of PM2.5 on incidence of diabetes
beta1_1 <- m.mediator1$gamma[1]
beta1_1.se <- sqrt(m.mediator1$var.gamma[1,1])

### fit the mediator model for incidence of AMI
m.mediator2<-aalen(Surv(start, stop_m2, M_td1_3)~const(PM25) + const(factor(cov1)) + const(cov2),
data=dat_boot, robust=0, n.sim = 500)

# extract the effect of PM2.5 on incidence of AMI
beta1_2<-m.mediator2$gamma[1]
beta1_2.se <-sqrt(m.mediator2$var.gamma[1,1])

###fit the mediator model for incidence of diabetes and AMI
m.mediator3<-aalen(Surv(start, stop_m2, M_td1_4)~const(PM25) + const(factor(cov1)) + const(cov2),
data=dat_boot, robust=0, n.sim = 500)

# extract the effect of PM2.5 on incident diabetes and AMI
beta1_3<-m.mediator3$gamma[1]
beta1_3.se<-sqrt(m.mediator3$var.gamma[1,1])
```

```
### rescale the effects

days <- 3796   # mean follow up days of the cohort
pm_unit <- 10   # every 10 unit increase in PM2.5

theta1   <- theta1 * days * pm_unit
theta2_1 <- theta2_1 * days
theta2_2 <- theta2_2 * days
theta2_3 <- theta2_3 * days

beta1_1 <- beta1_1 * days * pm_unit
beta1_2 <- beta1_2 * days * pm_unit
beta1_3 <- beta1_3 * days * pm_unit

### estimate direct, indirect, total effects, and proportion mediated

# natural direct effect
NDE <- theta1
# indirect effect via incident diabetes
NIE1 <- beta1_1 * theta2_1
# indirect effect via incident AMI
NIE2 <- beta1_2 * theta2_2
# indirect effect via the interaction between diabetes and AMI
NIE3 <- beta1_3 * theta2_3
# total indirect effect
NIE.T <- NIE1 + NIE2 +NIE3
# total effect
TE <- NDE + NIE1 + NIE2 +NIE3

# proportion of direct effect
prop_NDE <- NDE / TE
# proportion of indirect effect via incident diabetes
prop_NIE1 <- NIE1 / TE
# proportion of indirect effect via incident AMI
prop_NIE2 <- NIE2 / TE
# proportion of indirect effect via the interaction between diabetes and AMI
prop_NIE3 <- NIE3 / TE
# pproportion of total indirect effect
prop_NIE.T <- NIE.T / TE

###end
```

Table S1. The time period of data collection for each cycle

	The time period of data collection
1996/1997 cycle of National Population Health Survey	June 1996 - August 1997
Canadian Community Health Surveys	
2000/2001 cycle	September 2000 - November 2001
2003 cycle	January 2003 - November 2003
2005 cycle	January 2005 - December 2005
2007/2008 cycle	January 2007 - December 2008
2009/2010 cycle	January 2009 - December 2010
2011/2012 cycle	January 2011 - December 2012
2013/2014 cycle	January 2013 - December 2014

Table S2. Summary of variables

Variable	Timing	Level	Data source*
Exposure			
PM2.5	1995 to the end of follow-up	Postal code level†	Annual estimates of ground-level PM2.5 concentrations developed by van Donkelaar et al. (2019)
Outcome and mediators			
Non-accidental deaths	from the baseline to the end of follow-up‡	Individual level	Office of the Registrar General Vital Statistics Death Database (ORGD)
Incident diabetes	from the baseline to the end of follow-up	Individual level	Hospital discharge abstracts from the Canadian Institute for Health Information, physician service claims from the Ontario Health Insurance Plan database and claims for prescription drugs from the Ontario Drug Benefit database
Incident AMI	from the baseline to the end of follow-up	Individual level	Hospital discharge abstracts from the Canadian Institute for Health Information
Incident stroke	from the baseline to the end of follow-up	Individual level	Hospital discharge abstracts from the Canadian Institute for Health Information, physician service claims from the Ontario Health Insurance Plan database
Incident cardiovascular events	from the baseline to the end of follow-up	Individual level	Hospital discharge abstracts from the Canadian Institute for Health Information, physician service claims from the Ontario Health Insurance Plan database
Risk factors			
Age	at baseline	Individual level	The 1996/1997 cycle of National Population Health Survey (NPHS); the 2000/2001, 2003, 2005, 2007/2008, 2009/2010, 2011/2012, and 2013/2014 cycles of the Canadian Community Health Surveys (CCHS)
Sex	at baseline	Individual level	
Marital status	at baseline	Individual level (self-reported)	
White or non-white ethnicity	at baseline	Individual level (self-reported)	
Immigrants or non-immigrants	at baseline	Individual level (self-reported)	
Education	at baseline	Individual level (self-reported)	
Household income adequacy	at baseline	Individual level (self-reported)	
Smoking status	at baseline	Individual level (self-reported)	
Smoking pack years (only available in CCHS)	at baseline	Individual level (self-reported)	

Type of drinker	at baseline	Individual level (self-reported)	
Body mass index (BMI)	at baseline	Individual level (self-reported)	
Physical activity (based on energy expenditure)	at baseline	Individual level (self-reported)	
Nearest Census based neighbourhood % of recent immigrants	in 1996, 2001, 2006, and 2016	Census Division level[§]	Canadian Census data
Nearest Census based neighbourhood % of population aged ≥15 years who had not completed high school	in 1996, 2001, 2006, and 2016	Census Division level	
Nearest Census based neighbourhood unemployment rate	in 1996, 2001, 2006, and 2016	Census Division level	
Nearest Census based neighbourhood income quintile	in 1996, 2001, 2006, and 2016	Census Division level	
Rural residence (the community size is <= 10,000)	from the baseline to the end of follow-up	Postal code level	Postal Code Conversion File (PCCF)
Northern or Southern Ontario	from the baseline to the end of follow-up	Postal code level	Postal Code Conversion File (PCCF)

[*]These datasets were linked using unique encoded identifiers and analyzed at ICES.
[†]A single six-digit residential postal code can correspond to one side of a city street between consecutive intersections or a community mailbox or an apartment/business building. Using the Postal Code Conversion File (PCCF), six-character postal codes can be lined to standard Census geographic areas (e.g., dissemination areas, census tracts).
[‡]Eligible respondents were followed up from the time of survey (baseline) and censored when reaching the end of follow-up (December 31, 2017), becoming ineligible for provincial health insurance, or death.
[§]Census divisions are provincially legislated areas (equivalent to counties).

Table S3. Demographic characteristics of the study cohorts

Characteristics	Diabetes-AMI cohort				Diabetes-stroke cohort			
	All participants	Participants who were diagnosed with diabetes during follow-up	Participants who were diagnosed with AMI during follow-up	Participants who died due to non-accidental causes during follow-up	All participants	Participants who were diagnosed with diabetes during follow-up	Participants who were diagnosed with stroke during follow-up	Participants who died due to non-accidental causes during follow-up
N	192,362	18,739	4,014	21,318	190,052	18,444	8,608	20,258
PM2.5 exposure at entry [Mean ± SD]	8.46 ± 2.47	8.70 ± 2.68	8.59 ± 2.75	8.71 ± 2.78	8.45 ± 2.47	8.69 ± 2.68	8.69 ± 2.76	8.70 ± 2.78
Individual-level risk factors [Mean ± SD or percent]								
Age at entry (years)	49.7 ± 17.5	55.5 ± 14.6	63.7 ± 14.5	70.0 ± 12.8	49.4 ± 17.4	55.2 ± 14.4	66.1 ± 13.9	69.7 ± 12.9
Males	44.7	48.9	57.7	45.7	45.0	49.4	44.4	46.6
Immigrants	20.3	25.0	22.1	22.6	20.2	24.9	23.4	22.6
Nonwhite	11.2	12.7	5.7	4.2	11.3	12.8	5.7	4.2
Living in an urban area	74.0	74.0	70.7	72.4	74.0	73.8	73.7	72.4
Living in southern region	86.3	85.0	81.8	84.7	86.3	84.7	85.0	84.5
Marital status								
Married or common law	59.4	61.4	56.8	48.4	59.6	61.8	54.8	48.6
Single (never married)	19.1	12.9	9.3	8.4	19.2	12.8	8.1	8.5
Separate, widowed, or divorced	21.5	25.7	33.9	43.2	21.2	25.4	37.1	42.9
Education								
Less than high school	16.6	25.4	33.6	38.0	16.4	25.3	31.9	38.0
High school graduation	19.5	19.1	18.6	18.3	19.5	19.1	17.9	18.2
Some form of postsecondary education	8.0	8.3	7.6	7.3	8.0	8.4	7.6	7.4
Postsecondary graduation	56.0	47.2	40.2	36.4	56.1	47.3	42.6	36.5
Family income								
≤$29999	20.4	27.3	33.3	41.4	20.1	26.7	33.6	40.9
$30,000-$79,999	43.2	44.3	40.2	37.1	43.2	44.5	41.1	37.4
≥$80,000	29.3	19.2	15.8	8.8	29.6	19.6	13.9	9.1
Missing	7.1	9.2	10.7	12.7	7.1	9.2	11.4	12.5
Smoking (in NPHS)								
Never smoker	40.6	37.6	33.6	35.2	40.5	37.0	41.5	24.9

Daily smoker	24.9	24.9	30.1	24.5	24.9	24.8	19.7	31.8
Occasional smoker	2.0	1.8	1.1	1.7	2.0	2.0	1.7	1.7
Always occasional smoker	1.5	1.5	0.9	1.0	1.5	1.6	0.8	1.0
Former daily smoker	24.3	28.7	28.9	31.9	24.2	29.0	30.7	5.5
Former occasional smoker	6.8	5.6	5.4	5.6	6.8	5.6	5.5	35.1
Smoking (in CCHS)								
Never smoker	33.9	30.6	24.7	27.0	33.8	30.2	30.6	26.6
Current smoker of <10 pack-years	8.3	4.7	2.9	2.1	8.4	4.8	2.6	2.1
Current smoker of 10-20 pack-years	5.1	4.5	5.6	3.4	5.1	4.6	3.9	3.5
Current smoker of ≥20 pack years	9.2	12.9	19.1	17.2	9.2	13.1	12.6	17.3
Current smoker with missing pack-years	0.0	1.0	1.0	1.3	0.9	1.0	1.0	1.2
Former Smoker who quitted within 5 years	6.8	7.4	6.1	6.5	6.9	7.7	4.8	6.6
Former Smoker who quitted > 5 years	22.5	27.9	31.3	33.5	22.4	27.7	33.3	33.6
Missing	13.3	11.0	9.4	9.1	13.3	11.0	11.1	9.0
BMI (kg/m2)								
<18.5	2.1	0.6	1.3	3.3	2.0	0.6	2.0	3.2
18.5-25.0	42.9	19.0	33.7	37.3	42.9	18.7	36.7	37.2
25.0-30.0	33.9	36.7	36.3	29.0	33.9	36.7	33.5	29.2
30.0-35.0	11.9	23.6	13.5	10.1	11.9	23.9	11.4	10.2
≥35.0	4.3	12.6	4.0	3.7	4.3	12.7	3.6	3.8
Missing	4.9	7.5	11.3	16.7	4.8	7.4	12.8	16.5
Physical activity								
Active	24.3	18.0	18.7	15.5	24.5	18.2	19.5	16.0
Moderate active	25.5	22.9	24.0	20.0	25.6	23.0	23.5	20.3
Inactive	50.1	59.1	57.3	64.6	50.0	58.8	57.0	63.7
Area-level risk factor* [Mean ± SD or percent]								
Percentage of recent immigrants	2.6 ± 2.9	2.6 ± 3.0	2.0 ± 2.6	2.2 ± 2.7	2.6 ± 2.9	2.5 ± 3.0	2.2 ± 2.7	2.2 ± 2.7
Percentage of the population ≥15 years of age without employment	7.0 ± 1.4	7.0 ± 1.4	7.1 ± 1.5	7.0 ± 1.4	7.0 ± 1.4	7.0 ± 1.4	7.1 ± 1.4	7.0 ± 1.4
Percentage of the population ≥15 years of age with less than a high school education	26.8 ± 4.4	26.8 ± 4.4	27.5 ± 4.1	27.2 ± 4.1	26.8 ± 4.4	26.8 ± 4.3	27.2 ± 4.1	27.3 ± 4.1
Income quintiles								
Lowest	18.4	22.0	22.9	24.7	18.4	21.9	22.5	24.7

Lower	19.7	21.1	21.6	21.3	19.6	21.0	20.7	21.3
Middle	19.8	20.5	19.3	19.3	19.8	20.6	19.7	19.2
Upper	22.1	19.7	19.6	18.4	22.2	19.8	19.2	18.5
Uppermost	20.0	16.6	16.7	16.3	20.0	16.7	17.9	16.3

AMI: acute myocardial infarction; BMI: body mass index; CCHS: Canadian Community Health Surveys; NPHS: National Population Health Survey; SD: standard deviation.

* From Canadian Census 2001, at the dissemination area level.

Table S4. Proportions (%) of the direct and indirect effects derived from the sensitivity analysis

Path-specific effect	Main analysis*	Sensitivity analysis					
		Using the restricted cubic splines of age†	Further adjusting for a linear term for time†	Further adjusting for alcohol consumption†	Further working status†	Further adjusting for sense of belong to local community†	Using a method of dynamic path analysis‡
Mediators: Incidence of diabetes and AMI							
Direct effect§	81.4	83.1	84.8	82.4	78.7	82.3	77.0
Total indirect effect‖	18.6	16.9	15.2	17.6	21.3	17.7	23.0
Indirect effect via diabetes#	4.2	2.6	1.6	3.9	5.7	4.1	2.0
Indirect effect via AMI**	11.5	10.8	10.4	10.9	12.3	10.8	15.1
Indirect effect via the interaction between diabetes and AMI††	2.9	3.6	3.3	2.8	3.4	2.8	3.2
Mediators: Incidence of diabetes and stroke							
Direct effect	76.0	78.5	78.4	76.8	73.7	77.0	70.6
Total indirect effect	24.0	21.5	21.6	23.2	26.3	23.0	29.4
Indirect effect via diabetes	5.3	3.6	2.3	4.7	6.9	4.9	5.4
Indirect effect via stroke	13.9	15.4	16.7	13.9	13.5	13.5	18.6
Indirect effect via the interaction between diabetes and stroke	4.8	2.5	2.6	4.6	5.8	4.6	5.4
Mediators: Incidence of diabetes and cardiovascular events‡‡							
Direct effect	68.3	70.5	70.9	69.8	65.8	69.9	61.7
Total indirect effect	31.7	29.5	29.1	30.2	34.2	30.1	38.3
Indirect effect via diabetes	4.5	2.9	1.8	3.9	5.7	4.1	2.1
Indirect effect via cardiovascular events	22.8	23.9	24.6	22.1	23.4	21.8	31.5
Indirect effect via the interaction between diabetes and cardiovascular events	4.5	2.7	2.7	4.1	5.0	4.2	4.7

AMI: acute myocardial infarction; BMI: body mass index.

*Results from the main analysis were presented in Table 2.
†Aalen additive hazards models adjusted for the selected individual-level covariates measured at baseline (i.e., age, sex, marital status, education, immigration status, household income adequacy, smoking status, smoking pack years, type of drinker, daily consumption of total fruits and vegetables, physical activity, and BMI) and time-varying area-level variables (i.e., education, income, % of unemployment, % of immigrants, indicators for rural/urban and north/south).

‡These effects were estimated using a method of dynamic path analysis (Aalen et al., 2019). This approach estimates cumulative direct and indirect effects of $PM_{2.5}$ on non-accidental deaths over follow-up based on the additive hazards model and a sequential linear model for the mediator process. We were unable to perform bootstrapping to derive confidence intervals for the point estimates because of the computational challenges.
§The effect of $PM_{2.5}$ on non-accidental deaths not mediated by selected mediators.
‖The total effect of $PM_{2.5}$ on non-accidental deaths mediated by selected mediators.
#The effect mediated through the first mediator alone.
**The effect mediated through the second mediator alone.
††The effect mediated through the interaction between the two mediators.
‡‡A composite indicator of incident AMI and stroke (whichever occurred first).

Table S5. The association between PM2.5 and deaths derived from the Aalen additive model and Cox PH model without adjusting for mediators

	The association between PM2.5* and non-accidental death			
	Aalen additive model		Cox PH model	
	Rate difference (per 1,000 person-years)	95% CI (per 1,000 person-years)	Hazard ratio	95% CI
The DM-CVD cohort†	1.79	0.93-2.65	1.16	1.08-1.25
The cohort with prevalence cases added back	2.26	1.33-3.20	1.17	1.09-1.26

AMI: acute myocardial infarction; CVD: cardiovascular events (AMI or stroke); PM2.5: fine particulate matter.

*per 10 $\mu g/m^3$ increase.

†Excluding people who had prior diagnosis of diabetes and cardiovascular events (AMI or stroke).

Figure S1. The total effect of exposure to fine particulate matter ($PM_{2.5}$) on non-accidental all-cause deaths which is estimated as the sum of the direct effect and indirect effects through mediators